The Methods of Ethics

Henry Sidgwick

[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 bullyingly 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

Glossary

affection: 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 enoying 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 'indiffferent 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 I

Chapter 1: Introduction

1. The phrase 'method of ethics' here refers to any rational procedure by which we determine what individual human beings 'ought' to do or what it is 'right' for them to do, or to try to bring about by voluntary action.[1] This distinguishes the study of ethics from that of politics, which concerns the right public conduct not of •individuals but of •societies. I see both ethics and politics differing from positive [see Glossary] sciences by having as their primary concern what •ought to be, not merely what is or has been or will be.

The student of ethics pursues *systematic and precise* general knowledge of what ought to be, and in this sense his aims and methods can be called 'scientific'; but I call ethics a 'study' rather than a 'science' because many people think that a science must be concerned with some part of what actually exists. . . . It's true that the two kinds of study—the positive and the practical [see Glossary]—are very closely interrelated. On any theory, our view about what ought to be must be largely derived, in its details, from our knowledge of what is: we can't know how to achieve our ideal except by careful study of actual phenomena; and anyone asking himself 'What ought I to do?' should examine the answers that others have given to this question. Still it seems clear that

> the general laws or uniformities that explain the varieties of human conduct, and of men's sentiments and judgments regarding conduct

is an entirely different topic from

> the truths about which among these varieties of conduct is right and which of these divergent judgments valid.

It's the systematic study of the latter topic that constitutes the special and distinct aim of ethics and politics.

2. . . . .Ethics is sometimes seen as directed to •the true moral laws or rational precepts of conduct, and sometimes as an inquiry into •the ultimate end of reasonable human action—the good or 'true good' of man—and how to achieve it. Both views are familiar, and will be carefully considered; but the former looms larger in modern ethical thought, and is easier to connect with modern ethical systems generally. For the *good* investigated in ethics is limited to good that is at least partly attainable by human effort; so knowledge of the end is sought in order to discover what actions are right for achieving it. Thus however prominent the notion of an ultimate good may be in an ethical system, and however we interpret this notion, it won't be any use to us unless we can discover some rules telling us how to behave.

On the other hand, the idea of ethics as an investigation of man's 'ultimate good' and how to achieve it doesn't easily connect up with what I'll call the 'intuitional' view of morality—the view that conduct is right when it conforms to certain precepts or principles of duty [see Glossary] that are intuitively [see Glossary] known to be unconditionally binding. On this view, the conception of *ultimate good* is not necessarily of fundamental importance in settling what conduct is right. (It would be important, of course, if •man's ultimate good *was* •acting rightly or •having the sort of

[1] I'll discuss the exact relation of 'right 'and 'what ought to be' in chapter 3. I here assume that they are equivalent for most purposes.

character that right action produces. But that view of man's good isn't an essential part of the intuitional view of ethics; and I don't think it squares with the moral common sense of modern Christian communities. We commonly think that the complete notion of human good or well-being must include •the attainment of happiness as well as •the performance of duty; even if we hold that men shouldn't make doing their duty conditional on their knowing that it will favour their happiness, because (as Butler put it) 'the happiness of the world is the concern of the lord and proprietor of it'. For those who adopt this position, there may sometimes be no logical connection between •what men ought to take as the practically ultimate end of their action and standard of right conduct and •the conception of ultimate good for man; so that in such cases this latter conception wouldn't help in the methodical discovery of what conduct is right, however indispensable it might be to the completeness of an ethical system.

[Sidgwick now explains why he doesn't define *ethics* as the 'art of conduct'. He is using 'art' in a now somewhat outmoded sense that involves knowledge of the rules or procedures to achieve some result. On the intuitional view of morality, we know basic moral truths 'intuitively'—i.e. straight off, without appealing to rules and procedures. And Sidgwick plans to deal at length with the intuitional view of morality (it will be the topic of Book III), so he doesn't want to exclude it by his preliminary throat-clearing, and stands by his initial definition:] Ethics is the science or study of what is right or what ought to be ·the case·, so far as this depends on the voluntary actions of individuals.

3. If we accept this view of what ethics is, why is it commonly taken to consist to a large extent of psychological discussion about 'the nature of the moral faculty'? Why, in particular, have I thought it right to include some discussion of this ·psychological· kind in the present work? It isn't immediately obvious why this should belong to ethics, any more than discussion of the mathematical faculty belongs to mathematics or discussion of the faculty of sense-perception belongs to physics. Why don't we simply start with certain premises saying what ought to be done or sought, without considering the faculty by which we see their truth?

One answer is that the moralist has a practical aim: we want knowledge of right conduct in order to act on it. Now, we can't help believing things that we see to be true, but we can help performing actions that we see to be right or wise, and in fact we often do things that we know to be wrong or unwise. This forces us to notice that we contain irrational springs of action, conflicting with our knowledge and preventing its practical realisation [see Glossary]; and just because our practical judgments are connected so imperfectly with our will, we are driven to look for more precise knowledge of what that connection is.

But that's not all. Men never ask 'Why should I believe what I see to be true?' But they often ask 'Why should I do what I see to be right?' It's easy to reply that the question is futile: it could only be answered in terms of some other principle of right conduct, and the question would then arise about that, and so on. But the question is asked, widely and continually, so this demonstration of its futility isn't satisfying; we want also some explanation of its persistence.

Here's one explanation: we are moved to action not by moral judgment alone but also by desires and inclinations that operate independently of moral judgment; so the answer we really want to the question 'Why should I do it?' is one that doesn't merely prove •a certain action to be right but also arouses in us a predominant inclination to perform •it.

This explanation is indeed true for some minds in some moods. But I think that when someone seriously asks why he

should do anything, he commonly assumes that he can act in any way that is shown by argument to be reasonable, even if it's very different from what his non-rational inclinations point to. And most of us agree that any reasonable decision about how to act will involve

(a) moral principles, and

(b) the agent's inclination independently of moral considerations,

and that **(b)** is only one element among several that have to be considered, and usually not the most important one. But when we ask what **(a)** these principles are, we get

- a great variety of answers from the systems and basic formulae of moral philosophers, and
- *the same* variety in the ordinary practical reasoning of men generally.

Between these two groups there's a difference: the philosopher seeks unity of principle and consistency of method, at the risk of paradox; the unphilosophical man is apt to hold different principles at once, and to apply different methods in more or less confused combination. Perhaps this explains the persistence of the 'Why?'-question we have been looking at, ·and the persistent unsatisfactoriness of answers to it·: if implicit in the thought of the ordinary person there are different views about the ultimate reasonableness of conduct, and if the person doesn't bring them into clear relation to each other, no one answer to the 'Why?'-question will completely satisfy him because it will be given from only one of these points of view, always leaving room to ask the question from some other.

I'm convinced that this is the main explanation of the phenomenon; and the present work is structured accordingly. Of course if any reasonings lead to conflicting conclusions we can't regard them as valid; I assume as a basic postulate of ethics when two methods conflict, one of them must be modified or rejected. But I think it's fundamentally important at the outset of ethical inquiry to recognise that a variety of methods are at work in ordinary practical thought.

4. Then what are these different methods? What are the different practical principles that the common sense of mankind is *prima facie* prepared to accept as ultimate? This has to be answered with care, because it often happens that we •prescribe that this or that 'ought' to be done or aimed at without explicitly mentioning an ulterior end, while we •are tacitly presupposing some such end. It is obvious that such prescriptions are merely 'hypothetical imperatives' (Kant's phrase); they are addressed only to those who have already accepted the end.

[Sidgwick gives examples: you *ought to* do such-and-such if you want

- to produce a good picture,
- to make an elegant table,
- to get your health back,
- to be found socially acceptable,
- to be happy,

and so on. The last of these is connected with] many rules prescribing so-called 'duties to oneself'; it may be said that such rules are given on the assumption that a man regards his own happiness as an ultimate end; that if anyone doesn't have that as an end, he doesn't come within their scope; in short, that the 'ought' in all such formula is implicitly relative to an *optional* end.

But it seems to me that this account doesn't get to the bottom of the matter. We don't all look with mere indifference [see Glossary] on a man who declines to take the right means to achieve his own happiness simply because he doesn't care about happiness. Most men would disapprove of such a refusal, regarding it as irrational; thus implicitly endorsing Butler's statement that 'interest, one's own happiness, is

a manifest obligation'. In other words, they would think that a man *ought* to care for his own happiness, and here 'ought' is no longer relative: happiness now appears as an ultimate end, the pursuit of which—within any limits imposed by other duties—appears to be prescribed by reason 'categorically' [see Glossary] (as Kant would say), i.e. with no tacit assumption of a further end. It has been widely held by even orthodox moralists that all morality ultimately rests on the basis of 'reasonable self-love' (Butler's phrase), i.e. that its rules are binding on any individual only to the extent that it's in his over-all interests to obey them.

Still, common moral opinion certainly regards the duty or virtue of prudence as only a part of duty or virtue in general, and not the most important part. Common moral opinion recognises and teaches other fundamental rules, such as those of justice, good faith, and veracity. In its ordinary judgments on particular cases, common morality is inclined to treat these as binding without qualification, and without regard to consequences. And the intuitional view of ethics, in the ordinary form of it, explicitly and definitely maintains the 'categorical' version of such rules, doing this as a result of philosophical reflection. And it holds that acting virtuously, at least for the virtues I have just mentioned, consists in strict and unswerving conformity to such rules.

On the other hand, many utilitarians hold that all the rules of conduct that men prescribe to one another as moral rules are really, though in part unconsciously, prescribed as means to the general happiness of •mankind or of •the totality of sentient beings; and even more of them hold that such rules, however they may originate, are valid only to the extent that obeying them is conducive to general happiness. Later on I'll examine this contention with due care. Here I'll just say this: if the duty of aiming at the general happiness is thus taken to include all other duties—these being subordinate applications of it—we seem to have circled back to the notion of happiness as an ultimate end, categorically prescribed, except that now it is general happiness and not the private happiness of any individual. This is the view that I take of the utilitarian principle.

When we are investigating right conduct, considered in relation to the end of private or of general happiness, we don't have to assume that the end itself is determined or prescribed by reason; all we have to assume is that it is adopted as ultimate [i.e. not a means to some further end] and paramount [i.e. not open to challenge from any rival consideration that is equally or more morally weighty]. For if a man accepts any end as ultimate and paramount, he implicitly accepts as his 'method of ethics' whatever process of reasoning enables him to determine what actions are most conducive to this end. (See the last paragraph of chapter 3 [page 18].) In pursuing these matters, we shan't attend to *every* end that someone has in practice adopted as ultimate, subordinating everything else to it under the influence of a 'ruling passion'; every difference in ultimate ends generates some difference in the 'methods' of moral inquiry, so that if we tackled them all our task would be very complex and extensive. But if we confine ourselves to ends that ordinary common sense seems to accept as reasonable ultimate ends, our task will be of a manageable size, because this criterion will exclude much of what men in practice seem to regard as paramount. For example, many men sacrifice health, fortune and happiness to the achievement of *fame*, but I don't know of anyone who has deliberately maintained that it is reasonable for men to seek fame for its own sake. It commends itself to thoughtful people only as

- •a source of happiness for the person who gains it,
- •a sign of that person's moral or intellectual excellence, or

> •a testimony to some beneficial achievement by the person and an encouragement to him and to others to achieve more.

And in the last of those, the conception of *beneficial* would lead us again to happiness or excellence of human nature, because it is commonly thought that a man benefits others either by making them happier or by making them wiser and more virtuous.

Are these two the only ends that can be reasonably regarded as ultimate? I'll investigate that in chapter 9 [page 49] and III/14 [page 191]; but I'll say right away that *prima facie* the only ends with a strongly and widely supported claim to be regarded as rational ultimate ends are the two just mentioned, happiness and perfection or excellence of human nature. . . . The adoption of the happiness end leads us to two *prima facie* distinct methods, depending on whether the individual is to aim at happiness •for everyone or •for himself alone. No doubt a man *can* often best promote his own happiness by what he does and refrains from doing for the sake of others; but our ordinary notion of self-sacrifice implies that the actions that do most for general happiness don't—in this world, at least—always produce the greatest happiness of the agent.[1] And among those who hold that 'happiness is our being's end and aim' there's a basic difference of opinion about whose happiness it is ultimately reasonable to aim at. For to some it seems, in Bentham's words, that 'the constantly proper end of action on the part of any individual at the moment of action is his real greatest happiness from that moment to the end of his life', whereas others hold that reason's view is essentially universal, so that it can't be *reasonable* to take as an ultimate and paramount end the happiness of any one individual rather than that of any other,. . . .so that general happiness must be the 'true standard of right and wrong, in the field of morals' no less than of politics (Bentham again). One can of course adopt an intermediate end, aiming at the happiness of some limited group such as one's family or nation or race; but any such limitation seems arbitrary, and probably few would maintain it to be reasonable except as the best route to the general happiness or to one's own.

The case seems to be otherwise with excellence or perfection.[2] It might seem at first sight that the excellence aimed at could be taken either individually or universally; and it's conceivable that a man might think that he could best promote the excellence of others by sacrificing his own. But no moralist who takes excellence as an ultimate end has ever approved of such sacrifice, at least so far as moral excellence is concerned. . . . So when we are attending to the view that right conduct aims at the production of excellence, we don't have to look separately at two versions of it—•one focusing on the excellence of the individual and •the other on the excellence of the human community. Now, virtue is commonly conceived as the most valuable element of human excellence—one with no viable competitors—so any method that takes perfection or excellence of human nature as ultimate end will *prima facie* coincide to a great extent with the method based on what I called •the intuitional view; so I have decided to treat it as a special form of •this latter.[3] The two methods that take happiness to be ultimate will be distinguished here as 'egoistic hedonism' and 'universalistic

[1] For a full discussion of this question, see II/5 [page 75] and the concluding chapter of this work [page 241].

[2] In my usage, 'perfection' stands for the ideal complex of mental qualities that we admire and approve in human beings and 'excellence' stands for any approximation to the ideal that we actually find in human experience.

[3] See III/14 [page 191] where I explain why I give only a subordinate place to the conception of perfection as ultimate end.

hedonism'. The latter is what Bentham and his successors taught, now generally called 'utilitarianism', a usage I shall follow. It's hard to find a good one-word label for egoistic hedonism, and I shall often call it simply 'egoism'. . . .

5. . . . .I shall offer now an explanation not of •the nature and boundaries of ethics but rather of •the plan and purpose of this work.

There are several recognised ways of treating this subject, none of which I have chosen to adopt. **(i)** We can start with existing systems, and either •study them historically, tracing the changes in thought through the centuries, or •compare and classify them according to resemblances, or •criticise their internal coherence. **(ii)** Or we can try to add a new system, and claim after so many failures to have finally achieved the one true theory of the subject. . . . The present work does neither of these things. I shall try to define and unfold not one method of ethics, but several; and I shall look at them not

- •historically, as methods that have actually been adopted, but rather
- •as alternatives between which we have to choose when trying to construct a complete and consistent system of practical maxims.

. . . .Men commonly seem to guide themselves by a mixture of methods, disguised under ambiguities of language. Everyone gives some acceptance to the impulses or principles from which the methods arise, and to various claims about which ends are rational; but we also feel the need to harmonise these different elements because it's a postulate of practical reason that two conflicting rules of action can't both be reasonable. The result is usually either •a confused blending or •a forced and premature reconciliation of the different principles and methods. And these defects have turned up in systems framed by professional moralists. These writers have usually rushed to •synthesis without adequate •analysis, ·i.e. putting a system together without first carefully examining its parts·; because they have felt the practical demand for a synthesis more urgently than the theoretical need for analysis. This is one of the places where •practical considerations have hindered the development of the •theory of ethics; and in this area a more complete detachment of theory from practice might be desirable even for the sake of practice. A treatment that is a combination of the scientific and the hortatory [here = 'urging, recommending, scolding'] is apt to spoil both: the mixture is bewildering to the brain and not stimulating to the heart. Here as in other sciences it would be useful to draw as sharp a line as possible between the known and the unknown, because the clear indication of an unsolved problem is a step to its solution. Ethical treatises, however, have tended to keep the difficulties of the subject out of sight, either

- •unconsciously, from an unconscious belief that the questions the writer can't answer satisfactorily must be ones that oughtn't to be asked, or
- •consciously, so that he won't shake the sway of morality over the minds of his readers.

The latter precaution often defeats itself: the difficulties concealed in exposition are liable to reappear in controversy, where they are. . . .magnified for polemical purposes. And so we get on one hand •vague and hazy reconciliation, and on the other •loose and random exaggeration of discrepancies; and neither process dispels the original vagueness and ambiguity lurking in our basic moral notions. My one immediate purpose in this work is to eliminate or reduce this indefiniteness and confusion. So I shan't aim for a complete and final solution of the chief ethical difficulties that would convert •my exposition of various methods into •the development of a harmonious system. But I hope I'll

be helping the construction of such a system, because it seems easier to judge the mutual relations and conflicting claims of different modes of thought if one has first examined, fairly and rigorously, their logical consequences. Practical principles that we unhesitatingly assent to at first sight, ones involving only notions that are familiar and apparently clear, often look different and somewhat dubious when we look carefully into their consequences. It seems that most of the practical principles that have been seriously put forward are fairly satisfactory to common sense as long as they have the field to themselves; their basic assumptions are all ones that we're inclined to accept, and that to some extent govern our habitual conduct. When I am asked

- •Don't you think it is ultimately reasonable to seek pleasure and avoid pain for yourself?
- •Don't you have a moral sense?
- •Don't you intuitively pronounce some actions to be right and others wrong?
- •Don't you agree that general happiness is a paramount end?

I answer 'Yes' to each question. My difficulty begins when I have to choose between the different principles or inferences drawn from them. We accept that when they conflict we *have to* choose—that it's irrational to let sometimes one principle prevail and sometimes another—but the choice is a painful one, and before making it we should have the completest possible knowledge of each candidate.

My aim here is •to expound as clearly and fully as I can the different methods of ethics that I find implicit in our common moral reasoning, •to point out their inter-relations, and •to clarify what is going on where they seem to conflict. This will lead me to discuss the considerations that I think should be decisive in the choice of ethical first principles; but I shan't try to establish any such principles, or to supply a set of practical directions for conduct. I want to direct your attention to the processes of ethical thought rather than to their results; so I shan't identify any positive practical conclusion as something I accept, unless by way of illustration; and I shan't dogmatically decide any controverted points, except where the controversy seems to arise from lack of precision or clearness in the formulation of principles, or lack of consistency in reasoning.

Chapter 2: The relation of ethics to politics

1. I have spoken of ethics and politics as practical studies, both concerned with the determination of ends to be sought, or rules to be unconditionally obeyed. Before going on, I should sketch the inter-relations of these two studies, seen from the point of view of ethics.

In my introductory account of them, •ethics aims at determining what ought to be done by individuals, while •politics aims at determining what the government of a state ought to do and how it ought to be constituted.

This may seem to make politics a branch of ethics. All the actions of government are actions of individuals, and so are the politically significant actions of the governed; and it would seem that such actions ought to be justifiable on ethical principles. . . . But this argument is not decisive,

for by similar reasoning ethics would have to include all the liberal and industrial arts [see Glossary]. . . . It is an important part of every adult's moral duty to take care of his health,. . . .but we don't consider ethics to include the art of medicine.

The specially important connection between ethics and politics arises in a different way. It is the business of government to lay down and enforce laws that regulate the outward conduct of the governed—their conduct in *all* their social relations so far as such conduct is a proper subject for coercive rules. ·Morality comes into this in two ways·. **(a)** This regulation ought to be in harmony with morality, for obviously people ought not to be compelled to do what they ought not to do. **(b)** To an important extent the law of a man's state will determine the details of his moral duty, even beyond the sphere of legal enforcement. [Sidgwick doesn't present the ensuing example in first-person terms.] For example: we commonly regard it as an individual's moral duty to 'give to every man his own', and this should govern my behaviour towards you. It may happen that you can't legally enforce your right to 'your own', but when I am thinking about what counts as your own I ought—we generally think—to be guided by the law of the state. If that changed, my moral duty would change with it. Similarly, the mutual moral duties of husbands and wives, and of children and parents, will vary in detail with the variations in their legal relations.

But when we look more closely into these matters, we find a need to distinguish •actual or positive [see Glossary] law from •ideal law, i.e. law as it ought to be. Political theory lays down principles for ideal law; but what primarily determines right conduct for an individual here and now is positive law, actual law. If positive law seems to me to diverge widely from ideal law—e.g. if I'm convinced by political theory that the law of property should be fundamentally changed—this will influence my view of my moral duty under the existing law; but the extent of this influence is vague and uncertain. Suppose I'm a slave-owner in a society where slavery is established, and I become convinced that private property in human beings should be abolished by law; it doesn't follow that I'll regard it as my moral duty to set my slaves free at once. [At this point Sidgwick switches from •freeing my slaves to •working for the freeing of all slaves. The switch is his, not an artifact of this version.] I may think that there's no hope of immediate general abolition of slavery, and even that it wouldn't work well for the slaves themselves, who require a gradual education for freedom; so that it would be better for the present •to aim at legal changes removing the worst evils of slavery, and •to set an example of humane treatment of bondsmen. Similar reasonings might be applied to the abolition of •private ownership of the instruments of production or of •appointments to positions in the government or the church. How far should political ideals influence moral duty? That seems to depend on

- •how far one seems to be from achieving the ideal,
- •how pressing the need for it is, and
- •how satisfactory the immediate realisation of the ideal would be.

The force attached to these considerations is likely to vary with the political method adopted; so that it's for politics rather than ethics to determine them more precisely.

So we have to distinguish clearly between two questions regarding the determination of right conduct for an individual here and now:

(a) How far should it be influenced by positive laws, and by other commands of government as actually established?

(b) How far should it to be influenced by political theory concerning the functions and structure of government

as it ought to be?

As regards question **(a)**: it is clearly up to ethics to determine the grounds and limits of obedience to government, and to develop a general conception of political duty that goes beyond mere obedience and duly recognizes the large variations due to the varying political conditions of different states. (A 'good citizen' in the United States will reasonably form a conception of his actual political duty that differs widely from that of a good citizen in Russia.) This will be the primary business of ethics on the political side of life. The discussion of political ideals will come into the picture only in a more indefinite and indirect way, reflecting the fact that such ideals are sure to have *some* influence on the determination of political duty under existing conditions.

2. Some thinkers take a view of ethical theory that gives it a relation to political theory quite different from the one I have presented. They hold that theoretical or 'absolute' ethics ought to investigate not •what ought to be done here and now, but •what ought to be the rules of behaviour in a society of ideally perfect human beings. This makes the subject-matter of our study doubly ideal, prescribing what ought to be done in a society that ought to exist. . . .

Those who take this view[1] adduce the analogy of geometry to show that ethics ought to deal with ideally perfect human relations just as geometry treats of ideally straight lines and perfect circles. But the irregular lines we meet with in experience have spatial relations that geometry doesn't entirely ignore; it ascertains them with enough accuracy for practical purposes. Another example: astronomy would be an easier study if the planets moved in circles, as was once believed; but the fact that they move in ellipses, and irregular ellipses at that, doesn't take them out of the sphere of scientific investigation; we have learned how to calculate even these more complicated motions. It may be useful for teaching purposes to assume that the planets move in perfect ellipses; but what we as astronomers want to know is the **actual** motion of the planets and its causes; and similarly as moralists we naturally ask what ought to be done in the **actual** world we live in. Our general reasonings—in astronomy or in ethics—can't possibly capture the full complexity of the actual considerations; but we try to approximate to it as closely as we can. That is the only way we can get to grips with the question to which mankind generally require an answer: 'What is a man's duty *in his present condition*?'. . . .

The inquiry into the morality of an ideal society might be conducted as a preliminary investigation, to be followed by the step from the ideal to the actual. How desirable is such a preliminary construction? The different methods of ethics answer this differently. Intuitionists generally hold that true morality—as far as determinate duties are concerned—prescribes absolutely what is right in itself and under all social conditions: truth should always be spoken and promises kept, justice should be done 'though the sky should fall', and so on. From this point of view, the general definitions of duties won't require any basic distinction between •the actual state of society and •an ideal state—for example, justice will be the same for both and will be equally stringent in both. Still, even an extreme intuitionist would admit that the *details* of justice and other duties will vary with social institutions; and it's plausible to suggest that getting a clear view of the 'absolute' justice of an ideal community would help us to achieve the merely 'relative'

[1] [Sidgwick has here a long footnote explaining that the present section was primarily aimed at the theory of ethics in Herbert Spencer's *Social Statics*; that Spencer replied to the section (which was also in Sidgwick's earlier editions); and that in this reply Spencer quietly moved to a much less extravagantly 'ideal' position.]

justice that is all we can have under existing conditions. How far this plausible suggestion is true is something we'll be better placed to judge when we have examined the definition of justice from an intuitional point of view.

For the method that proposes universal happiness as an ultimate end and supreme standard, the question is simpler: In our efforts to promote human happiness here and now, how much are we likely to be helped by systematically considering the social relations of an ideally happy group of human beings? I shan't deny here that this approach might be useful, but it's easy to show that it is involved in serious difficulties.

Just as in ordinary deliberation, so also when thinking about the ideal society, we have to consider what is best under certain conditions of human mental or physical life. We need to focus our thought less on

- the desired end, namely the most pleasant consciousness conceivable, lasting as long and as continuously as possible, than on
- some method by which human beings might achieve it under conditions not too remote from our own, so that we can at least try to imitate them.

So we have to know how far our actual circumstances are modifiable; a difficult question, as we can see from the ideal societies that have actually been invented! For example, Plato's Republic differs a good deal from reality, and yet he thinks of *war* as a permanent unalterable fact that has to be allowed for in the ideal state;. . . .yet any modern Utopia—even one that wasn't at all bold or flashy—would include the suppression of war. Indeed, two thinkers constructing ideal states may head out from the actual in diametrically opposite directions. For example, permanent marriage-unions now cause some unhappiness because conjugal affection isn't always permanent, but they are thought to be necessary, partly to protect men and women from harmful ups and downs of passion, but chiefly to secure good upbringings for children. There are two ways an 'ideal state' theorist could go from there:

(a) In an ideal state of society we could trust more to parental affections, needing less control over the natural play of emotion between the sexes; so that 'free love' is therefore the ideal.

(b) Permanence in conjugal affection is natural and normal, and any exceptions to it will disappear as we approach the ideal state.

Another example: in our actual society over-all happiness is lessened by unequal distribution of the means of happiness, and the division between rich and poor. We can think of this as remedied

(a) by the rich becoming more willing to redistribute their share, or

(b) by enabling the poor to secure more for themselves.

The voluntary and casual almsgiving that now goes on would be increased in **(a)**, extinguished in **(b)** .

When we abandon the firm ground of actual society, then, we enter an illimitable cloudland, in which we can construct •any variety of pattern states but •no definite ideal to which the actual world undeniably approximates, as the straight lines and circles of the physical world approximate to those of scientific geometry.

You may think that we can reduce this variety by studying mankind's past history, as a basis for predicting to some extent their future manner of existence. [Sidgwick says that this won't help us to know what to do *now*. Even supposing we get excellent evidence that mankind has been following definite lines of progress: this may help us to see what is likely to come next, but it tells us nothing about what line of further development would be a move closer to an ideal

state of society. He adds in a footnote:] This question will be further discussed in chapter IV/2.2 [page 227].

* * * * *

[If you consult a printed copy of this work you won't find the next four paragraphs in this position. See the note on page 223 for an explanation.]

If we consider the relation of ethics to politics from a utilitarian point of view, the question *What rules of conduct for the governed should be fixed by legislators and applied by judges?* will be answered in the same way as are all questions of private morality, namely by predictions regarding consequences—attempts to estimate and balance against each other the effects of such rules on the general happiness. But if we divide utilitarianism into two parts—•the theory of private conduct and •the theory of legislation—and ask which of these two is prior, we seem to get different anwers for different parts of the legal code.

(i) To a large extent the rules laid down in a utilitarian code of law will be ones that anyone sincerely wanting to promote general happiness would generally try to follow, even if they weren't legally binding. Examples of this include:

- •the rule of not inflicting any bodily harm or needless annoyance on anyone, except in self-defence or as retribution for wrong;
- •the rule of not interfering with anyone else's pursuit of the means of happiness, or with his enjoyment of wealth acquired by his own labour or the free consent of others;
- •the rule of fulfilling all engagements freely entered into with any person x unless the fulfilment would be harmful to others or much more harmful to oneself than beneficial to x, or there were good grounds for supposing that x would not perform his share of a bilateral contract; and
- •the rule of supporting one's children while they are helpless and one's parents if they are decrepit, and of educating one's children suitably to their future life.

As regards rules like these, utilitarian ethics seems independent of politics, and naturally prior to it; we *first* consider what conduct is right for private individuals, and *then* consider how much of this they can advantageously be compelled to by legal penalties.

(ii) There are other rules that would serve general happiness only if everyone was forced to follow them; for example, •refraining from personal retaliation for injuries, and •a more general and unhesitating fulfilment of contracts than would be expedient [see Glossary] if they were not legally enforced.

(iii) When it comes to settling all the claims that the members of society have against one another,

- •the great differences in relevant facts about different situations imply that on many points the utilitarian theory of right private conduct apart from law would lead to different answers in different cases,

while at the same time

- •uniformity is either indispensable, to prevent disputes and disappointments, or at least highly desirable so as to maintain rules of conduct that are usually though not always expedient.

Examples of this are •exact fixing of the limits of copyright in literary compositions and patents of technical inventions; and •a large part of the law of inheritance, and of •the law regulating family relations. In such cases. . . .utilitarian ethics seems to blend with utilitarian politics in a rather complicated way; because we cannot determine the right conduct for a private individual in any particular case without first considering what rule (if any) it would be on the whole expedient to maintain, in the society of which he is a member, by legal penalties as well as by the weaker and less

definite sanctions of moral opinion. In any paticular case this problem is further complicated by the delicate relations between positive law and positive morality (as we may call the actual moral opinions generally held in a given society at a given time). On one hand, it is dangerous for legislation to go beyond positive morality by prohibiting conduct that is generally approved or tolerated; on the other hand, up to the point at which this danger becomes serious, legislation is an effective instrument for modifying or intensifying public opinion in a desirable direction. Leaving this difficult question of social dynamics, we may say that normally in a well-organised society the most important and indispensable rules of social behaviour will be legally enforced and the less important left to be maintained by positive morality. Law will constitute, as it were, the skeleton of social order, clothed in the flesh and blood of morality.

Chapter 3: Ethical judgments

1. I have spoken of actions that we judge to be right and that ought to be done as being 'reasonable' or 'rational', and of ultimate ends as 'prescribed by reason'; and I contrasted the corresponding motives with 'non-rational' desires and inclinations. This way of speaking is employed by writers of various schools, and fits with ordinary language and ordinary thought. We do commonly think that wrong conduct can be shown by argument to be essentially irrational. We don't think that men are influenced to act rightly by reason alone, but we hold that appeals to reason are an essential part of all moral persuasion—the part that concerns the moralist or moral philosopher as distinct from the preacher or moral rhetorician. But many people think that, as Hume says, 'Reason, meaning the judgment of truth and falsehood, can never of itself be any motive to the will', and that every motive to action is some non-rational desire, taking 'desires' to include the impulses to action given by present pleasure and pain. Before going further, we should look carefully at the grounds of this contention.

Let us first define the issue as clearly as we can. Most of us have felt •a bodily appetite [see Glossary] prompting us to indulgences that we judge to be imprudent, and •anger prompting us to acts that we disapprove of as unjust or unkind. In such a conflict the desires are said to be irrational because they drive us to volitions that are opposed to our deliberate judgments; and it's when we succeed in resisting such irrational desires that their impulsive force is most definitely felt, because in resisting them we have make a voluntary effort somewhat like that of muscular exertion. Desires of this kind are often at work when we aren't giving any thought to our duty or our interests, as when an ordinary healthy man eats his dinner. It seems best to call these desires 'non-rational' rather than 'irrational'. Despite these labels, the more important of these desires are normally accompanied by intellectual processes. Some impulses to action do indeed seem to operate 'blindly' or 'instinctively', with no definite thought about either •the end the action aims at or •the means by which it is to be

attained; but this happens only with impulses that •don't occupy consciousness for long and •require only very familiar and habitual actions to achieve their immediate ends. In ethical discussion we are chiefly concerned not with those cases but with ones where the intended result and at least some of the means to it are more or less clearly represented in consciousness before the volition that starts the action. So the resultant forces of 'nonrational' desires and the volitions prompted by them are continually modified by intellectual processes in two ways; •by new perceptions or representations of means to the desired ends, and •by new presentations or representations of relevant facts, especially ones about probable consequences of the contemplated action.

Is that an exhaustive account of the influence of the intellect on desire and volition? Is the so-called 'conflict of desire with reason' really just a conflict among desires and aversions, with reason's only role in it being to confront the mind with ideas of actual or possible facts that modify the resultant force of our various impulses?

I say No: the ordinary moral or prudential judgments that have at least some influence on volition in most minds aren't judgments concerning •present or future human feelings or •any facts about the sensible world; •the basic notion of *ought* or *right* which such judgments involve is **essentially different from** any notion representing facts of physical or mental experience. To support this claim I shall have to appeal to you to reflect on your own practical judgments and reasonings; and the best start to this appeal is to criticise all attempts to explain the practical judgments that employ •this basic notion without recognising its unique character as above **negatively** defined. Such explanations have an element of truth: they highlight feelings that undoubtedly *accompany* moral or prudential judgments and ordinarily have some effect on the person's willingness to do what he judges to be right; but considered as interpretations of what such judgments *mean* they are complete failures.

In this context we have to consider 'moral' judgments separately from 'prudential' ones. As I have pointed out, there's a strongly supported opinion that all valid moral rules ultimately have a prudential basis; but in ordinary thought we sharply distinguish •judgments of duty from •judgments as to what 'is right' or 'ought to be done' in view of the agent's private interests or happiness; and the depth of the distinction won't be lessened by the closer examination of these judgments that we are now to conduct.

[Sidgwick acknowledges that we do sometimes use 'right' to mean 'the best or only way to get x' and similarly with 'ought'. He continues:] But it seems clear **(1)** that certain kinds of actions under the names 'justice', 'veracity', 'good faith' etc. are commonly held to be right •unconditionally and not merely •right-if-you-want-to-achieve-x; and **(2)** that we similarly regard as 'right' the adoption of certain ends such as the common good of society. In either of these cases the above ·means-to-an-end· interpretation is clearly wrong.

So we have to find a meaning for 'right' or 'what ought to be' other than the notion of fitness for some end. Here's a proposal that has been made:

> The judgments that we commonly call 'moral' in the narrower sense really affirm only that the speaker has a specific emotion. When I say 'Truth ought to be spoken' or 'Truthspeaking is right', I mean only that the idea of truthspeaking arouses in my mind a feeling of approval or satisfaction.

It's probably true that most moral judgments on real cases are *accompanied by* some degree of such emotion, i.e. of so-called 'moral sentiment'; but it's absurd to say that what the proposition 'Truth ought to be spoken' *means* is that the

speaker has a feeling of approval of truthspeaking. If that were what it meant, this could happen:

> One man says 'Truth ought to be spoken', another says 'Truth ought not to be spoken', and they are both right!

This is so *obviously* absurd that we must suppose that no-one has really meant to maintain it, and that the thesis we are up against here is not the one I have just exposed but rather this:

> The subjective fact of my approval is all that there's any evidence for;

or perhaps this:

> The subjective fact of my approval is all that any reasonable person is prepared on reflection to affirm.

There certainly are many statements, objective in their form, that we usually aren't willing to defend as more than subjective if their validity is questioned. If I say 'The air is sweet' or 'The food disagreeable', it's not the case that all I mean is that I like the air or dislike the food; but if my statement is challenged, I'll probably settle for reporting the existence of such feelings in my mind. But this case differs fundamentally from the case of moral feelings. The unique emotion of moral approval is, in my experience, inseparably bound up with the implicit or explicit conviction that the conduct approved is 'really' right, i.e. that anyone who disapproves of it is in error. If for any reason I give up this conviction, I may still have a feeling prompting me to the conduct in question, or (perhaps more often) a feeling of repugnance to the opposite conduct; but this feeling will no longer have the special quality of 'moral sentiment' strictly so-called. . . . Take the case of a man who has been habitually influenced by a moral sentiment in favour of veracity, but becomes convinced that in his present special circumstances speaking the truth is not right but wrong. He will probably still feel a revulsion against violating the rule of truthspeaking; but this will be different in kind and degree from the feeling that prompted him to veracity as virtuous. We might call the one a 'moral' and the other a 'quasi-moral' sentiment.

That argument holds equally against this:

> Approval or disapproval is not the mere •liking or aversion of an individual for certain kinds of conduct, but •this mixed in with a sympathetic [see Glossary] representation of similar likings or aversions felt by others.

No doubt such sympathy normally comes along with moral emotion; and when it doesn't, it is harder to maintain the moral position. But that is partly because our moral beliefs commonly agree with those of others in our society, and our confidence in the truth of these beliefs depends greatly on this agreement. But suppose that we are led by argument to a new moral belief, opposed to our own habitual moral sentiment and to that of the society we live in; this is a crucial experiment [see Glossary] that proves the existence in us of 'moral sentiments' as I have defined them, colliding with the sympathies of our fellow-men as much as with our own mere likings and aversions. And even if we imagine that the whole human race has sympathies opposed to our newly-acquired convictions, so that we see ourselves as standing against the world, still, so long as our conviction of duty is firm, our moral emotion stands out as quite distinct from the complex sympathy opposed to it, however much we extend, complicate and intensify the latter.

2. . . . .There's another account of 'ought' in which the likings and aversions that people generally have for certain kinds of conduct enter the picture in a different way: on this account, when x says that y 'ought to' do A he is thinking of the general aversion to A as a basis for holding that if y does A he will *suffer* directly or indirectly from the dislike of his

fellow-creatures.

This interpretation expresses a part of what 'ought' and 'duty' mean in ordinary thought and discourse. When we are talking about someone's *duty* or what he *ought* to do, we often express this by saying that he has a 'moral obligation' to do it, which suggests an analogy between this notion and that of *legal* obligation; and in the case of positive [see Glossary] law there's an essential connection between 'obligation' and punishment: a law isn't established in a society if those who break it are never punished. But further thought shows that •the use of 'ought' that fits this account, though it really does occur, must be distinguished from •the special ethical use of the term. In common thought the conceptual distinction between legal rules and merely moral ones lies in just this connection of punishment with legal rules and *not* with moral ones. We think there are some things that a man ought to be compelled to do or refrain from, and others that he ought to do or refrain from without compulsion, and that only the former lie within the sphere of ·positive· law. No doubt we also think that in many cases where the compulsion of law is undesirable, the fear of moral censure and its consequences supplies a normally useful constraint on the will of any individual. But when we say that a man is 'morally though not legally bound' to do A, we don't mean merely that he'll be punished by public opinion if he doesn't do A; for we often say things of the form 'He ought to do A *and* he'll be punished by public opinion if he doesn't', meaning this as two statements, not one. Also, public opinion is fallible and we often judge that men 'ought' to do A while we're perfectly aware that they won't pay much of a social price if they don't. . . .

Admittedly we quite often make judgments that sound like moral judgments in form, and aren't distinguished from them in ordinary thought, though on reflection we realise that they really depend on the existence of current public attitudes. Modern civilised societies have codes of public opinion, enforced by social penalties, which no thoughtful person •thinks are moral codes or •regards as unconditionally binding. Any such code varies through time, and at a single time is different for different classes, professions, social circles. Such a code always supports to a considerable extent the commonly received code of morality; and most thoughtful people think it generally reasonable, for prudential or moral reasons, to conform to the dictates of public opinion—that is

- •to the 'code of honour', as it may be called in serious matters, or
- •to 'the rules of politeness or good breeding' in lighter matters

—whenever these don't positively conflict with morality. But less thoughtful people don't distinguish the duties imposed by social opinion from moral duties; and the common meaning of many terms captures this failure-to-distinguish. If we say that a man has been 'dishonoured' by a cowardly act, it's not quite clear whether we mean that he has *incurred* contempt or that he has *deserved* it, or both; and this becomes evident when the code of honour conflicts with morality. Take the case of a man who refuses a duel on religious grounds: some would say that he was 'dishonoured' but had acted rightly; others would say that there couldn't be real dishonour in a virtuous act. . . .

Another way of interpreting 'ought' as involving penalties is less easy to meet by a crucial psychological experiment. The moral imperative may be taken to be a law of God, who will punish breaches of it. . . . But this belief is not shared by everyone whose conduct is influenced by independent moral convictions that may not be supported by •the law of the land or •the public opinion of their community. And even for

those who *do* fully believe in the moral government of the world, the judgment **(i)** 'I ought to do A' can't be identified with the judgment **(ii)** 'God will punish me if I don't do A', because believing **(i)** is clearly seen to be an important part of the grounds for believing **(ii)**. Also, when Christians say that God's 'justice' (or any other moral attribute) is shown in punishing sinners and rewarding the righteous, they obviously mean not merely that God *will* thus punish and reward but that it is 'right 'for him to do so; and of course they can't mean that he will be punished if he doesn't!

3. So the notion of *ought* or *moral obligation* as used in our common moral judgments doesn't merely mean something about emotions, the speaker's and/or those of others; and doesn't mean something about penalties, whether social or divine. Then what does it mean? What definition can we give of 'ought', 'right', and other terms expressing the same basic notion? My answer is that the notion in question is too elementary, too simple, to be capable of any formal definition. I'm not implying that it belongs to the mind's original constitution, i.e. that its presence in consciousness is not the result of a process of development. I'm sure that the whole structure of human thought, including the most simple and elementary conceptions, has grown through a gradual mental process out of some lower life in which thought, properly so-called, had no place. But it doesn't follow that no notion is really simple though some appear to be so. Water results from hydrogen and oxygen in such a way that it has these elements *in* it, ·which means that it isn't simple·; but I don't know any reason for transferring this pattern from chemistry to psychology, maintaining that mental items contain *as parts* the mental items out of which they grew.[1] In the absence of such reasons, a psychologist must accept as ·simple or· elementary anything that careful introspection declares to be so. . . . The *ought* notion that we are dealing with here can be made clearer in only one way, namely by determining its relation to •other notions with which it is connected in ordinary thought, especially to •notions with which it is liable to be confused.

We need to distinguish two senses that the word 'ought' can have: **(i)** When we judge that A 'ought to be' done, in the narrowest ethical sense, we're thinking of A as something that can be brought about by the volition of anyone who 'ought' to do it. I can't conceive that I 'ought' to do something that I don't think I *can* do. **(ii)** But there's also a useful place for a wider sense that is at work when, for example, I judge that I 'ought' to know what a wiser man would know, or to feel as a better man would feel in my place, while knowing that I couldn't directly acquire any such knowledge or feeling by any effort of will. In this use, the word merely implies an ideal or pattern which I 'ought' (now in the stricter sense) to imitate as far as possible. And the word is normally being used in this wider sense. . . .in political judgments, as when I judge that the laws of my country 'ought to be' other than they are. I don't of course mean that anyone's individual volition can directly bring about the change, or even that any group of individuals could produce *all* the changes that I think ought to occur; but my judgment points to a pattern to which my country could approximate. In each sense I imply that •what ought to be is •something that can be known, i.e. that what I judge ought to be must—unless I am in error—be similarly judged by all rational beings who judge truly of the matter. In the present work 'ought' will always be used in the stricter sense except where the context clearly points to

[1] In Chemistry the compound weighs exactly as much as the elements making it up, and usually we can dismantle the compound and get the separate elements back. There's no analogue to any of this in the relation between a mental item and the 'elements' it has grown from.

the wider sense—that of the political 'ought'.

In referring such judgments to 'reason', I'm not prejudging the question of whether valid moral judgments are normally reached by •a process of reasoning from universal principles or axioms, or by •direct intuition of the particular duties of individuals. Many people hold that our moral faculty deals primarily with individual cases as they arise, applying the general notion of *duty* to each case and deciding intuitively what this person ought to do in these circumstances. On this view, the grasping of moral truth resembles sense-perception more than it does rational intuition as this is commonly understood,[1] so that the label 'moral *sense*' might seem more appropriate. But 'sense' suggests a capacity for feelings that may vary from one person to another without either being in error, rather than a faculty of cognition;[2] and I think it's very important to avoid this suggestion. So I think it is better to use the term 'reason', with the explanation given above, to name the faculty of moral cognition; adding, as a further justification of this use, that even when a moral judgment relates primarily to some particular action we commonly regard it as applicable to any other action belonging to a certain definable class; so that the moral truth apprehended is implicitly conceived to be intrinsically universal, though particular in our first apprehension of it. (A further justification for this extended use of the term 'reason' will be suggested in chapter 8.3 [page 46].)

Also, when I speak of the cognition or judgment that 'A ought to be done ' (in the stricter ethical sense of 'ought') as a 'dictate' or 'precept' of reason to the persons to whom it relates, I imply that in any rational being this cognition would give an impulse or motive to action; though for human beings this is only one motive—often not a predominant one—among others that may conflict with it. In fact, this possibility of conflict is conveyed by the words 'dictate' and 'imperative', which likens

•how reason relates to non-rational impulses etc.

to

•how the will of a superior relates to the wills of his subordinates.

This conflict seems also to be conveyed by the terms 'ought', 'duty' and 'moral obligation', as used in ordinary moral discourse; so that these aren't applicable to the actions of rational beings who don't have impulses conflicting with reason. But we can say of such beings that their actions are 'reasonable' or (in an absolute sense) 'right'.

4. Some people will want to answer the whole of my line of thought by denying that they can find in their consciousness any such unconditional or categorical imperative as I have been trying to exhibit. If this really is the final result of self-examination for any person, there's no more to be said. I, at least, don't know how to convey the notion of moral obligation to someone who is entirely devoid of it. But I think that many of those who give this denial really mean only to deny that they have any consciousness of *moral obligation to actions without reference to their consequences*, and wouldn't deny that they recognise some universal end—e.g. universal happiness or well-being—as being what it's ultimately reasonable to aim at, giving this aim preference over any personal desires that conflict with it. But this view (as I said before) clearly involves the unconditional

[1] We don't say that particular physical facts are grasped by reason; we consider •discursively used reason as dealing with relations among judgments or propositions, and •intuitive reason as restricted to the seeing of universal truths such as the axioms of logic and mathematics.

[2] By 'cognition' I always mean what some would call 'apparent cognition'; i.e. I don't mean to affirm the validity of the cognition, but only its existence as a mental fact and its *claim* to be valid.

imperative regarding the end, and recognises an obligation to perform the acts that most conduce to it. The obligation isn't 'unconditional', but it's not conditional on any non-rational desires or aversions. And nothing I've been saying is meant as an argument for intuitionism against utilitarianism or any other method that treats moral rules as relative to general good or well-being. For instance, nothing I've said is inconsistent with the view that truthspeaking is valuable only as a means to the preservation of society; but then the preservation of society (or some further end to which it is a means) must be intrinsically valuable and therefore something that a rational being ought to aim at. . . .

And even those who hold that moral rules are obligatory only because it is in the individual's interest to conform to them. . . .don't get rid of the 'dictate of reason' if they think that private ·self·-interest or happiness is an end that it's ultimately reasonable to aim at. . . . Kant and others maintain that it can't be a man's duty to promote his own happiness, because 'what everyone inevitably wills cannot be brought under the notion of duty'. But even if we grant (and I'll show in chapter 4 that I don't) that it's in some sense true that a man's volition always aims at attaining his own happiness, it doesn't follow that a man always does what he thinks will lead to his own *greatest* happiness. As Butler emphasized, we are all familiar with cases where someone gives way to some appetite or passion while knowing that this is clearly opposed to what they conceive to be their interests. So the notion *ought*—as expressing how rational judgment relates to non-rational impulses—will find a place in the practical rules of any egoistic system just as it does in the rules of ordinary morality where it prescribes duty without reference to the agent's interests.

This may be maintained:

> Egoism doesn't regard the agent's greatest happiness as what he *ought* to aim at, but only as what he predominantly *wants*. This desire may be temporarily overcome by passions and appetites, but ordinarily it regains the upper hand when these passing impulses have spent their force.

I know that many people take this view of egoistic action, and I'll consider it in chapter 9. But even if we hold that *no* end of action is unconditionally or 'categorically' prescribed by reason, this won't remove the *ought* notion from our practical reasonings; it still remains in the 'hypothetical imperative' prescribing the best means to any end selected end. When a physician says **(i)** 'If you wish to be healthy you ought to rise early', this is *not* the same thing as saying **(ii)** 'Early rising is essential for the attainment of health'. What **(ii)** does is to express the fact on which **(i)** is based, but the word 'ought' doesn't merely indicate this fact; it also implies that it's unreasonable to adopting a certain end while refusing to adopt the means needed for its attainment. Possible objection:

> It's not just unreasonable—it's impossible! Adopting an end means having a preponderant desire for it, and if aversion to the essential •means causes •them not to be followed although recognised as indispensable, the desire is no longer preponderant and stops being adopted.

This objection arises from a defective psychological analysis. What I find when I look into my own consciousness is that •adopting an end is a quite different mental phenomenon from •having a desire; it's a **(a)** kind of volition, though not of the **(c)** kind that initiates a particular immediate action. Intermediate between **(a)** and **(c)** is **(b)** a resolution to act in a certain way at some future time. Sometimes when the

time comes to act on such a resolution we act otherwise, under the influence of passion or habit, without consciously cancelling our previous resolve. Our practical reason condemns this inconsistency of will as irrational, quite apart from any judgment of approval or disapproval on either volition considered by itself. There is a similar inconsistency between the adoption of an end and a general refusal to take the means we see to be indispensable to achieving it; and if when the time comes we •don't follow those means yet also •don't consciously retract our adoption of the end, it can hardly be denied that we *ought* in consistency to act otherwise than we do. And we are all familiar with such contradictions between a general resolution and a particular volition.

Chapter 4: Pleasure and desire

1. I haven't described •the emotional characteristics of the impulse that prompts us to obey the dictates of reason. That is because •these seem to be different in different minds, and even in one mind at different times, without any change in the volitional direction of the impulse. The ruling impulse in the mind of a rational egoist is generally what Butler and Hutcheson call a **'calm' or 'cool' self-love**; whereas in someone who takes universal happiness as the end and standard of right conduct, the desire to do what he judges to be reasonable is usually combined with some degree of **sympathy and philanthropic enthusiasm**. Someone who thinks of the dictating reason (whatever its dictates may be) as external to himself, the thought of rightness is accompanied by a sentiment of **reverence for authority**; which some may think of as impersonal but more regard as the authority of a supreme Person, so that the sentiment. . . .becomes religious. This conception of reason as an external authority against which the self-will rebels is often irresistibly forced on the reflective mind; but at other times the identity of •reason and •self presents itself as an immediate conviction, and then reverence for authority passes over into **self-respect**; if we see the rational self as liable to be enslaved by the force of sensual impulses, the opposite and even stronger sentiment of **freedom** is called in. Quite different again are the emotions of **aspiration or admiration** aroused by the conception of virtue as an ideal of moral beauty (I'll return to this in chapter 9). . . . There are important differences in the moral value and efficacy of these different emotions; but their primary practical effect seems to remain the same as long as the cognition of *rightness* remains unchanged. The chief concern of ethics, in my view, is with these cognitions; it aims to free them from doubt and error, and to systematise them as far as possible.

But one view of the feelings that prompt us to voluntary action is sometimes thought to cut short all controversy over the principles that ought to govern such action. I mean the view that volition is always determined by pleasures or pains, actual or prospective. I call this doctrine 'psychological hedonism'; it is often connected with—and quite often confused with—the method of ethics that I have called 'egoistic hedonism'; and it does at first sight seem natural to think that if the psychological doctrine is true then the ethical one

must be also—if one end of action. . . .is definitely determined for me by unvarying psychological laws, a different end can't be prescribed for me by reason.

When you think about it, though, you'll see that this inference assumes that a man's pleasure and pain are determined independently of his moral judgments; whereas it's plainly possible that our expectation of pleasure from doing A may largely depend on whether we think that A is the right thing to do. And in fact psychological hedonism requires us to suppose that this is the case with people who habitually act in accordance with their moral convictions. . . .

So •the psychological thesis that pleasure or absence of pain to myself is always the actual ultimate end of my action has no necessary connection with •the ethical thesis that my own greatest happiness or pleasure is for me the *right* ultimate end. [Sidgwick turns to the version of the psychological thesis which says that each of us is psychologically determined to seek his own *greatest possible* pleasure (or least pain), and agrees with Bentham that anyone who believes this will have no room in his scheme of things for the ethical doctrine that Sidgwick calls 'egoistic hedonism', because:] a psychological law that my conduct invariably conforms to can't be conceived as 'dictate of reason'; this latter must be a rule from which I am conscious that I could deviate. [But, Sidgwick continues, 'writers who now maintain psychological hedonism' wouldn't accept Bentham's super-strong version of it. He quotes Mill as saying that people often fail to act towards their greatest happiness, this failure being due to 'infirmity of character'. So Sidgwick will now take psychological hedonism not in the strong Bentham form of it but in the weaker form in which 'greatest' doesn't occur.]

So egoistic hedonism becomes a possible ethical ideal to which psychological hedonism seems to point. If the ultimate aim of each of us in acting is always solely some pleasure (or absence of pain) to himself, there's a strong suggestion that each of us *ought* to seek his own *greatest* pleasure (or, strictly, greatest surplus of pleasure over pain). . . .—the mind has a natural tendency to pass from the one position to the other. . . . This psychological doctrine seems to conflict with an ethical view widely held by morally developed people, namely that for an act to be in the highest sense virtuous it mustn't be done solely for the sake of the pleasure it brings, even if it's the pleasure of the moral sense; that is, a truly virtuous act won't be done solely so as to get a glow of moral self-approval. [Sidgwick puzzlingly introduces that last sentence with 'Further. . . ', as though he were continuing or reinforcing an immediately preceding argument.]

So it seems important to subject psychological hedonism, even in its more indefinite ·non-Bentham· form, to a careful examination.

2. Let us start by defining the question at issue more precisely.

> **pleasure**: a kind of feeling that stimulates the will to actions tending to sustain it if it's actually present, and to produce it if it's only being thought about.
>
> **pain**: a kind of feeling that stimulates the will to actions tending to remove or avert it.[1]

It's convenient to call the felt volitional stimulus 'desire' in one case and 'aversion' in the other; though 'desire' is usually restricted to the impulse felt when pleasure is only thought about and not actually present.[2] So the question at issue is. . . .this:

[1] [A footnote here indicates that this account of pain isn't right as it stands, and will be returned to in II/22.2 .]

[2] [A footnote refers to the use of 'desire' in cases where the person knows that the desired state of affairs can't possibly be achieved. Sidgwick will set this aside, he says, and mentions it only because the psychologist Alexander Bain so defines 'desire' that *only* the forlorn-hope cases are desires.]

> Are there desires and aversions that *don't* have pleasures and pains for their objects, i.e. conscious impulses to produce or avert results other than the agent's own feelings?

Mill explains that '•desiring a thing and •finding it pleasant are two ways of naming the same psychological fact'. If that were right, the question we are asking can be answered without resorting to the 'practised self-consciousness and self-observation' ·that Mill invokes in his *Utilitarianism*·, because the answer Yes would involve a contradiction in terms. The discussion of this question has been confused by an ambiguity in 'pleasure'. When we speak of a man doing something 'at his pleasure' or 'as he pleases', we usually mean only that this was a voluntary choice, not that he was aiming at some feeling for himself. Now, if by 'pleasant' we mean merely what influences choice, exercises an attractive force on the will, then the statement *We desire what is pleasant* or even *We desire a thing in proportion as it appears pleasant* is perfectly safe because it is tautological [see Glossary]. But if we take 'pleasure' to denote the kind of feeling defined above, then there's a real *question* as to whether the end to which our desires are always consciously directed is our coming to have such feelings. And *this* is what we must understand Mill to regard as 'so obvious that it will hardly be disputed'.

It is rather curious to find one of the best-known English moralists supporting the exact opposite of what Mill thinks to be so obvious, supporting it not merely as a universal fact of our conscious experience but even as a necessary truth! Butler distinguishes self-love, i.e. the impulse towards our own pleasure, from 'particular movements towards particular external objects—honour, power, the harm or good of another'—which lead to actions that 'aren't ·self·-interested except in the ·debased· sense in which every action of every creature *must* be self-interested because no-one can act except from a desire or choice or preference of his own'. Such particular passions or appetites are, he goes on to say, '*necessarily presupposed by the very idea* of a ·self-·interested pursuit; because the very idea of interest or happiness consists in the success of some appetite or affection [see Glossary] in getting what it aims at'. We couldn't pursue pleasure at all unless we had desires for something other than pleasure, because pleasure consists precisely in the satisfaction of these 'disinterested' impulses.

Butler has certainly over-stated his case, so far as my own experience goes; for many pleasures—especially those of sight, hearing and smell, together with many emotional pleasures—occur to me without any relation to previous desires, and it seems quite *conceivable* that our primary desires might all be directed towards pleasures like these. But as a matter of fact it appears to me that throughout the whole scale of my impulses—sensual, emotional, and intellectual—I can pick out desires for things other than my own pleasure.

Let's start with an illustration from the impulses commonly placed lowest in the scale. The appetite [see Glossary] of hunger strikes me as a direct impulse to eat food. Such eating will usually be accompanied by an agreeable feeling, whether weak or strong; but it can't be strictly said that •this agreeable feeling is the object of hunger, or that •the thought of this pleasure is what stimulates the will of the hungry man. Careful introspection seems to show that although •hunger is frequently and naturally accompanied by •anticipation of the pleasure of eating, the two aren't inseparable. And even when they do occur together, the pleasure seems to be the object not of the primary appetite but of a secondary desire that can be distinguished from the other. Someone who gets tremendous pleasure from eating may be led by this to

stimulate his hunger and delay its satisfaction in order to prolong and vary the process of satisfying it.

Indeed it's so obvious that hunger is different from the desire for anticipated pleasure that some writers have regarded its volitional stimulus (and that of desire generally) as a case of *aversion from present pain*. But this is a definite mistake in psychological classification. It's true that in desire as in pain we feel a stimulus prompting us to pass from the present state into a different one. But aversion from pain is an impulse to get out of the present state and pass into some *other* state, the only requirement being the •negative one that it be *other, not this!*; whereas in desire the primary impulse is towards the achievement of some •positive future result. When a strong desire is somehow blocked from causing action, that is usually somewhat painful; that generates a secondary aversion to the state of desire; this blends with the desire itself and may easily be confused with it. To see how different these two are, consider the fact that there are two different ways of acting on one's aversion to the pain of ungratified desire: work harder to get the desire gratified, or get rid of the pain by suppressing the desire.

Does desire have in some degree the quality of pain? The question is of psychological rather than ethical interest. . . .·but I don't mind answering it·. Speaking for my own case I have no hesitation in answering No. Consider hunger again: I certainly don't find hunger painful in normal circumstances—only when I am ill or half-starved. Generally speaking, indeed, if D is some desire that isn't blocked from generating relevant actions, then

> D is not itself a painful feeling, even when one is a long way from satisfying it,

and indeed it's often the case that

> D is an element of a state of consciousness that is over-all highly pleasurable.

Indeed, the pleasures provided by the consciousness of eager activity, in which desire is an essential element, are a large component in the total enjoyment of life. It is almost a cliché to say that such 'pleasures of pursuit' (as we may call them) are more important than the pleasures of attainment; and very often what motivates us to engage in the pursuit is precisely the pleasures of pursuit. [Sidgwick illustrates this at some length. Then:]

An interesting contrast now comes to light. In the case of hunger, the appetite of hunger is distinct from the desire for the pleasure of eating, but there's no difficulty about their both being present in full strength in one person. But with the pleasures of pursuit there does seem to be a certain incompatibility: it seems that a certain subordination of self-regard is needed if the person is to have full enjoyment. Take the case of a man engaged in pursuing some goal who keeps his main conscious aim perpetually fixed on the pleasure he expects to get from succeeding. He won't catch the full spirit of the chase; his eagerness will never get just the sharpness of edge that gives the pleasure of pursuit its highest zest. This brings us to what we may call the fundamental paradox of hedonism, that if the impulse towards pleasure is too predominant it will defeat its own aim. This effect is scarcely visible in the case of passive sensual pleasures. But it's certainly true of our active enjoyments generally—whether associated with bodily or intellectual activities—as well as of many emotional pleasures, that we can't attain them in their highest degree as long as we keep our main conscious aim concentrated on them. It's not just that •the exercise of our faculties isn't sufficiently stimulated by a mere desire for the pleasure of •it, and can't be fully developed without other more objective 'extra-regarding' [see Glossary] impulses; it's also, further, the case that these other impulses must be temporarily predominant and absorbing

if the exercise and the pleasure of it are to attain their full scope. Many middle-aged Englishmen would say business is more agreeable than amusement; but they wouldn't find it so if they transacted their business with a perpetual conscious aim at the pleasure of doing so. The pleasures of thought and study, also, can be enjoyed in the highest degree only by those who have an eagerness of curiosity that temporarily carries the mind away from self and its sensations. . . .

The important case of the benevolent affections is at first sight more doubtful. When those whom we love are pleased or pained, we ourselves feel sympathetic [see Glossary] pleasure and pain, and the flow of love or kindly feeling ·involved in any benevolent action· is itself highly pleasurable. So that it's at least plausible to think that benevolent actions aim ultimately at getting one or both of these two pleasures—·the flow-of-love pleasure and the sympathetic pleasure·—or at ·getting the flow-of-love pleasure and· averting sympathetic pain. But ·there are three reasons not to accept this as a full account of benevolent motives·. **(a)** The impulse to beneficent action produced in us by sympathy is often vastly stronger than any consciousness of sympathetic pleasure and pain in ourselves, so that it would be paradoxical to regard this latter—·i.e. getting the sympathetic pleasure or averting the sympathetic pain·—as its object. Often, indeed, a tale of actual suffering excites us in a way that is more pleasurable than painful. . . .and yet it also stirs in us an impulse to relieve the suffering, even when this relief is painful and laborious and involves various sacrifices of our own pleasures. **(b)** We can often free ourselves from sympathetic pain most easily by turning our thoughts away from the other person's suffering; and we sometimes feel an egoistic impulse to do this, which we can then clearly distinguish from the sympathetic impulse prompting us to relieve the suffering. **(c)** It seems that the much-commended pleasures of benevolence don't amount to much unless we already had a desire to do good to others for *their* sake. As Hutcheson explains, we can *cultivate* benevolent affection for the sake of the pleasures that come with it (just as the glutton cultivates appetite), but we can't produce it at will, however strongly we desire these pleasures; and when a benevolent affection exists, even if it arose from a purely egoistic impulse, it is still essentially a desire to do good to others for their sake and not for ours.

The self-abandonment and self-forgetfulness that seemed essential for the full development of the other elevated impulses don't normally and permanently characterise benevolent affection, because strong love seems naturally to involve a desire for reciprocated love. . . ., and thus the consciousness of self and of one's own pleasures and pains seems often to be strengthened by the intensity of the affection that binds one to others. Still, this self-suppression—this filling of one's consciousness with the thought of other people and their happiness—is a common feature of all strong affections; and it is said that those who love intensely sometimes feel •a conflict between the egoistic and altruistic elements of their desire, and •an impulse to suppress the egoistic side, which can show itself in acts of fantastic and extravagant self-sacrifice.

So if reflection on our moral consciousness seems to show that (as William Lecky put it) 'the pleasure of virtue can be obtained only on the express condition of its not being the object sought', we need not distrust this result of observation on grounds that it is abnormal. It is in fact merely another instance of a psychological law that we have seen at work across the whole range of our desires. In the promptings of the senses, no less than in those of intellect or reason, we find the phenomenon of *strictly disinterested impulse*; it's not only sublime and ideal ends that excite desires of

this kind—low and trivial ones can do it too. It is true of some pleasures of the merely animal life, as well as of the satisfactions of a good conscience, that they can be obtained only if they are not directly sought.

3. I have stressed the felt incompatibility of 'self-regarding' and 'extra-regarding' impulses because I wanted to show their essential distinctness. I don't wish to overstate this incompatibility; I believe that it is usually very transient and often only momentary, and that our greatest happiness (if that is what we are after) is generally achieved through a sort of alternating rhythm of the two kinds of impulse in consciousness. Our conscious desires are, more often than not, chiefly extra-regarding, but where there's a strong desire in any direction there is commonly a keen openness to the corresponding pleasures; and the most devoted enthusiast is sustained in his work by the recurrent consciousness of such pleasures. But the familiar and obvious instances of conflict between self-love and some extra-regarding impulse are not paradoxes and illusions that we have to explain away; rather, they are. . . .just what one might expect. If we're continually acting from impulses whose immediate objects are something other than our own happiness, it is quite natural that we should occasionally yield to such impulses when that involves losing some pleasure. Thus a man with weak self-control who has fasted for too long may easily indulge his appetite for food to an extent that he knows is unhealthy; not because the pleasure of eating appears to him at all worthy of consideration in comparison with the harm to his health; but merely because he feels an impulse *to eat food*, which prevails over his prudential judgment. Another example: men have sacrificed all life's enjoyments and even life itself to obtain posthumous fame, not from •any illusory belief that they could derive pleasure from it, but from •a direct desire for the future admiration of others and a preference for that over their own pleasure. And yet another: when someone makes a sacrifice for some ideal end—e.g. truth, freedom, religion—it may be a real sacrifice of the person's happiness and not the preference for one highly refined pleasure (or of the absence of one special pain) over all the other elements of happiness. . . .

To sum up: •we are conscious of having ever so many extra-regarding impulses, i.e. ones that are directed towards something other than pleasure and relief from pain; •much of our pleasure depends on the existence of such impulses; •in many cases there isn't room in the person's mind for the extra-regarding impulse *and* the desire for his own pleasure; and less often (though not *rarely*) the two come into irreconcilable conflict, prompting the person to opposite courses of action; •and this incompatibility is specially prominent when. . . .the extra-regarding impulse is the love of virtue for its own sake, i.e. the desire to do what is right just because it is right.

4. . . . .The conclusion I have reached has been subject to two attacks, not trying to falsify it outright but to weaken its force. **(a)** It has been maintained that pleasure, though not the only conscious aim of human action, is always the result that it is *unconsciously* directed to. It would be hard to disprove this: no-one denies that some pleasure normally accompanies the achievement of a desired end; and there seems to be no clear method of determining whether the pleasure is aimed at if it's acknowledged not to be *consciously* aimed at. That also makes the proposition hard prove, but I have more to say against it than that. If we try to take seriously the notion of the unconscious aspect of human action, we can only conceive it as a combination of movements of material organism's parts; and the 'end' of any such movements (it's reasonable to think) must be some physical state of the organism, a state that favours the survival either of the

individual organism or of its species. In fact, the doctrine that pleasure (or the absence of pain) is the end of all human action can't be supported by •introspection or by •external observation and inference; it seems to come from an arbitrary and illegitimate combination of the two.

(b) It is sometimes said that our original impulses—the ones we had when very young—were all directed towards pleasure or from pain, and that any extra-regarding impulses are derived from these by 'association of ideas'. [Sidgwick replies •that it seems to be false, because children have many extra-regarding inferences, and •that in any case it is irrelevant to his thesis] that men do not *now* normally desire pleasure alone. . . . To say in answer to this that all men *once* desired pleasure is, from an ethical point of view, irrelevant; except on the assumption that there is an original type of man's appetitive nature to which, just because it is original, he ought to conform. But probably no hedonist would explicitly claim this, though writers of the intuitional school often make such an assumption.

[The chapter ends with a long 'Note' on the thesis that 'desire is essentially painful'. Sidgwick argues against this, focussing especially on Alexander Bain's defence of it.]

Chapter 5: Free-will

1. I have treated rational action and then disinterested action, without raising the vexed question of the freedom of the will. The long history of debates about this question reveals that it is full of difficulties, and I want to keep these within tight limits so as to reduce their disturbing influence on my topic. Now, I can't see any psychological basis for identifying •disinterested action with either •'free' action or •'rational' action; and identifying rational action with free action is at least misleading, and tends to obscure the real issue raised in the free-will controversy. In chapter 4 I tried to show that strictly disinterested action—i.e. action that isn't motivated by any foreseen balance of pleasure to the agent—is found in the most •instinctive as well as in the most •deliberate and self-conscious region of our volitional experience. And when individual's conduct is made rational by causes external to his own volition, it is still rational; so the conception of acting rationally, as explained in chapter 3, is *not* tied to the notion of acting 'freely', as libertarians generally maintain against determinists. I don't say 'all libertarians', because what Kant's disciples say about how freedom is connected to rationality seems to me to involve a confusion between two meanings of 'free', meanings that ought to be carefully distinguished in any discussion of free-will. When a Kantian[1] says that a man 'is a free agent

[1] I choose to exclude the Kantian conception of free-will from this chapter: •because of the confusion mentioned in the text, and •because it depends on the notion of a causality that isn't subject to time-conditions—a notion that I think is entirely untenable though a discussion of it doesn't fit anywhere into the plan for this work. Still, Kantian theory is having a large influence on current ethical thought, so I'll briefly discuss his conception of 'free-will' in an Appendix [not included in this version].

in so far as he acts under the guidance of reason', it's easy to agree because, as William Whewell says:

> 'We ordinarily identify ourselves with our reason rather than with our desires and affections. We speak of desire, love, anger, as •mastering us or of ourselves as •controlling them. If we decide to prefer some remote and abstract good to immediate pleasures, or to conform to a rule that brings us present pain (a decision requiring the exercise of reason), we regard those acts as more particularly our own acts.'

So I don't have any ordinary-language objection to this use of the term 'free' to label voluntary actions in which the pull of appetite or passion is successfully resisted; and I'm aware of what a help it is to moral persuasion if the powerful sentiment of liberty is enlisted in this way on the side of reason and morality.[1] But if we say that a man is a 'free' agent to the extent that he acts rationally, we can't also say—in the same sense of 'free'—that when he acts irrationally he is doing this by his own 'free' choice; yet that is just what the libertarians have usually wanted to say. They have thought it important to show that any moral agent is 'free', because of the connection they think exists between freedom and moral responsibility; but obviously a freedom connected in that way with responsibility isn't the 'freedom' that shows up in rational action; it's the freedom to *choose between* right and wrong, which shows up in either choice. The Christian notion of 'willful sin' implies that men do deliberately and knowingly choose to act irrationally. It's not merely that they prefer self-interest to duty, for that's a conflict of claims to rationality than clear irrationality; but rather that they prefer sensual indulgence to health, revenge to reputation etc., though they know that this is opposed to their true interests as well as to their duty.[2] So our experience as a whole doesn't present the conflict between •reason and •passion as a conflict between •'ourselves' and •a force of nature. We may speak of being 'the slaves of our desires and appetites', but we must admit that we *chose* our slavery. Well, can we say this ↓ about the willful wrongdoer?

> His choice was 'free' in this sense: he could have chosen rightly even if all the antecedents of his choice had been what they actually were.

I take this to be the substantial issue raised in the free-will controversy; and I'll now briefly discuss it, since it is widely believed to be of great ethical importance.

2. The predicates 'right' and 'what ought to be done'—when taken in the strictest ethical sense—are applicable to voluntary actions and to nothing else; all methods of ethics agree about this. Let us start, then, by defining this notion of *voluntary action* more exactly. In the first place, voluntary action is *conscious*, which marks it off from unconscious or mechanical actions or movements of the human organism. The person whose organism [Sidgwick's phrase] makes such movements doesn't become aware of them until after they have been made; so they are not imputed to him as a person, or judged to be morally wrong or imprudent; though they

[1] But it's also true, as I'll show later, that we sometimes identify ourselves with passion or appetite in conscious conflict with reason; and in those cases the rule of reason is apt to seem like an external constraint, and obedience to it a servitude if not a slavery.

[2] The difficulty that Socrates and the Socratics had in conceiving a man to choose deliberately what he knows to be bad for him—a difficulty that drives Aristotle into real determinism in his account of purposive action, at the same time explicitly maintaining the 'voluntariness' and 'responsibility' of vice—seems to be much reduced for the modern mind by the distinction between •moral and •prudential judgments, and the *prima facie* conflict between 'interest' and 'duty'. Because we are thus familiar with the conception of deliberate choice consciously opposed *either* to interest *or* to duty, we can quite easily conceive of such choice in conscious opposition to both. See chapter 9.3.

may sometimes be judged to have good or bad consequences, implying that they ought to be encouraged or checked as far as this can be done indirectly by conscious effort.

Someone who performs a conscious action isn't regarded as morally culpable, except in an indirect way, for entirely unforeseen effects of his action. When a man's action causes unforeseen harm, he is often blamed for carelessness; but thoughtful people would generally agree that if in such a case the agent is morally to blame for anything it must be that his carelessness resulted from some willful neglect of duty. So it seems that the proper immediate objects of moral approval or disapproval are always the results of a man's volitions that he *intended*, i.e. that he thought of as certain or probable upshots of his volitions. Or, more strictly, (dis)approval attaches to the volitions themselves in which such results were thus intended: if external causes prevent the agent's volition from producing its intended result, we don't excuse him on that account.

This seems to differ from the common opinion that the morality of acts depends on their 'motives', if by 'motives' we mean the desires that we feel for some of the foreseen consequences of our acts. But I don't think that those who have this opinion would deny that we are blameworthy for any bad result that we *foresaw* in willing, whether or not we *wanted* it. No doubt it is commonly held that acts, similar as regards their foreseen results, may be 'better' or 'worse' through the presence of certain desires or aversions. (More about this in chapter 9.) Still so far as these feelings are not altogether under the control of the will, the judgment of 'right' and 'wrong' doesn't strictly apply to the feelings themselves but rather to the exertion or omission of voluntary effort to check bad motives and encourage good ones. . . .

So judgments of right and wrong relate to volitions accompanied by intention, whether the intended effects are •external or •something involving the agent's own feelings or character. This excludes conscious actions that aren't strictly intentional, as when sudden strong feelings of pleasure or pain cause movements that we are aware of making but aren't preceded by any thought of those movements or of their effects. . . .

Our common moral judgments distinguish •impulsive wrongdoing from •deliberate wrongdoing, condemning the latter more strongly. The line between them isn't sharp; but we can define 'impulsive' actions as ones where a feeling prompts the action so simply and immediately that there's almost no sense of choosing the intended result. In a deliberate volition there's a conscious *selection* of the result.

With volitions that are objects of moral condemnation or approval, the concept *volition* seems to include

- •intention, or thought of the results of the action, and
- •awareness of oneself as choosing, deciding, determining these results.

What I take to be at issue in the free-will controversy is the question: Which of these two is true?

(1) The self which I credit with making my deliberate volitions has strictly determinate moral qualities, a definite character—partly inherited, partly formed by my past actions and feelings, and by physical influences—so that my voluntary action at any moment is completely caused by the qualities of this character together with the external influences acting on me at that moment (including the present state of my body).

(2) There is always a possibility of my choosing to act in the way that I now judge to be reasonable and right, whatever my previous actions and experiences may have been.

In **(1)** a materialist would substitute 'brain and nervous

system' for 'character', and thereby obtain a clearer notion; but I have left •materialism out of this because •determinism doesn't require it. For my present purposes, the substantial dispute relates to the whether every volition depends causally on the state of things at the preceding instant—and it makes no difference whether that state of things consists in character and circumstances or brain and environing forces.[1]

On the determinist side there is a cumulative argument of great force. All competent thinkers believe that events are determined by the state of things immediately preceding them—all kinds of events except human volitions. This belief has steadily grown •in clarity and certainty and •in the scope of its application, as the human mind has developed and human experience has been systematised and enlarged. Lines of thought conflicting with this have, step by step in successive branches of science, receded and faded, until at length they have vanished everywhere, except from this mysterious citadel of *will*. Everywhere else the belief is so firmly established that some declare its opposite to be •inconceivable, and others even maintain that it always was •so. Every scientific procedure assumes it; each success of science confirms it. We are finding more and more proof not only •that events are determined in discoverable ways but also •that the different sorts of determination of different kinds of events are all inter-connected and are basically the same. So we are increasingly convinced of the essential unity of the knowable universe, which increases our unwillingness to credit human action with the exceptional character claimed for it by libertarians.

Again, we see that the portion of human action that is originated unconsciously is admittedly determined by physical causes; and we find that no clear line can be drawn between acts of this kind and ones that are conscious and voluntary. [Sidgwick develops this point with examples of kinds of situation where the unconscious/conscious line is especially hard to draw.]

Further, we always explain the voluntary action of everyone but ourselves on the principle of causation by character and circumstances.[2] Indeed social life would be impossible if we didn't; for the life of man in society involves daily a mass of tiny forecasts of the actions of other men, based on experience of •mankind generally or of •particular classes of men or of •individuals; so that individuals are necessarily regarded as things having determinate properties, causes whose effects are calculable. With people we know, we usually infer their future actions from their past actions; and when our forecast turns out to be wrong, we explain this in terms not •of the disturbing influence of free-will but •of gaps in our knowledge of their character and motives. And passing to whole communities: whether or not we believe in a 'social science', we all take part in discussions of social phenomena in which the same principle is assumed; and however we may differ as to particular theories, we never doubt the validity of the assumption; and if we find anything inexplicable in history, past or present, it never occurs to

[1] Some determinists conceive of each volition as connected by uniform laws with our past state of consciousness. But any uniformities we might trace among a man's past states, even if we knew them all, would still give us very incomplete guidance to his future actions, because there would be left out of account •all inborn tendencies that hadn't yet completely shown themselves, and •all past physical influences whose effects hadn't been perfectly represented in consciousness.

[2] I don't mean that this is the only view that we take of the conduct of others; in judging their conduct morally, we ordinarily apply the conception of free-will. But we don't ordinarily regard this as one kind of causation that limits and counteracts the other kind. More about this later.

us to attribute it to an extensive exercise of free-will in a particular direction. Indeed, even as regards our own actions: however 'free' we feel ourselves to be at any moment T—however intensely our *choice* seems to be unconstrained by present motives and circumstances, and unfettered by our previous actions and feelings—when we later look back on •our choice at T and put it in the series of our actions, its causal relations and similarities to other parts of our life appear, and we naturally explain •it as an effect of our nature, education, and circumstances. Indeed we even think in that way about our future actions, and the more developed our moral sentiments are the more inclined we are to do this. [Sidgwick explains why: with increasing moral seriousness we acquire a growing sense of the role any choice might have in affecting our future thoughts and actions. But, he adds,] we habitually adopt at the same time the opposite, libertarian, view about our future ·choices·; we believe, for example, that we are perfectly able from now on to resist temptations that we have continually succumbed to in the past. But moralists of all schools admit and even insist that this belief is largely illusory. Though •libertarians contend that we *can* at any moment act in a manner opposed to our past customs and present tendencies, •they join the determinists in teaching that breaking away from the subtle unfelt drag of habit is much harder than it is usually thought to be.

3. Against •the formidable case for determinism there is •what consciousness tells us at the moment of deliberate action. When I'm sharply aware of having a choice between two ways of behaving, one of which I think to be right or reasonable, I can't help thinking that I can now choose to do the right thing if there's no external obstacle to my doing it—that I *can* do this, however strongly I want to do the wrong thing and however often I have yielded to such wants in the past. I realize that each concession to vicious desire increases the difficulty of resisting it next time around, but the •difficulty always seems to be entirely different from •impossibility. I admit that in some cases a certain impulse—e.g. aversion to death or extreme pain, or a craving for alcohol or opium—becomes so intense that it is felt as irresistibly dominating voluntary choice. We usually hold that a person is not morally responsible for what he does under such a dominating impulse; but the moral problem *that* raises is very exceptional; in ordinary cases of giving in to temptation there's no sense of an irresistible impulse. Ordinarily, however strong the rush of appetite or anger that comes over me, it doesn't present itself as irresistible; and if I deliberate at such a moment I can't regard the mere force of the impulse as a reason for doing what I judge to be unreasonable. I can suppose that •my conviction of free choice *may* be illusory; that •if I knew my own nature I *might* see it to be already settled that on this occasion I am going to act against my rational judgment. But when I think of myself as seeing this, I have to think of myself as having a fundamentally altered conception of what I now call 'my' action; I can't conceive that if I saw the actions of my organism in this ·determinist· light I would attribute them to my 'self'—i.e. to the mind so contemplating—in the way in which I now attribute them. It's not surprising that the theoretical question about the freedom of the will is still answered differently by reputable thinkers; and I don't want to answer it now. But it may be useful for me to show that the ethical importance of answering it is liable to be exaggerated, and that anyone who considers the matter carefully will find this importance to be very limited.

Libertarians are most likely to exaggerate the ethical importance of the free-will question. Some libertarian writers maintain that the conception of the freedom of the will, alien

as it may be to positive [see Glossary] science, is indispensable to ethics and legal theory, because in judging that I 'ought' to do something I imply that I 'can' do it, and similarly in praising or blaming an action of yours I imply that you 'could' have acted otherwise. So some people say this:

> If a man's actions are mere links in a chain of causation that ultimately goes back to events that occurred before he was born, he can't really have either merit or demerit; and if he has neither, it's against the common moral sense of mankind to reward or punish—or even to praise or blame—him.

Let us clear the ground by assuming that for present purposes we are confident and agreed about what it is right to do, except for rights or wrongs that arise from the present question. And let us tackle the question of the importance of free-will in relation to moral action *generally*, setting aside the *special* question of its importance in relation to punishing and rewarding; because in punishing and rewarding the focus is not on the present freedom of the agent but the past freedom of the person now being acted on.

As regards action generally, the determinist accepts that a man is morally bound to do x only if doing x is 'in his power', which he explains as meaning that x will be done if the man chooses to do it. I think this is the sense in which the proposition *What I ought to do I can do* is commonly taken; it means 'can do if I choose', not 'can choose to do'. Still the question remains 'Can I *choose to do* x, which in ordinary thought I judge to be right to do?' I hold that within the limits I have explained I inevitably conceive that I can choose to do x; but I can envisage •regarding this conception as illusory and •judging on the basis of my past record that I certainly won't choose x and therefore that such a choice is not really possible for me. If I do get into that frame of mind, this judgment will cancel or weaken the operation of the moral motive in the case of x, because one or other of these two must happen:

(i) I don't judge it be reasonable to choose to do x, or

(ii) I do make that judgment, but I also judge that the conception of duty it involves is just as illusory as the conception of freedom.

I go that far in conceding the libertarian view about the demoralising effect of a really firm belief in determinism. But there are very few cases where, even on determinist principles, I can legitimately conclude that it is certain—not just highly likely—that I will deliberately choose to do something that I judge to be unwise.[1] Ordinarily the legitimate inference from •a man's past experience and •from his general knowledge of human nature would not take him further than a very strong probability that he will choose to do wrong; and a mere strong probability that I shan't will to do right can't be regarded by me in deliberation as a reason for not willing. What it *does* provide a good reason for is willing strongly. . . . [Sidgwick remarks that the question *What is the moral effect of thinking it highly probable that one will not choose to act rightly?* is one that both libertarian and determinist might usefully think about. He concludes:] In all ordinary cases, therefore, it's not relevant to ethical deliberation to find the answer to

[1] When a man yields to temptation, judging that it's 'no use trying to resist', I think he is probably engaging in semi-conscious self-sophistication [see Glossary], due to the influence of appetite or passion on his reasoning. This self-sophistication will probably take a determinist form in the mind of a determinist, but a libertarian is in equal danger of self-sophistication, though in his case it will take a different form. Where a determinist would reason 'I certainly shall take my usual glass of brandy to-night, so there is no use resolving not to take it', the libertarian would reason 'I mean to stop taking brandy, but it will be just as easy to stop tomorrow as today; I will therefore have one more glass, and stop tomorrow.'

> 'Regarding my sense of being free to choose whatever I may conclude is reasonable—is it metaphysically valid?'

unless the answer, one way or the other, changes my view of *what* it would be reasonable to choose to do if I could so choose.

There shouldn't be any change of view concerning the ultimate **ends** of rational action—the ones that I took (in chapter 1) to be commonly accepted. If •private or general happiness is taken to be the ultimate end of action on a libertarian view, the adoption of determinism provides no reason to reject it; and if •excellence is in itself admirable and desirable, it remains so whether or not any individual's approximation to it is entirely determined by inherited nature and external influences—unless the notion of excellence includes that of free-will, ·which it doesn't!· Free-will is obviously not included in our common ideal of physical and intellectual perfection; and I can't see that it is included, either, in the common notions of the excellences of character that we call 'virtues'. The instances of courage, temperance, and justice don't become less admirable because we can trace their antecedents in a happy balance of inherited dispositions developed by a careful upbringing.[1]

Well, then, can affirming or denying free-will affect our view of the best **means** for attaining either end? That may depend on what our grounds are for that view. **(a)** They may involve a belief that the world has a moral government: according to the usual form of *that*, doing one's duty is the best means to happiness because the world has a moral government through which God will reward virtue and punish vice in an after-life. If •free-will is essential to the moral government of the world and an after-life for men, that obviously gives it basic ethical importance—not in determining our duty but in reconciling it with our interests. This is the main element of truth in the view that denying free-will is removing motives to doing our duty; and I admit that this is right, to the extent that

> **(1)** if we set aside theological considerations, the course of action conducive to our interest diverges from our duty, and
>
> **(2)** free-will is an essential part of the theological reasoning that removes this divergence.

I'll examine **(1)** in II/5. But **(2)** really lies outside the scope of this work.[2]

(b) If our belief about the best means to happiness is based on empirical grounds, it seems not to let the issue over free-will into the picture. A man is deliberating on whether to do A; if he does, what will the consequences be? It's not plausible to say that that depends on whether his doing A was pre-determined! But you may say:

> In considering how to act, we should take into account the probable future conduct of ourselves and others; and for this we need an answer to the question of free-will, so that we can know whether the future *can* be predicted from the past.

But I can't see that this has any definite practical upshots. However far we go in admitting free-will as a cause that might kick in and falsify the most scientific forecasts of human action, it would be an absolutely unknown cause, so that our recognition of it couldn't lead us to change *what* we predict, though it might reduce our confidence in our predictions.

[1] The ordinary notion of merit *does* become inapplicable. But I can't see that perfection becomes less an end to be aimed at because we stop regarding the attainment of it as meritorious. God's actions aren't thought of as having 'merit', but no-one infers from this that he isn't perfect.

[2] Though an important section of theologians who have had the most intense belief in the moral government of the world have been determinists.

Suppose we were convinced that all the planets have free-will, and are kept in their courses only by the continual exercise of free choice in resistance to strong centrifugal or centripetal forces. Our general confidence in the future of the solar system might reasonably be impaired, though it's hard to say how much; but the details of our astronomical •calculations wouldn't be affected because we couldn't re-do •them so as to take free-will into account. And the situation will be like that in the forecast of human conduct if psychology and sociology ever become exact sciences. At present, however, they are so far from exactness that this additional element of uncertainty—coming from crediting humans with free-will—can hardly have even any emotional effect.

•When we reason to any definite conclusions about how we or others will act, we have to consider such actions as determined by strict laws. If they aren't perfectly strict, our reasoning is to that extent liable to error, but it's not as if we could choose to reason in some other way. •When we are trying to decide (on some basis or other) how it would be reasonable to act right now, determinist conceptions are irrelevant—whereas in the preceding case they were inevitable. Thus, deciding the metaphysical issue about free-will has no practical importance in the general regulation of conduct, unless—moving across from ethics to theology—we base the reconciliation of duty and ·self·-interest on a theological argument that requires free-will as a premise.

4. I have argued that a man's adopting determinism shouldn't affect •his view of what it's right for him to do or •his reasons for doing it (except in certain exceptional circumstances or on certain theological assumptions). But this may be said:

> Granting that the reasons for right action aren't altered ·by believing in determinism·, the motives that prompt to it will be weakened, because a man won't feel remorse [see Glossary] for his actions if he regards them as necessary results of causes that existed before he did.

The sentiment of remorse implies self-*blame*, so I admit that it must tend to vanish from the mind of a convinced determinist. Still I don't see why a determinist's imagination shouldn't be as vivid as a libertarian's, his sympathy as keen, and his love of goodness as strong; so I don't see why his •dislike for the damaging qualities of his character that have caused him to act badly shouldn't be as effective a source of moral improvement as •remorse would be. Men generally seem to take no more trouble to cure •moral defects than they do to cure equally damaging •defects in their circumstances, their bodies and their intellects that don't cause them remorse.

This brings up the issue of the effect of determinism on the assignment of punishment and reward. For it must be admitted that the common retributive view of punishment—and the ordinary notions of merit, demerit, and responsibility—involve the assumption of free-will. If the wrong act and the bad character shown by it are seen as inevitable effects of causes right outside the agent, he can't be morally responsible (using the ordinary notion of this) for the harm caused by them. But the determinist can give to 'He deserves punishment' and 'He is responsible' etc. meanings that are •clear and definite and from a utilitarian point of view •the only suitable meanings. When they are in play, a determinist can say that someone 'is responsible for' a harmful act ·and deserves to be punished for it·, meaning that it is right to punish him for it—primarily so that the fear of punishment may prevent him and others from acting like that in future. This view of punishment is in theory very different from the common view; but when in I come in III/5 to examine in detail the current conception of justice

I'll argue that the difference can have hardly any practical effect, because in rewarding services or punishing bad acts it's practically impossible to be guided by any considerations except those embodied in the determinist interpretation of desert [see Glossary]. For instance, the treatment of legal punishment as deterrent and reformative, rather than retributive, seems to be forced on us by the practical demands of the order and well-being of society, quite apart from any determinist philosophy.[1] Moreover, as I shall show in III/5.5, if the retributive view of punishment is taken strictly—with no input from the preventive view—it puts our conception of justice into conflict with benevolence because it presents punishment as a purely useless evil. In the sentiments expressed in moral praise and blame there *is* ·a difference between the determinist and the libertarian·. Where the libertarian seeks to express something about what the person *deserves*, the convinced determinist wants only to encourage good conduct and prevent bad; but I don't see why the determinist's moral sentiments shouldn't promote virtue and social well-being as effectively as the libertarian's.

5. *How far does the power of the will actually extend?* The answer to this defines the range within which ethical judgments have a proper place, so it's obviously important for us to know what it is. The question is independent of the free-will question; we can state it in determinist terms thus: *What effects can be caused by human volition if adequate motives are in place?* These effects are mainly of three kinds: **(i)** changes in the external world consequent on muscular contractions; **(ii)** changes in the series of ideas and feelings that constitutes our conscious life; and **(iii)** changes in the tendencies to act in certain ways under certain circumstances.

(i) The most conspicuous work done by volitional causation consists in events that can be produced by muscular contractions. It is sometimes said that what we really will is •the muscular contraction and not •its effects. That is because •the latter involve a contribution from other causes, so that we can never know for sure that •they will follow. But strictly speaking it's not certain that the muscular contraction will follow, because our limb may be paralysed, etc. The *immediate* upshot of the volition is some molecular change in the motor nerves; but when we will to do something we aren't aware of changes in our motor nerves or indeed (usually) of the muscular contractions that follow them; so it seems wrong to speak of either of those as what our mind is aiming at in willing; what we consciously will and intend is almost always some effect that is further along the causal chain than those. Still, some contraction of our muscles is required for almost all effects of our will on the external world; and when that contraction is over, our part in the causation is completed.

(ii) We can to some extent control our thoughts and feelings. A good deal of what we commonly call 'control of feeling' belongs in **(i)**: by controlling our muscles we can keep down the expression of a feeling and resist its promptings to action; and—because freely expressing a feeling usually sustains and prolongs it—this muscular control amounts to a power over the emotion. There's no such connection between our muscular system and our thoughts; but experience shows that most men can, to a greater or lesser extent, *voluntarily* direct their thoughts and pursue *at will* a given line of meditation. It seems that in these cases the effort of will produces a concentration of our consciousness on a

[1] Thus we find it necessary to punish •negligence when its effects were very serious, even when we can't trace •it to willful disregard of duty; and to punish rebellion and assassination even if we know that they were prompted by a sincere desire to serve God or to benefit mankind.

part of its content, so that this part grows more vivid and clear while the rest tends to recede into the shadows and eventually to vanish. (This voluntary exertion is often needed only to start a series of ideas which then continues without effort. . . .) By concentrating our minds in this way we can free ourselves of many thoughts and feelings that we don't want to dwell on; but our power to do this is limited, and if a feeling is strong and its cause is persistent we need a very unusual effort of will to banish it in this way.

(iii) Another effect of volition, which deserves special attention, is the alteration in men's tendencies to future action. This is presumably an effect of general resolutions about future conduct, insofar as these have any effects. Even a resolution to do a particular act. . . .must be supposed to produce a change of this kind; it must somehow modify the person's present tendencies to act in a certain way on a foreseen future occasion. But the practical importance of knowing what is within the power of the will mainly concerns *general* resolutions for future conduct. [Sidgwick now devotes a page to refuting this:

> The thesis that we have free-will implies that any effort of will we make to amend our future behaviour will be completely effective

—which he says is 'sometimes vaguely thought'. He remarks, among other points, that the free-will thesis should make one less, not more, confident that one's present volition will succeed. The target thesis seems so implausible that we can safely excuse ourselves from following Sidgwick's detailed destruction of it. He goes on to explain why he cares about this:]

I hope that this discussion will dispel any lingering doubts you have concerning my thesis that the free-will controversy has little or no practical importance. You may have had such doubts because you vaguely thought this:

> On the determinist theory it is sometimes wrong to perform a single act of virtue because we have no reason to believe we will follow through with it; but on the assumption of freedom we should always boldly do what would be best if it were consistently followed through with, because we are conscious that such consistency is ·always· in our power.

But this supposed difference vanishes when we recall that any effort of resolution *now* can produce only a certain limited effect on future actions, and that immediate consciousness can't tell us that this effect will be adequate to the occasion—or indeed how great it will really prove to be. For the most extreme libertarian must then allow that before pledging ourselves to any future course of action we ought to estimate carefully—from our experience of ourselves and of people in general—how likely we are to keep our present resolutions in the circumstances we are likely to be in. Of course we shouldn't peacefully accept any weakness or lack of self-control; but the fact remains that such weakness can't be cured by a single volition; and whatever we can do towards curing it, by any effort of will at any moment, is as clearly enjoined by reason on the determinist theory as it is on the libertarian. Neither theory makes it reasonable for us to •deceive ourselves about the extent of our weakness, or •ignore it in forecasting our own conduct, or •suppose it to be more easily remediable than it really is.

Chapter 6: Ethical principles and methods

1. The results of the three preceding chapters can be briefly stated as follows. Ethics aims to systematise and free from error the apparent cognitions that most men have of the rightness or reasonableness of conduct, whether the conduct be considered as right •in itself or •as a means to some end commonly regarded as ultimately reasonable.[1] These cognitions are normally accompanied by emotions known as 'moral sentiments'; but an ethical judgment doesn't merely affirm the existence of such a sentiment; indeed it's an *essential* characteristic of a moral feeling that it is bound up with an apparent cognition of something more than mere feeling. I call such cognitions 'dictates' or 'imperatives' because when they are brought into practical deliberation they are accompanied by a certain impulse to do the acts recognised as right. . . . As long as this impulse is effective in producing right volition, it is not of primary importance for ethical purposes to know exactly what emotional states precede such volitions. And this remains true even if the force actually operating on the person's will is mere desire for the pleasures that he thinks the right conduct will bring or aversion to the pains that he thinks it will prevent; though in that case his action doesn't fit our common notion of strictly virtuous conduct, and though there's no evidence that such desires and aversions are the sole—or even the normal—motives for human volitions. Something else that it's not generally important to know: whether we are always metaphysically speaking 'free' to do what we clearly see to be right. What I 'ought' to do, in the strictest use of the word 'ought' is always 'in my power', in the sense that there is no obstacle to my doing it except absence of adequate motive; and when I am deliberating what to do, it is usually impossible for me to regard such an absence of motive as a reason for not doing what I otherwise judge to be reasonable.

What do we commonly regard as valid ultimate reasons for acting or abstaining? This, as I said, is the starting-point for the discussions of the present work, which is not primarily concerned with

> proving or disproving the validity of any such reasons,

but rather with

> expounding the different 'methods' or rational procedures for determining right conduct in cases where this is logically connected with various widely accepted ultimate reasons.

I showed in chapter 1 that such reasons were supplied by the notions of

(i) happiness and
(ii) excellence or perfection (prominently including virtue or moral perfection), and
(iii) duty,

with **(i)** and **(ii)** regarded as ultimate ends, and **(iii)** prescribed by unconditional rules. This three-part conception of the ultimate reason for conduct corresponds to a three-part way of looking at human existence. We distinguish **(ii)** the conscious being from the stream of conscious experience, and we distinguish this stream into **(iii)** acting and **(i)** feeling. . . . Other reasons have also been widely accepted as ultimate

[1] As I have already said, we could determine right conduct relative to an ultimate end—whether happiness or perfection—without regarding the end as prescribed by reason; all that's needed is for it to be adopted as ultimate and paramount. But in the present work I confine my attention to ends that are widely accepted as reasonable; and in III/12 I shall try to exhibit the self-evident practical axioms that I think are implied in this acceptance.

grounds of action:

(a) many religious people think that the highest reason for doing anything is that it's God's will;

(b) for others 'self-realisation' is the really ultimate end,

(c) and for others again it's 'life according to nature'.

And it's not hard to see why conceptions like these are thought to provide deeper and more completely satisfying answers to the basic question of ethics than the three I have focused on. It's because they represent what *ought to be* in an apparently simple relation to what actually *is*. The fundamental facts of existence are **(a)** God, **(c)** Nature and **(b)** Self. The knowledge of **(a)** what will accomplish God's will, **(c)** what is according to Nature, **(b)** what will realise the true Self in each of us—these seem to solve the deepest problems of metaphysics as well as of ethics. But just because these notions do combine the ideal with the actual, they properly belong not in ethics as I define it, but in philosophy, the central and supreme study of the relations among all objects of knowledge. Introducing these notions into ethics is liable to create a deep confusion between 'what is' and 'what ought to be', destructive of all clearness in ethical reasoning; and if that confusion is avoided and the strictly ethical import of **(a)**–**(c)** is made explicit, they appear always to lead us to one or other of **(i)**–**(iii)** .

There's the least danger of confusion over 'God's will', because here the connection between 'what is' and 'what ought to be' is perfectly clear and explicit. We think of the content of God's will as existing now, •as an idea; the end to be aimed at is making it •actual. A question arises: how could God's will fail to be realised [see Glossary]? If it can't fail to be realised, whether we act rightly or wrongly, how can realisation provide the ultimate motive for acting rightly? But this difficulty is for theology to solve, not ethics. The practical question—·the one that can't be shunted off to theology·—is this:

> Assuming that God's will somehow creates the facts about what we ought to do, how are we to discover *what* he wills in any particular case?

This must be either by **(1)** revelation or by **(2)** reason or by **(3)** both combined. **(1)** If an external revelation is proposed as the standard, we are obviously carried beyond the range of our present study ·which concerns ethics·. **(2)** If we try to discover the divine will by reason,. . . .the situation is this: we select as being in accordance with the divine will whatever *we* know to be dictated by reason. [To make sure that this is clear: we answer 'Does reason tell us that God wants us to do A?' by answering 'Does reason tell us to do A?'] So it is usually assumed **(i)** that God desires the happiness of men,. . . .or **(ii)** that he desires their perfection,. . . .or **(iii)** that whatever his *end* may be (into which perhaps we have no right to inquire) his laws are immediately knowable, being in fact the first principles of intuitional morality. Or perhaps it is explained that God's will is to be learned by examining our own constitution or that of the world we are in; so that **(a)** conformity to God's will seems to come down to **(b)** self-realisation or **(c)** life according to nature. In any case, this conception ·of God's will·, however important it may be in supplying new motives for doing what we believe to be right, doesn't suggest any special criterion of rightness unless revelation is brought in.

2. Let us now consider the notions *nature*, *natural* and *conformity to nature*. To get a principle distinct from *self-realisation* (which I'll deal with in chapter 7), we should take it that the *nature* to which we are to conform is not each person's individual nature but human nature generally. . . ., and that we are to find the standard of right conduct in a certain type of human existence that we can somehow abstract from observation of actual human life. Every rational man must of course 'conform to nature' in the

sense of adapting his efforts (whatever goal he is aiming at) to the particular physical and mental conditions of his existence. But if he is to go beyond this and look to 'Nature' for guidance in choosing an ultimate end or paramount standard of right conduct, that must be on the basis either of •strictly theological assumptions or at least •a more or less definite recognition of *design* in the empirically known world. If we find no design in nature, and think of the world's processes as an orderly but aimless drift of change, knowing these processes and their laws may •limit the aims of rational beings but I can't see how it could •determine the ends of their action or be a source of unconditional rules of duty. And those who use *natural* as an ethical notion do commonly suppose that by attending to the actual play of human impulses or the physical constitution or social relations of man we can find principles that completely settle the kind of life man was designed to live. But every such attempt to derive *what ought to be* from *what is* obviously collapses as soon as it is freed from fundamental confusions of thought. If for example we want to get practical guidance from the conception of *human nature* regarded as a system of impulses and dispositions, we must obviously give a special precision to the meaning of 'natural'. Why? Because every impulse is 'natural' in a sense (Butler's point), and there's no guidance for us in this: the question of duty is never raised except when we want to know which of two conflicting impulses we ought to follow. 'The supremacy of reason is natural'—it's no use saying *that*, because we have started by assuming that reason tells us to follow nature, so that our line of thought would become circular—·*Nature points to Reason, which points to Nature*·. The Nature that we are to follow must be distinguished from our practical reason, if it's to become a guide to it. Then how *can* we distinguish the 'natural impulses' that are to guide rational choice from the unnatural ones? The friends of *the natural* seem usually to have interpreted 'natural' to mean either

- •*common* as opposed to the rare and exceptional, or
- •*original* as opposed to what develops later, or
- •*not an effect of human volition* as opposed to the artificial.

But I have never seen any basis for the view that nature abhors the exceptional, or prefers the earlier in time to the later; and looking back over human history we find that some admired impulses—e.g. the love of knowledge, enthusiastic philanthropy—are both rarer and later in their appearance than others that all judge to be lower. [Sidgwick goes on to argue that if we take 'natural' to mean 'not produced by the institutions of society', the injunction to follow nature will produce some morally absurd results. And tying 'natural' to the natural processes of our bodies the prescription 'follow nature' will give us very little guidance, because:] almost always the practical question is not •*whether* we are to use our organs or leave them unused but •*how* we are to use them. . . .

A last try: consider man in his social relations as father, son, neighbour, citizen, and try to determine the 'natural rights and obligations attaching to such relations. In this context the concept *natural* presents a problem and not a solution. To an unreflective mind what is customary in social relations usually appears natural; but no reflective person would present 'conformity to custom' as a basic moral principle; so we have to look for guidance in *selecting* the customary obligations that have moral force [and this, Sidgwick says, throws us back onto one of the other guides to basic morality.]

The more modern view of nature regards the organic world as exhibiting. . . .a continuous and gradual process of changing life; and this 'evolution', as the name implies, goes

no merely •from old to new but •from fewer to more of certain definite characteristics. But it would surely be absurd to *infer* that these characteristics are ultimately good and that our whole moral project should be to accelerate the arrival of an inevitable future!. . . .

Summing up: I don't think that any definition of *natural* shows this notion to be capable of providing an independent ethical first principle. (·For some concepts, the lack of such a definition is not a defect or drawback; for example· the notion of *beautiful* is indefinable yet clear, because it is derived from a simple unanalysable impression; but no-one maintains that the notion of *natural* is like that.) So I don't see how it could provide a definite practical criterion of the rightness of actions.

3. What emerges from that discussion is that the different views about the ultimate reason for doing what is concluded to be right don't all generate different methods of arriving at this conclusion; indeed, almost any method can be linked with almost any ultimate reason through some assumption. That's why it is hard to classify and compare ethical systems: the comparisons go differently depending on whether we go by •method or by •ultimate reason. I am taking difference of •method as the main consideration; and that's why I have treated the view that is the ultimate end as a variety of the intuitionism that fixes right conduct in terms of intuitively known axioms of duty, and have sharply separated •Epicureanism or egoistic hedonism from the •universalistic or Benthamite hedonism to which I propose to restrict the term 'utilitarianism'.

These two methods are commonly treated as closely connected, and it's easy to see why. •They both prescribe actions as means to an end that is distinct from the actions; so that they both lay down rules that are valid only if they conduce to the end. •The ultimate ends in both are the same in quality, namely pleasure—or strictly the greatest possible surplus of pleasure over pain. •And the conduct recommended by one principle largely coincides with that taught by the other. In a tolerably well-ordered community, intelligent self-interest nearly always leads to the fulfilment of most of one's social duties. And conversely, a universalistic hedonist can reasonably think that his own happiness is the part of the universal happiness that he is best placed to promote, and thus is especially entrusted to his charge. And the practical blending of the two systems is sure to go beyond their theoretical coincidence. It is easier for a man to zigzag between egoistic and universalistic hedonism than to be in practice a consistent follower of either. Few men are so completely selfish that they won't sometimes have a sympathetic impulse to promote the happiness of others, without basing this on any Epicurean calculation. And probably even fewer are so resolutely unselfish that they are never guilty of deciding rather too easily that something that *they* want is for 'the good of everybody'!

Bentham's psychological doctrine that everyone always *does* aim at his own greatest apparent happiness seems to imply that it is useless to show a man the conduct that would conduce to •the general happiness unless you convince him that it would also conduce to •his own. Hence on this view egoistic and universalistic considerations must be combined in any practical treatment of morality; so it was expectable that Bentham or his disciples would try to base the universalistic hedonism that they approve and teach on the egoism that they accept as inevitable. And so we find that Mill does try to connect the psychological and ethical principles that he shares with Bentham, and to convince his readers that because each man naturally seeks his own happiness, *therefore* he ought to seek other people's. I'll discuss this argument in III/13.

But I'm sure that the practical affinity between utilitarianism and intuitionism is really much greater than that between the two forms of hedonism. I'll defend this at length in subsequent chapters. Here I will just say this: many moralists who have •maintained as practically valid the rules that common-sense morality seems intuitively to come up with have nevertheless •regarded general happiness as an end to which those rules are the best means, and have •held that a knowledge of these rules was implanted by nature or revealed by God for the attainment of this end. On this view, though I'm obliged to act in conformity to a rule that is (for me) absolute, the natural or divine *reason for* the rule is utilitarian. This certainly rejects the *method of utilitarianism*, and it doesn't link right action and happiness through any process of reasoning. Without outright rejecting the *utilitarian principle*, it holds that the limitations of the human reason prevent it from properly grasping the real connection between the true principle and the right rules of conduct. But there has always been a considerable recognition by thoughtful people that obedience to the commonly accepted moral rules tends to make human life tranquil and happy. Even moralists, like Whewell, who are most strongly opposed to utilitarianism have been led to stress utilitarian considerations when trying to exhibit the 'necessity' of moral rules. [When a few lines back Sidgwick writes of 'the real connection between the true principle and the right rules of conduct', one would expect him to mean '. . . the true principle (whatever it is). . . ', but the rest of the paragraph shows that he expects to be understood to be referring to the utilitarian principle. It is puzzling that at this stage in the work he should label it as 'true'.]

During the first period of ethical controversy in modern England, after Hobbes's bold assertion of egoism had kicked off an earnest search for a philosophical basis for morality, utilitarianism appears in friendly alliance with intuitionism. When Cumberland declared that 'the common good[1] of all rational beings' is the end to which moral rules are the means, he wasn't trying to supersede the morality of common sense but rather to support it against the dangerous innovations of Hobbes. We find him quoted with approval by Clarke, who is commonly taken to represent an extreme form of intuitionism. And Shaftesbury in introducing the theory of a 'moral sense' never dreamed that it could ever lead us to act in ways that weren't clearly conducive to the good of the whole; and his disciple Hutcheson explicitly identified the promptings of the moral sense with those of benevolence. Butler seems to have been our first influential writer who dwelt on the discrepancies between •virtue as commonly understood and 'conduct most likely to produce a surplus of happiness over pain'.[2] When Hume presented utilitarianism as a way of explaining current morality, it was suspected to have a partly destructive tendency. But it wasn't until the time of Paley and Bentham that utilitarianism was presented as a method for determining conduct—a method that was to overrule all traditional precepts and supersede all existing moral sentiments. And even this final antagonism concerns theory and method rather than

[1] Neither Cumberland nor Shaftesbury uses the noun 'good' in an exclusively hedonistic sense. But Shaftesbury uses it mainly in this sense, and Cumberland's 'good' includes happiness along with perfection.

[2] See the second appendix, 'The Nature of Virtue', to his *Analogy of Religion*. There was a gradual change in Butler's view on this important point. In the first of his *Fifteen Sermons on Human Nature*, published a few years before the *Analogy*, he doesn't notice—any more than Shaftesbury and Hutcheson did—any possible lack of harmony between conscience and benevolence. But a note to the twelfth sermon seems to indicate a stage of transition between the view of the first sermon and the view of the appendix to the *Analogy*. [Each of the Butler passages referred to on this page can be found on the website from which the present text was taken.]

practical results; in the minds of ordinary folk that practical conflict is mainly between •self-interest and •social duty, however that is determined. From a practical point of view, indeed, the principle of aiming at the 'greatest happiness of the greatest number' seems to be more definitely opposed to egoism than the morality of common sense is. The latter seems to leave a man free to pursue his own happiness under certain definite limits and conditions, whereas utilitarianism seems to require that self-interest be *always* subordinate to the common good. And thus, as Mill remarks, utilitarianism is sometimes attacked from diametrically opposite directions: from a confusion with egoistic hedonism it is called base and groveling; while it is also, more plausibly, accused of setting a standard of unselfishness that makes exaggerated demands on human nature.

There's much more still to be said to clarify the principle and method of utilitarianism, but it seems best to defer that until I come to investigate its details. It will be convenient to take that as the final stage—Book IV—of my examination of methods. It will simplify things if egoistic hedonism (Book II) is discussed before universalistic hedonism; and we should have the pronouncements of intuitive morality (Book III) in as exact a form as possible before we compare them with the results of the more doubtful and difficult calculations of utilitarian consequences.

I'll try in the remaining chapters of Book I to remove certain unclarities concerning the general nature and relations of egoism and intuitionism, before examining them more fully in Books II and III.

[The chapter ends with a page-long note defending the attribution to Bentham of the doctrine Sidgwick is calling 'utilitarianism'. Writings of Bentham's that weren't published in his lifetime seem to show him favouring utilitarianism for public morality and egoistic hedonism for private morality, with the two being harmonized by an argument for the thesis—which Sidgwick thinks is wrong—that the best route to private happiness is through working for the happiness of everyone.]

Chapter 7: Egoism and self-love

1. I have been using the term •'egoism' in the usual way, as denoting a system that tells each person to act in pursuit of his own happiness or pleasure. The ruling motive here is usually called •'self-love'. But each of those terms can be taken in other ways, which we should identify and set aside before going on.

For example, the term 'egoistic' is ordinarily and not improperly applied to the basis on which Hobbes tried to construct morality, and which he regarded as the only firm grounding for the social order, enabling it to escape the storms and convulsions that the vagaries of the unenlightened conscience seemed to threaten it with. But the first of the precepts of rational egoism that Hobbes calls 'laws of nature', namely *Seek peace and follow it*, doesn't focus on the end of egoism as I have defined it—the greatest attainable pleasure for the individual—but rather on '**self**-preservation',

or perhaps on a compromise between the two,[1] as the ultimate end and standard of right conduct.

In Spinoza's view the (egoistic) principle of rational action is, as Hobbes thought, the impulse of **self**-preservation. He holds that everything, including the individual mind, does its best to stay in existence; indeed, this effort is a thing's very essence. It's true that the object of this impulse can't be separated from pleasure or joy, because pleasure or joy is 'a passion in which the soul passes to higher perfection'. But what the impulse primarily aims at is not •pleasure but •the mind's perfection or reality. These days we would call it **self**-realisation, ·and I now explain why·. The highest form of it, according to Spinoza, consists in a clear comprehension of all things in their necessary order as states of the one divine being—·which Spinoza calls 'God, i.e. Nature'·—and the willing *acceptance of everything* that comes from this comprehension; in this state the mind is purely active, with no trace of passivity; and thus in achieving this state the thing realises its essential nature to the greatest possible degree.

This is the notion of **self**-realisation as defined not only *by* but *for* a philosopher! It would mean something quite different in the case of a man of action whom the reflective German dramatist Schiller (in his play *Wallenstein*) introduces thus:

> I cannot,
> Like some hero of big words, like one who babbles
> of virtue,
> Get warmed by my will and my thoughts. . .
> If I no longer *work* I shall be nothing.

Many an artist sees his production of beautiful art as a realisation of **self**; and moralists of a certain kind down through the centuries have similarly seen the sacrifice of inclination to duty as the highest form of **self**-development, and held that true self-love prompts us always to obey the commands issued by the governing principle—reason or conscience within us—because in such obedience, however painful it is, we'll be realising our truest **self**.

So the term 'egoism', taken as implying only that first principles of conduct refer to **self**, doesn't imply anything about the content of such principles. Except when we're aware of a conflict between two or more of them, *all* our impulses—high and low, sensual and moral—are related to self in such a way that we tend to identify ourselves with each as it arises. So **self** can come to the fore when we are letting *any* impulse have its way; and egoism, considered as merely implying this fact, is common to all principles of action.

Someone might object:

> 'Properly understood, to "develop" or "realise" oneself is not merely to let one's currently predominant impulse have its way, but to exercise all the different faculties, capacities, and propensities of which our nature is made up, each in its due place and proper degree.'

But what is meant here by 'due proportion and proper degree'? [The switch from 'place' to 'proportion' is Sidgwick's.] **(a)** These terms may imply an ideal that the individual mind has to be trained up to by restraining some of its natural impulses and strengthening others, and developing its higher faculties rather than its lower. **(b)** Or they may merely refer to the combination and proportion of tendencies that the person is born with, the thesis being that he should do

[1] Thus the end for which an individual is supposed to renounce the unlimited rights of the state of nature is said to be 'nothing else but the security of a man's person in this life, and the means of preserving life so as not to be weary of it' (*Leviathan*, chapter 14).

his best to manage the situations he gets into and the functions he chooses to exercise in ways that enable him to 'be himself', 'live his own life', and so on. According to **(a)** rational self-development is just the pursuit of perfection for oneself: while **(b)** seems to present self-development not as an absolute end but rather as a means to happiness. If a man has inherited propensities that clearly tend to his own unhappiness, no-one would recommend him to develop these rather than modifying or subduing them in some way. Is it true that giving free play to one's nature is the best way to seek happiness? I'll address that question when I come to examine hedonism.

The upshot seems to be that the notion of self-realisation is too indefinite for use in a treatise on ethical method. And similarly we must discard a common account of egoism which describes its ultimate end as the 'good' of the individual, because the term 'good' can cover all possible views of the ultimate end of rational conduct. Indeed it may be said that egoism in this sense was assumed in the whole ethical controversy of ancient Greece: it was assumed by everyone that a rational individual[1] would make the pursuit of his own good his supreme aim, and the controversy all concerned the question of whether this good was rightly conceived as pleasure or virtue or some combination of those. [Sidgwick develops this theme at some length, remarking that even Aristotle's view that the desirable end is *Eudaimonia* is open to different interpretations, as is 'happiness', the English word by which *Eudaimonia* usually translated. On this word he remarks:] It seems to be commonly used in Bentham's way, as equivalent to 'pleasure' or rather as denoting something whose constituents are all pleasures; and that's the sense in which I think it is most convenient to use it. Sometimes in ordinary talk 'happiness' is used to denote a special kind of agreeable consciousness, which is calmer and more indefinite than specific pleasures such as the gratifications of sensual appetite or other sharp and urgent desires. We could call it the feeling that accompanies the normal activity of a 'healthy mind in a healthy body'; specific pleasures seem to be stimulants of it rather than parts of it. Sometimes, though with a more obvious departure from ordinary language, 'happiness' or 'true happiness' is understood in a definitely non-hedonistic sense as referring to something other than any kind of agreeable feeling.[2]

2. To be clear, then, I specify that the object of self-love and the goal of egoistic hedonism is

> pleasure, in the widest sense of the word—including *every kind* of delight, enjoyment, or satisfaction, except for kinds that are incompatible with some greater pleasures or productive of pain.

That's how self-love seems to be understood by Butler and other English moralists after him—as a desire for one's own pleasure •generally, and for as much of it as possible, whatever its source is. (Butler in the eleventh of his *Fifteen Sermons* writes of 'the cool principle [see Glossary] of self-love or •general desire for our own happiness'.) In fact, the 'authority' and 'reasonableness' attributed to self-love in Butler's system are based on this •generality and comprehensiveness. . . . When conflicting impulses compete for control of the will, our desire for pleasure in general leads us to compare the impulses in terms of the pleasures we think

[1] I'll try later on to explain how it comes about that in modern thought the proposition 'My own good is my only reasonable ultimate end' is not a mere tautology even if we define 'good' as 'what it is ultimately reasonable to aim at'. See chapter 9 and III/13–14.

[2] [Sidgwick adds a note to this, criticising T. H. Green for such a departure, and also Mill, though in Mill's case he thinks it is merely 'looseness of terminology, excusable in a treatise aiming at a popular style'.]

each offers, and to go with the one that offers most—which happens because that impulse is reinforced by our self-love, i.e. our desire for pleasure in general. So self-love is called into play whenever impulses conflict, and this—as Butler argues—involves it in regulating and directing our other springs of action. On this view, self-love makes us merely consider the *amount* of pleasure or satisfaction—as Bentham put it: 'Quantity of pleasure being equal, push-pin is as good as poetry.'

But many people find this offensively paradoxical; and Mill in developing Bentham's doctrine in *Utilitarianism* chapter 2, thought it desirable to abandon it by bringing in *quality* of pleasure as well as *quantity*—·i.e. to allow for better/worse pleasures as well as for greater/lesser ones·. Two points to note here: **(i)** It is quite consistent with Bentham's view to describe some kinds of pleasure as better than others, if by 'a pleasure' we mean (as is often meant) a whole state of consciousness that is only partly pleasurable; and still more if we take account of subsequent states. For many pleasures are accompanied by pain and many more have painful consequences. In valuing such an 'impure pleasure' (Bentham's word) we must give a negative value to the pain; so this is a strictly •quantitative basis for saying that such a pleasure is inferior in •kind. **(ii)** We mustn't confuse intensity of *pleasure* with intensity of *sensation*: a pleasant feeling may be strong and absorbing and yet not as pleasant as another that is more subtle and delicate. Given these two points, I think that to work out consistently the method that takes pleasure as the sole ultimate end of rational conduct we must accept Bentham's proposition, and regard all •qualitative comparisons of pleasures as really being •quantitative. For all the items called 'pleasures' are supposed to have a common property, *pleasantness*, and may be compared in respect of this common property. So if what we are seeking is pleasure as such, and nothing else, it's obvious that we must always prefer a more pleasant pleasure to a less pleasant one. No other choice seems reasonable unless we are aiming at something besides pleasure. And when we say that one kind of pleasure is better than another—e.g. that the pleasures of mutual love are superior in quality to the pleasures of gratified appetite—we mean that they are more pleasant. We could of course mean something else; we could for instance mean that they are nobler and more elevated, although less pleasant. But then we are clearly introducing a non-hedonistic ground of preference, and that involves a method that is a puzzling mixture of intuitionism and hedonism.

To sum up: If by 'egoism' we merely understand a method that aims at self-realisation, it seems to be a form that almost any ethical system can have without changing its essential nature. And even when further defined as egoistic *hedonism*, it is still not clearly distinct from intuitionism if it allows •quality of pleasures to over-rule •quantity. What is left is pure or quantitative egoistic hedonism: this method •is essentially distinct from all the others and •is widely held to be rational; so it seems to deserve a detailed examination. According to it, the only thing a rational agent regards as important in deciding what to do is quantity of consequent pleasure and pain to himself; and he always seeks the greatest attainable surplus of pleasure over pain—which we can appropriately call his 'greatest happiness'. This view and attitude of mind seems to be what is most commonly meant by the vaguer terms 'egoism' and 'egoistic'; so I shall allow myself to use these terms with this more precise meaning.

Chapter 8: Intuitionism

1. I have used the term 'intuitional' [see Glossary] to denote the view that the practically ultimate end of moral actions is *conformity to certain rules or dictates*[1] *of duty that are unconditionally prescribed*. But current ethical discussion reveals a considerable ambiguity in the terms 'intuition', 'intuitive' and their cognates, and we must now try to clear it up. Writers who maintain that we have 'intuitive knowledge' of the rightness of actions usually mean that an action it found to be right simply by 'looking at' •the action itself without considering •its consequences. This can't be meant for the whole range of duty, because there has never been a morality that didn't pay some attention to consequences. Prudence or forethought has commonly been regarded as a virtue; and all modern lists of virtues have included *rational benevolence*, which aims at the happiness of other human beings generally, and therefore has to consider even remote effects of actions. Also, it's hard to draw the line between an act and its consequences, because the effects resulting from each of our volitions form a continuous indefinitely long series, and we seem to be conscious of causing any of these effects that we foresee as probable at the moment of volition. However, in the common notions of different *kinds* of actions a line is actually drawn between •the results included in the notion and regarded as forming part of the act and •the results considered as its consequences. ·Take for example the action-kind *speaking truth*·. In speaking truth to a jury I may foresee that in the given situation my words are sure to lead them to a wrong conclusion about the guilt or innocence of the accused, but also sure to give them the right conclusion about the particular matter of fact that I am testifying about; and we would commonly consider the latter foresight or intention to fix the nature of my act as an •act of truth-telling, while the former merely relates to a •consequence. So we have to take it that the disregard of consequences that the intuitional view is here taken to imply relates only to certain specific kinds of action (such as truth-speaking) for which common usage settles which events are included in the general notions of the acts and which are merely consequences.

Men can and do judge remote as well as immediate results to be good, and such as we should try to bring about, without bringing in the feelings of sentient beings. I have taken this to be the view of those whose ultimate end is the general perfection—as distinct from the happiness—of human society; and it seems to be the view of many who concentrate their efforts on results such as the promotion of art or of knowledge. Such a view, if explicitly distinguished from hedonism, might properly be labelled as 'intuitional', but in a sense broader than the one defined at the start of this chapter. The point of calling such a view 'intuitional' is that according to it

> the results in question—·perfection, art, knowledge, whatever·—are judged to be good •immediately and not •by inference from experience of the pleasures that they produce.

So we have to admit a broader use of 'intuition', as meaning 'immediate judgment about to what ought to be done or aimed at'. But these days when writers contrast 'intuitive' or *a priori* with 'inductive' or *a posteriori* morality, there is often a confusion at work:

[1] I use the term 'dictates' to include the view that the ultimately valid moral imperatives are conceived as relating to particular acts.

- •What the 'inductive' moralist claims to know by induction is usually the conduciveness to pleasure of certain kinds of action.
- •What the 'intuitive' moralist claims to know by intuition is usually their rightness.

So there is no proper opposition. If hedonism claims to give authoritative guidance, it must be through the principle that pleasure is the only reasonable ultimate end of human action; and *this* can't be known by induction from experience. Experience could only tell us is that all men always do, not that they ought to, seek pleasure as their ultimate end (I have already tried to show that in fact it doesn't even show that they *do*). If this latter proposition—·that men ought to seek pleasure as their ultimate end·—is right as applied to private or general happiness, it must either be immediately known to be true—which makes it a moral intuition—or be inferred from premises including at least one such moral intuition; so either species of hedonism. . . .can legitimately be said to be *in a certain sense* 'intuitional'. But the prevailing opinion of ordinary moral persons, and of most writers who have maintained the existence of moral intuitions, is that certain kinds of actions are unconditionally prescribed without regard to their consequences; so I shall treat that doctrine as what marks off the intuitional method during the main part of my detailed examination of that method in Book III.

2. The common antithesis between 'intuitive' and 'inductive' morality is misleading in another way, because a moralist may •hold that an action can be known to be right without bringing in the pleasure produced by it, while •using a method that can properly called 'inductive'. For he may hold that, just as the generalisations of physical science rest on particular observations, so in ethics general truths can be reached only by induction from judgments or perceptions of the rightness or wrongness of particular acts.

That's what Aristotle was talking about when he said that Socrates applied inductive reasoning to ethical questions. Having discovered that men (including himself) used general terms confidently, without being able to explain their meanings, Socrates worked towards the true definition of each term by examining and comparing different instances of its application. Thus the definition of 'justice' would be sought by looking for a general proposition that fits all the different actions commonly judged to be just.

In the plain man's view of *conscience* it seems to be often implied that particular judgments are the most trustworthy. 'Conscience' is the man-in-the-street's label for the faculty of moral judgment as applied to one's own acts and motives; and we usually think of the dictates of conscience as relating to particular actions. When someone is told in a particular case to 'trust to his conscience', what usually seems to be meant is that he should form a moral judgment on this case without relying on general rules, and even in opposition to what follows by deduction from such rules. It's this view of conscience that most easily justifies the contempt often expressed for 'casuistry'; for if the particular case can be settled by conscience without bringing in general rules, then casuistry (i.e. the application of general rules to particular cases) is at best superfluous. But then on this view we'll have no •practical need for any such general rules. . . . We could of course form general propositions by induction from particular judgments of conscience, and arrange them systematically; but this would have •theoretical interest only. This may explain why some conscientious people are indifferent or hostile to systematic morality: they feel that they at least can do without it; and they fear that the cultivation of it may have an outright bad effect on the proper development of the practically important faculty that

is at work in particular moral judgments.

This view may be called, in a sense, 'ultra-intuitional', because in its most extreme form it recognises nothing but •simple immediate intuitions and sees no need for any kind of •reasoning to moral conclusions. We can see in it **one phase or variety of the intuitional method**, if the term 'method' can be stretch to cover a procedure that is completed in a single judgment.

3. ·Three reasons for not relying only on conscience:· **(i)** Probably all moral agents have experience of such particular intuitions, which are a large part of the moral phenomena of most minds; but relatively few people are so thoroughly satisfied with them that they don't feel a practical need for some further moral knowledge. For thoughtful people don't experience these particular intuitions as being beyond question; and when they have sincerely put an ethical question to themselves, they don't always find that their conscience gives them a clear immediate insight into the answer. **(ii)** When a man compares the utterances of his conscience at different times, he often finds it hard to reconcile them; the same conduct will have a different moral look at one time from what it had earlier, although the person's knowledge of its circumstances and conditions hasn't relevantly changed. **(iii)** We find that the moral perceptions—·the deliverances of conscience·—of different minds frequently conflict, though the minds seem equally competent to judge. These three factors create serious doubts about the validity of each man's •particular moral judgments; and we try to set these doubts at rest by appealing to general rules that are more firmly established on a basis of •common consent.

The view of conscience that I have been discussing is suggested by much untutored talk, but it's not the one that Christian and other moralists have usually given. They have likened the process of conscience to jural [see Glossary] reasoning such as what goes on in a court of law. Here we start with a system of universal rules, and a particular action has to be brought under one of them before it can be judged to be lawful or unlawful. Now an individual person can't learn the rules of positive law [see Glossary] by using his reason; this may teach him that law ought to be obeyed, but he has to learn what the law *is* from some external authority. And this is quite often what happens with the consciences of ordinary folk when some dispute or difficulty forces them to reason: they want to obey the right rules of conduct, but they can't see for themselves what these are, and have to consult their priest, or their sacred books, or perhaps the common opinion of the society they belong to. When that's what happens we can't strictly call their method 'intuitional': they haven't intuitively apprehended the rules they are following. Other people (perhaps *all* to some extent) do seem to see for themselves the truth and bindingness of such current rules. They may use 'common consent' as an argument for the rules' validity, but only as supporting the individual's intuition, not replacing it.

So this is **a second intuitional method**, which assumes that we can discern certain general rules with really clear and finally valid intuition. It involves the following theses:

- •General moral rules are implicit in the moral reasoning of ordinary men, who grasp them adequately for most practical purposes and can state them roughly.
- •To state them with proper precision requires a special habit of contemplating abstract moral notions clearly and steadily.
- •The moralist's job then is to do this abstract contemplation, to arrange the results as systematically as possible, and by proper definitions and explanations to remove vagueness and prevent conflict.

That's the kind of system that seems to be generally intended

when 'intuitive' or *a priori* morality is mentioned. It will be my main topic in Book III. [Strictly speaking, III/3–10 will be mainly concerned with trying to establish what common-sense morality *is*; the attempt to systematise it will come in III/11.]

4. Philosophic minds, however, often find the morality of common sense, even when made as precise and orderly as possible, to be unsatisfactory as a system, although they aren't disposed to question its general authority. They can't accept as scientific first principles the moral generalities that we reach by reflecting on the ordinary thought of mankind (ourselves included). Even if these rules can be formulated so that they cover the whole field of human conduct, answering every practical question without coming into conflict, still the resulting code looks like a jumble that stands in need of some rational synthesis. Without denying that conduct commonly judged to be right *is* right, we may want some deeper explanation of *why* it is so. This demand gives rise to **a third species or phase of intuitionism**. This accepts the morality of common sense as mainly sound, but tries to give it a philosophic basis that it doesn't itself offer—to get one or more absolutely and undeniably true principles from which the current rules might be deduced, either just as common sense has them or with slight modifications and rectifications.[1]

Those three phases of intuitionism can be seen as three stages in the formal development of intuitive morality, which we could call

- •perceptional,
- •dogmatic, and
- •philosophical

respectively. I have defined the third of them only in the vaguest way; in fact, I have presented it only as a problem and we can't foresee how many solutions of it may be attempted; but before investigating it further—·as I shall do in III/13·—I want to examine in detail the morality of common sense.

It must not be thought that these three phases are sharply distinguished in the moral reasoning of ordinary men; but then no more is intuitionism of any sort sharply distinguished from either species of hedonism. The commonest type of moral reasoning is a loose combination or confusion of methods. Probably most moral men believe •that their moral sense or instinct will guide them fairly rightly in any particular case, but also •that there are general rules for determining right action in different kinds of conduct; and •that these in turn can be given a philosophical explanation that deduces them from a smaller number of basic principles. Still for systematic direction of conduct, we require to know on what judgments we are to rely as ultimately valid. [That last sentence is as Sidgwick wrote it.]

I have been focusing on differences in intuitional method due to difference of generality in the intuitive beliefs recognised as ultimately valid. There's another class of differences arising from different views about the precise quality that is immediately apprehended in the moral intuition; but these are especially subtle and hard to pin down clearly and precisely, so I'll give them a chapter of their own.

* * * * *

NOTE

Intuitional moralists haven't always been careful to make clear whether they regard as ultimately valid •moral judgments on single acts, or •general rules prescribing particular

[1] It may be that such principles are not 'intuitional' in the narrower sense that excludes consequences but only in the broader sense as being self-evident principles relating to 'what ought to be'.

kinds of acts, or •more universal and fundamental principles. For example, Dugald Stewart [1753–1828] calls the immediate operation of the moral faculty 'perception'; but when he describes what is thus 'perceived' he always seems to have in mind general rules.

Still we can pretty well sort out English ethical writers into •those who have confined themselves mainly to defining and organising the morality of common sense and •those who have aimed at a more philosophical treatment of the content of moral intuition. We find that the distinction mainly corresponds to a difference of periods and—perhaps surprisingly—that the more philosophical school is the earlier. We can partly explain this by attending to the doctrines that the intuitional method was opposing in the various periods. In the first period all orthodox moralists were occupied in refuting Hobbism ·as presented in *Leviathan* [1651]·. But this system, though based on materialism and egoism, was intended as ethically constructive. Accepting the commonly received rules of social morality, it explained them as the conditions of peaceful existence that enlightened self-interest told each individual to obey, provided that the social order to which he belonged was an *actual* one with a strong government. This certainly makes the theoretical basis of duty seriously unstable; still, assuming a decently good government, Hobbism can claim to explain and establish the morality of common sense, not to undermine it. And therefore, though some of Hobbes's antagonists (such as Cudworth [1617–88]) settled for simply reaffirming the absoluteness of morality, the more thoughtful ones felt that system must be met by system and explanation by explanation, and that they must penetrate beyond the dogmas of common sense to some more solid certainty. And so, while Cumberland [1631–1718] found this deeper basis in the notion of 'the common good of all rational beings' as an ultimate end, Clarke [1675–1729] tried to exhibit the more basic of the commonly accepted rules as perfectly self-evident axioms that are *forced* on the mind when it contemplates human beings and their relations. But Clarke's results were not found satisfactory; and the attempt to exhibit morality as a body of scientific truth gradually fell into discredit, and the stress moved over into the **emotional** side of the moral consciousness. But when ethical discussion thus passed over into psychological analysis and classification, the idea of duty as objective, on which the authority of moral sentiment depends, gradually slipped out of sight. For example, we find Hutcheson [1694–1746] asking why the moral sense shouldn't vary in different human beings, as the palate does; he didn't dream that there's any peril to morality in such a view. But when the new doctrine was endorsed by the dreaded name of Hume [1711–76], its dangerous nature was clearly seen, as was the need to bring the **cognitive** element of moral consciousness back into prominence; and this work was undertaken as a part of the general philosophical protest of the Scottish school against the empiricism that had culminated in Hume. But this school claimed as its characteristic merit that it met empiricism on its own ground, and revealed assumptions that the empiricist repudiated among the facts of psychological experience that he claimed to observe. And thus in ethics the Scottish school was led to expound and reaffirm the morality of common sense, rather than offering any deeper principles that couldn't be so easily supported by an appeal to common experience.

Chapter 9: Good

1. Up to here I have spoken of the quality of conduct discerned by our moral faculty as 'rightness', which is the term commonly used by English moralists. I have regarded this term, and its everyday equivalents, as implying the existence of a dictate or imperative of reason that *prescribes* certain actions—either unconditionally or with reference to some end.

But there's a possible view of virtuous action in which, though the validity of moral intuitions isn't disputed, this notion of rule or dictate is at most only latent or implicit, the moral ideal being presented as *attractive* rather than *imperative*. That view seems to be at work when an action or quality of character is judged to be 'good' in itself (and not merely as a means to some further good). This was the basic ethical conception in the Greek schools of moral philosophy generally; and that includes even the Stoics, though the prominence their system gives to the conception of *natural law* makes the system a transitional link between ancient and modern ethics. And this historical illustration brings out an important result of substituting *goodness* for *rightness*—which might at first sight seem to be a merely verbal change. What mainly marks off ancient ethical controversy from modern comes from their use of a •generic notion instead of a •specific one in expressing the common moral judgments on actions. Virtue, or right action, was commonly regarded among the Greeks as only a species of the good; and so, on this view of what the basic moral input is, the first question that offered itself when they were trying to systematise conduct, was: What is the relation of this species of good to the rest of the genus? This was the question that the Greek thinkers argued about, from first to last. To understand their speculations we have to make the effort to throw aside the quasi-jural notions of modern ethics, and ask (as they did) not 'What is duty and what is its ground?' but 'Which of the objects that men think good is truly good or the highest good?' or, in the more specialised form of the question that the moral intuition introduces, 'How does the kind of good we call "virtue"—the qualities of conduct and character that men commend and admire—relate to other good things?'

That's one difference between the two forms of intuitive judgment. The recognition of conduct as *right* involves an authoritative prescription to do it; but when we have judged conduct to be *good* it's not yet clear that we ought to prefer it to all other good things; we have to find some standard for estimating the relative values of different goods.

So I'm going to examine the notion *good* across the whole range of its application. Because what we want is a standard for comparing the constituents of *ultimate* good, we aren't directly concerned with anything that is good only as a means to some end. Indeed, if we were considering *only* good-as-a-means, we could plausibly to interpret 'good' without reference to human desire or choice, as meaning merely 'fit' for the production of certain effects—a good horse for riding, a good gun for shooting, etc. But because we apply the notion of *good* also to ultimate ends, we must look for a meaning for it that will cover both applications.

2. Many people maintain that whenever we judge something to be 'good' we are implicitly thinking of it as a means to pleasure, even when we don't explicitly refer to this or any other end. On this view, comparing things in respect of their 'goodness' is really comparing them as sources of pleasure;

so that any attempt to systematise our intuitions of goodness, whether in conduct and character or in other things, must reasonably lead us straight to hedonism. And indeed in using 'good' outside the sphere of character and conduct, and not in application to things that are definitely regarded as means to some specific end—·'good knife', 'good candle' or the like·—our recognition of things as in themselves 'good' is closely correlated with our expectation of pleasure from them. The good things of life are things that give sensual or emotional pleasure—good dinners, wines, poems, pictures, music—and this gives *prima facie* support to the interpretation of 'good' as meaning 'pleasant'. But this isn't clearly supported by common sense. To see this, think about the application of 'good' to the cases most analogous to that of conduct, namely to what we may call 'objects of taste'. Granted that the judgment that something is good of its kind is closely connected with the expectation of pleasure derived from it, but it is usually to a specific *kind* of pleasure; and if the object happens to give us pleasure of a different kind, that doesn't lead us to call it 'good', or anyway not without some qualification. We wouldn't call a wine 'good' solely because it was very wholesome; or call a poem 'good' on account of its moral lessons! So when we consider the meaning of 'good' as applied to conduct, we've been given no reason to suppose that it refers to or corresponds with *all* the pleasures that may result from the conduct. Rather, the perception of goodness or virtue in actions seems to be like the perception of beauty[1] in material things; which is normally accompanied by a specific pleasure that we call 'aesthetic', but often has no discoverable relation to the general usefulness or agreeableness of the thing discerned to be beautiful; indeed, we often recognise this kind of excellence in things that are hurtful and dangerous.

Another point about aesthetic pleasures and aesthetic goodness: it is generally accepted that some people have more 'good taste' than others, and it's only the judgment of people with good taste that we recognise as valid regarding the real goodness of the things we enjoy. Each person is the final judge of what gives him pleasure; but for something to qualify as *good* it must satisfy some universally valid standard, and we get an approximation to this, we believe, from the judgment of those to whom we attribute good taste. And in this context 'good' clearly doesn't mean 'pleasant'; it conveys an aesthetic *judgment* that •is answerable to some standard and •is just *wrong* if it deviates from that standard. And the person with the best taste isn't always the one who gets the most enjoyment from good and pleasant things. Connoisseurs of wines, pictures, etc. often retain their intellectual ability to appraise and rank-order the objects that they criticise, even when their capacity for getting pleasure from these objects is blunted and exhausted. Indeed, someone whose feelings are full and fresh may get more pleasure from inferior objects than a connoisseur gets from the best.

[1] But we must distinguish the idea of moral goodness from that of beauty as applied to human actions; although they have much in common and have often been identified, especially by the Greek thinkers. In some cases the ideas are indistinguishable, and so are the corresponding pleasurable emotions; a noble action affects us like a scene, a picture, or a strain of music; and the depiction of human virtue is an important part of the artist's means for producing his special effects. But look closer: much good conduct isn't beautiful, or anyway doesn't sensibly impress us as such; and certain kinds of wickedness have a splendour and sublimity of their own. For example, a career such as Cesare Borgia's, as Renan says, is *beau comme une tempête, comme un abîme*—·beautiful like a tempest, like the Grand Canyon·. It's true that in all such cases the beauty comes from the conduct's exhibiting striking gifts and excellences mingled with the wickedness; but we can't sift out the wickedness without spoiling the aesthetic effect. So I think we have to distinguish the sense of beauty in conduct from the sense of moral goodness.

To sum up: the general admission that things called 'good' are productive of pleasure, and that goodness is inseparable in thought from pleasure, doesn't imply that ordinary estimates of the goodness of conduct are really estimates of the amount of pleasure resulting from it; ·and there are two reasons for thinking that they aren't·. **(i)** Analogy would lead us to conclude that attributing goodness to conduct, as to objects of taste generally, doesn't correspond to *all* the pleasure that is caused by the conduct but to some kind of pleasure—specifically the satisfaction a disinterested spectator would get from contemplating the conduct. [This is Sidgwick's first mention of the 'disinterested spectator' (a central figure in Adam Smith's *Theory of the Moral Sentiments*), and it's the last except for some sympathetic remarks on page 225. He apparently thinks of the disinterested spectator as relating to the moral evaluation of conduct in the way a connoisseur relates to the aesthetic valuation of music or painting or wine.] **(ii)** And it may not arouse even this specific pleasure in proportion to its goodness, but only (at most) in people who have good moral taste; and even with them we can distinguish •the intellectual apprehension of goodness from •the pleasurable emotion that often accompanies it.

When we pass from the *adjective* to the *noun* 'good', we see straight away that the noun can't be regarded as synonymous with 'pleasure' or 'happiness' by anyone—e.g. a hedonist—who •maintains that the pleasure or happiness of human beings is their ultimate good persons and who •takes this to be a significant proposition and not a mere tautology. This obviously requires that the two terms have different meanings; and the same presumably holds for the corresponding adjectives.

3. Then what *are* we to say about the general meaning of 'good'? Shall we say with Hobbes and his followers that 'whatever is the object of any man's desire is what he calls "good", and the object of his aversion is what he calls "evil"'? To simplify the discussion I'll consider only what a man desires

- •for itself, and not as a means to some end, and
- •for himself, and not benevolently for others.

—i.e. his own good[1] and ultimate good. First, there's the obvious objection: a man often desires what he knows is bad for him; the pleasure of drinking champagne that is sure to disagree with him, the gratification of revenge when he knows that his true interest lies in reconciliation. The answer is this:

> In such cases the desired result is accompanied or followed by other effects which, when they occur, arouse aversion stronger than the original desire. But although these bad effects are *foreseen* they are not *forefelt*; the thought of them doesn't adequately modify the predominant direction of present desire.

But now focus solely on the desired result, setting aside things that accompany it or follow from it; it *still* seems that what is desired at any time is merely an apparent good that may not be found good when the time comes to enjoy it. It may turn out to be a 'Dead Sea apple' [see Glossary], mere dust and ashes in the eating; more often, having it will partly correspond to expectation but fall significantly short of it. And sometimes—even while yielding to the desire—we're aware of the illusoriness of this expectation of 'good' that the desire carries with it. Thus, if we are to conceive of

[1] The common view of *good* seems to imply that sometimes an individual's sacrifice of his own over-all good would bring about greater good for others. Whether such a sacrifice *is* ever required, and whether (if it is) it's truly reasonable for the individual to sacrifice his own over-all good, are among the deepest questions of ethics; and I shall carefully consider them later on (especially in III/14 [page 191]. At present I want only to avoid any prejudgment of these questions in my definition of 'my own good'.

the elements of ultimate good as capable of quantitative comparison—as we do when we speak of preferring a 'greater' good to a 'lesser'—we can't identify the object of desire with 'good' but only with 'apparent good'.

Also, a prudent man will do his best to suppress desires for things that he thinks he can't achieve through his own efforts—fine weather, perfect health, great wealth or fame, etc.—but any reduction in the intensity of such desires has no effect in leading him to judge the desired objects as less 'good.'

If we are to interpret *good* in terms of *desire*, therefore, we must identify it not with what is *desired* but with what is *desirable*. And in this context 'x is desirable for person y' doesn't mean 'y ought to desire x' but rather

> 'y *would* desire x if he thought he could achieve it by his own efforts and if he had perfect foresight and forefeeling into what it would be like to achieve it'.

But that is still not right as an account of 'x is over-all good, or good on the whole, for y'; because even if what is chosen turns out to be just how it was imagined when it was desired, it may be over-all bad because of what accompanies or results from it. So we have to vary the formula displayed above by somehow limiting our view to desires that affect conduct by leading to volitions; because I might regard something as desirable while judging it to be on the whole imprudent to aim at it. But even with this limitation, the relation of my 'good on the whole' to my desire is very complicated. It isn't right to say that my good on the whole is

> what I would actually desire and seek if at that time I knew in advance and adequately imagined in advance what it would be like to have it.

If we are rational, our concern for a moment of our conscious experience won't be affected by the moment's position in time; ·so when a man is wondering whether to do x, thoughts and feelings that he expects to have later on should be given their due weight, and not discounted because they are off in the future·. But the mere fact that a man doesn't afterwards feel for the consequences of an action a strong enough aversion to make him regret it doesn't prove for sure that he has acted for his 'good on the whole'. Indeed, we often count it among the worst consequences of some kinds of conduct that they alter men's desires, making them prefer their lesser good to their greater; and we think it all the *worse* for a man. . . .if he is never roused out of such a condition and lives the life of a contented pig, when he could have been something better. So we have to say that a man's future *good on the whole* is **what he would now desire and seek on the whole, if all the consequences of all the different lines of conduct open to him were accurately foreseen and adequately ·forefelt, i.e.· realised in imagination at the present point of time**.

This is such an elaborate and complex conception that it's hardly believable that this is what we commonly *mean* when we talk of what is 'good on the whole' for someone. Still, this **hypothetical object of a resultant desire** provides an intelligible and admissible interpretation of the terms 'good' (noun) and 'desirable', as giving philosophical precision to the vaguer meaning they have in ordinary discourse; and it would seem that a calm comprehensive desire for 'good' conceived somewhat in this way, though more vaguely, is normally produced by intellectual comparison and experience in a thoughtful mind. This notion of *good* has an ideal element; it's something that isn't always actually desired and aimed at by human beings; but the ideal element is entirely interpretable in terms of actual or hypothetical *fact*, and doesn't introduce any value-judgment, let alone any 'dictate of reason'.

But it seems to me more in accordance with common sense to recognise as Butler does that the calm desire for my 'good on the whole' is *authoritative*, carrying with it implicitly a rational dictate ·or instruction· to aim at this end if a conflicting desire is urging the will in an opposite direction. Still we can keep the notion of 'dictate' or 'imperative' merely implicit and latent (as it seems to be in ordinary thoughts about 'my good' and its opposite) by interpreting '**ultimate good on the whole for me**' to mean

> what I would practically desire if my desires were in harmony with reason, assuming my own existence alone to be considered.

On this view, '**ultimate good on the whole**' with no reference to any particular subject ('me') must be taken to mean

> what as a rational being I would desire and try to bring about, assuming myself to have an equal concern for all existence.

When conduct is judged to be' good' or 'desirable' in itself, independently of its consequences, I think that this latter point of view is being taken. Such a judgment differs from the judgment that conduct is 'right' because it doesn't involve a definite order to perform it, since it leaves open the question of whether this particular kind of good is the greatest good we can obtain under the circumstances. And there's another difference: calling an action 'good' or 'excellent' doesn't imply that it is in one's power. . . .in the same strict sense as calling it 'right' does; and in fact there are many excellences of behaviour that we can't achieve by any effort of will, at least directly and at the moment. That's why we often feel that recognising goodness in someone else's conduct doesn't involve a clear precept to do likewise, but rather 'the vague desire that stirs an imitative will' [quoted from Tennyson's *In Memoriam*]; and why goodness of conduct becomes a distant end that is beyond the reach of immediate volition.

4. A final question: When conduct or character is intuitively judged to be good in itself,[1] how are we to compare this value with the value of other good things? I shan't now try to establish a standard for such comparisons; but we can limit considerably the range of comparison for which it is required. That's because when we judge something x to be good, where x isn't a quality of human beings, it always turns out that x's goodness has some relation to •human existence or at least to •some consciousness or feeling.[2]

For example, we often judge some inanimate objects—river-banks, hillsides, etc.—to be good because beautiful, and others bad because ugly; but no-one would think it made sense to aim at the production of beauty in external nature apart from any possible human experience of it. When beauty is said to be 'objective', what is meant is not that it exists as beauty out of relation to any mind, but only that there's some standard of beauty that is valid for all minds.

[Sidgwick remarks that a man might try to make beautiful things without having any thought of who if anyone might enjoy them, or pursue knowledge without caring about who is to possess it. He continues:] Still, I think it will be generally

[1] Character is only known to us through its manifestation in conduct; and I don't think that in our common recognition of virtue as having value in itself we distinguish character from conduct. Is character to be valued for the sake of the conduct that expresses it, or is conduct to be valued for the sake of the character that it exhibits? We don't ordinarily give any thought to this question. In III/2.2 and III/14.1 I'll consider how it should be answered.

[2] There's a point of view. . . .from which the whole universe and not merely a certain condition of sentient beings is seen as 'very good'. . . . But such a view can scarcely be developed into a method of ethics. For practical purposes we need to see some parts of the universe as less good than they might be; and we have no basis for saying this of any portion of the non-sentient universe considered in itself and not in relation to sentient beings.

held that beauty, knowledge, and other ideal goods—as well as all external material things—are reasonably to be sought by men only if they conduce either **(1)** to happiness or **(2)** to the perfection or excellence of human existence. I say 'human' because although most utilitarians consider the pleasure (and freedom from pain) of the lower animals to be included in the happiness that they take to be the proper end of conduct, no-one seems to contend that we ought to aim at perfecting brutes, except as a means to our ends or at least as objects of scientific or aesthetic contemplation for us. Nor can we include as a practical end the existence of beings above the human. We certainly apply the idea of *good* to God, just as we do to His work; and when it is said that 'we should do all things to the glory of God', this may seem to imply that God's existence is made better by our glorifying Him. But this inference appears somewhat impious, and theologians generally recoil from it, and don't base any human duty on the notion of a possible addition to the goodness of the Divine Existence. As for the influence of our actions on other extra-human intelligences, this can't at present be a topic of scientific discussion.

So I am confident in saying that if there is any good other than happiness to be sought by man as an ultimate practical end, it can only be the goodness, perfection, or excellence of human existence.

- What does this notion include in addition to virtue?
- What is its precise relation to pleasure?
- If we accept it as fundamental, what method ·of ethics· will that logically lead us to?

It will be more convenient to discuss these questions after our detailed examination of these two other notions, pleasure (Book II) and virtue (Book III).

BOOK II: Egoistic hedonism

Chapter 1: The principle and method of hedonism

1. My aim in this Book is to examine the method of determining reasonable conduct that I have already sketched under the name of 'egoism', using this term as short-hand for 'egoistic hedonism'—the thesis that the ultimate end of each individual's actions is his own greatest happiness. Ought this to be counted among the received 'methods of *ethics*'? There are strong grounds for holding that simple egoism can't be a basis on which to construct a system of morality that is satisfactory to the moral consciousness of mankind in general. In chapter 3.2 and chapter 5 I shall carefully discuss these reasons. At present I'll just point to the wide acceptance of the principle that it's reasonable for a man to act in the way that does most for his own happiness. It is explicitly accepted by leading proponents of intuitionism and of utilitarianism (which is my name for universalistic hedonism). I have already noticed [page 5] that Bentham, although he regards the greatest happiness of the greatest number as the 'true standard of right and wrong', regards it as 'right and proper' that each individual should aim at his own greatest happiness. And Butler is equally prepared to grant that

> 'Our ideas of happiness and misery are nearer and more important to us than any of our other ideas. . . . Virtue or moral rectitude does indeed consist in affection towards and pursuit of what is right and good, as such; but let us admit that when we sit down in a cool hour we can't justify to ourselves this or any other pursuit until we are convinced that it will be for our happiness, or at least not contrary to it.' [*Fifteen Sermons*, no. 11]

And even Clarke in his *Boyle Lectures*, despite having emphatically maintained that 'virtue truly deserves to be chosen for its own sake and vice to be avoided', still admits that it is 'not truly reasonable that men by •adhering to virtue should give up their lives, if by doing this they eternally deprive themselves of all possibility of getting any benefit from that •adherence'.

And, generally, in Christian times it has been obvious and natural to hold that achieving virtue is essentially an enlightened pursuit of happiness for the agent. And this has been held not only by coldly calculating people but also—emphatically—by such a chivalrous and high-minded preacher as Berkeley. This is only one side of the Christian view; the opposite doctrine, that an action done from self-interest is not properly virtuous, has continually asserted itself—either openly conflicting the former thesis or somehow reconciled with it. But although the former thesis is less refined and elevated, it seems to have been the commoner view. Common sense pretty well assumes that ·self·-interested actions that favour the agent's happiness are *prima facie* reasonable, and that the onus of proof lies with those who maintain that disinterested [see Glossary] conduct, as such, is reasonable.

But the common notions of *·self·-interest*, *happiness* etc. are somewhat vague and ambiguous; so that to make these terms usable in scientific discussion we must let them retain the main part of their meanings while trying to make them more precise. We get that result, I think, if by 'greatest possible happiness' we understand the *greatest attainable*

surplus of •*pleasure over* •*pain*, using •those two terms in a comprehensive way, to include respectively all kinds of agreeable and disagreeable feelings. And if we're to accept this quantitative definition of the goal, consistency requires that pleasures be sought in proportion to their pleasantness, so that a less pleasant consciousness mustn't be preferred to a more pleasant one because of some other qualities that it may have. The distinctions of •quality that Mill and others urge can be admitted as grounds of preference if, but only if, they can be resolved into distinctions of •quantity. Practical reasoning that is commonly called 'egoistic', once we have cleansed it of ambiguities and inconsistencies, tends to fit this pattern; and it's only in this more precise form that it's worth our while to examine such reasoning closely. So that's what we must understand an 'egoist' to be—a man who, when two or more courses of action are open to him, does his best to discover what amounts of pleasure and pain are likely to result from each and chooses the one that he thinks will yield him the greatest surplus of pleasure over pain.

2. Adopting the basic **principle** of egoism doesn't necessarily require the ordinary empirical **method** of seeking one's own pleasure or happiness. A man may aim at the greatest happiness within his reach, and yet not try to discover empirically what amount of pleasure and pain is likely to come from any given course of action, because he thinks he has a surer method—a deductive method—for identifying the conduct that will make him happiest in the long-run. He may believe this on grounds of **revealed religion**, because God has promised happiness as a reward for obedience to certain commands; or of **natural religion**, because a just and benevolent God must have organized the world in such a way that happiness will in the long run be distributed in proportion to virtue. It is by combining these two that Paley connects •the universalistic hedonism that he adopts as a method for determining duties with the •egoism that he thinks to be self-evident as a basic principle of rational conduct. Or a man may connect virtue with happiness by a strictly **ethical process of *a priori* reasoning**, as Aristotle seems to do by assuming that the *best* activity will always be unshakably accompanied by the greatest pleasure; with 'best' being fixed by reference to moral intuition, or to the common moral opinions of men generally or of well-bred and well-educated men. Or the deduction by which maximum pleasure is inferred to be the result of a particular kind of action may be **psychological or physiological**; we may have some general theory connecting pleasure with some other physical or mental fact, enabling us to deduce the amount of pleasure that will come from any particular kind of behaviour. Suppose for example that we hold (as many do) that the best chance of pleasure in the long run comes from a perfectly healthy and harmonious exercise of our bodily and mental functions. Given that view we may accept the hedonistic principle without being called on to estimate and compare particular pleasures; rather, we'll have to define the notions of *perfect health* and *harmony of functions* and to consider how those two goals may be achieved. Still those who advocate such deductive methods often appeal to ordinary experience, at least for confirmation, and they admit that only the individual who experiences a pleasure (or pain) directly knows how pleasant (or painful) it is. So it seems that the obvious method for egoistic hedonism is the one I'll call 'empirical-reflective; and I think it's the one that is commonly used in egoistic deliberation. So the next move should be to examine this method as to ascertain clearly the assumptions it involves, and estimate the exactness of its results.

Chapter 2: Empirical hedonism

1. The empirical method of egoistic hedonism, and indeed the very conception of *greatest happiness* as an end of action, rests on the basic assumption that pleasures and pains have definite quantitative relations to each other. If they don't, they can't be conceived as elements—·pleasures positive, pains negative·—of a total that we are to try to make as great as possible. What if some kinds of pleasure (·let's call them 'superpleasures'·) are so much pleasanter than others that the smallest conceivable amount of a superpleasure would outweigh the greatest conceivable amount of any other pleasure? ·That wouldn't wreck the calculation·, because if we knew that it was the case, we could handle any hedonistic calculation involving superpleasures by treating all other pleasures as practically non-existent.[1] But in all ordinary prudential reasoning, I think, we implicitly assume that all the pleasures and pains we could experience bear a finite ratio to each other in respect of pleasantness and its opposite. If we can make this ratio definite, we can balance the intensity of a pleasure (or pain) against its duration; for if finitely long pleasure (or pain) x is intensively greater than another, y, in some definite ratio, it seems to be implied in this conception that if y were continuously increased in extent without change in its intensity it would at a certain point just balance x in amount.[2]

If pleasures can be arranged in a scale, as greater or less in some finite degree, this leads to the assumption of a hedonistic zero—a perfectly neutral feeling—as a point from which the positive quantity of pleasures can be measured. And this emerges even more clearly in the balancing of pleasures against pains. For pain must be reckoned as the negative quantity of pleasure, to be subtracted from the positive in estimating over-all happiness; so we must conceive as at least theoretically possible a point of transition in consciousness at which we pass from the positive to the negative. We don't absolutely have to assume that this strictly neutral feeling ever actually occurs; but experience seems to show that a state close to it occurs quite commonly; and we certainly experience transitions from pleasure to pain and vice versa, and thus (unless all such transitions are abrupt) we must at least momentarily be in this neutral state.

---

[1] Some enthusiastic and passionate people have said that there are feelings so exquisitely delightful that one moment of their rapture is preferable to an eternity of agreeable consciousness of a lower kind. These assertions may have been meant as exaggerations and not intended as statements of fact; but in the case of pain, the thoughtful and subtle Edmund Gurney soberly maintained, as something with important practical implications, that 'torture 'so extreme as to be 'incommensurable with moderate pain' is an actual fact of experience. This doesn't fit my own experience, and I don't think it is supported by the common sense of mankind. . . .

[2] Bentham gives four qualities of any pleasure or pain (taken singly) as important in hedonistic calculation: •intensity, •duration, •certainty, •proximity. If we assume that intensity must be commensurable with duration, the influence of the other qualities on the comparative value of pleasures and pains is easy to determine: we are accustomed to estimate the value of chances numerically, and this method enables us to determine. . . .how much the doubtfulness of a pleasure detracts from its value. And proximity is a property that it's reasonable to disregard except in its effect on uncertainty. My feelings next year should be just as important to me as my feelings next minute, if only I could be equally sure of what they will be. This impartial concern for all the ·temporal· parts of one's conscious life is a prominent element in the common notion of the •rational as opposed to the •impulsive pursuit of pleasure.

This implicitly denies the paradox of Epicurus:

> The state of painlessness is equivalent to the highest possible pleasure. If we can obtain absolute freedom from pain, the goal of hedonism is reached; and then we can vary our pleasure but we can't increase it.

This doctrine contradicts common sense and common experience. But it would be equally wrong to regard this neutral feeling—'hedonistic zero', as I have called it—as the normal condition of our consciousness, out of which we occasionally sink into pain and occasionally rise into pleasure. Nature hasn't been as niggardly to man as this. In my experience, as long as health is retained, and pain and burdensome work banished, the mere performance of everyday functions is a frequent source of moderate pleasures, alternating rapidly with states that are nearly or quite indifferent [see Glossary]. Many Greek moralists in the post-Aristotelian period regarded *apatheia* as the ideal state of existence, but they weren't thinking of it as 'without one pleasure and without one pain', but rather as a state of peaceful intellectual contemplation, which might in philosophic minds reach a high degree of pleasure.

2. I haven't yet made the notions of pleasure and pain precise enough for quantitative comparison. In this and in the rest of the discussion of hedonism I shall mostly part speak only of pleasure, assuming that pain can be regarded as the negative quantity of pleasure, so that any statements about pleasure can through obvious verbal changes be applied also to pain.

[Sidgwick now embarks on a long discussion of proposed definitions of 'pleasure' (and analogous ones of 'pain') by his contemporaries Herbert Spencer and Alexander Bain. Those definitions both imply that to have pleasure is be in a state that one is disposed actively to protect. Sidgwick points to •pleasures of relaxation ('a warm bath'), •the pleasure of food for a temperate person who never eats to the point of satiety, and other objections; conscientiously suggests ways for Spencer and Bain to cope with his counterexamples; and concludes that *for purposes of measurement* (his italics) this approach won't do: the intensity of pleasure/pain is wildly different from any measure of will to continue/cease. With all that out of the way, he returns to his own investigation.]

Shall we then say that the word 'pleasure' names a measurable quality of feeling that is •independent of its relation to volition and •simple in a way that makes it strictly indefinable?—like the quality of feeling named by 'sweet', which we are also conscious of in varying degrees of intensity. Some writers seem to think so; but when I reflect on the notion of *pleasure*—still using that word to cover the most refined and subtle intellectual and emotional gratifications as well as the coarser and more definite sensual enjoyments—the only common quality that I can find is the relation to desire and volition expressed by the general term 'desirable', in the sense I explained [on page 52]. So I shall define *pleasure*—when we are considering its 'strict value' for purposes of quantitative comparison—as a feeling which, when experienced by thinking beings, is at least implicitly taken to be desirable or (in cases of comparison) preferable.

Now a problem arises. When I said in chapter 1, as a fundamental thesis of hedonism, that it's reasonable •to prefer pleasures in proportion to their intensity and •not to allow this to be outweighed by any merely qualitative difference, I implied that it is actually *possible* to prefer pleasures on the non-quantitative grounds that they are 'higher' or 'nobler'; and it is indeed commonly thought that non-hedonistic preferences happen frequently. But my definition of *pleasure* as the kind of feeling that we take to be desirable or preferable seems to make it a contradiction in terms to say that the less pleasant feeling can ever be

thought preferable to the more pleasant.

Here's how to avoid this contradiction. You'll agree that the pleasantness of a feeling is only directly knowable by the individual who feels it, ·and he knows it only· at the time of feeling it. Now, when an estimate of pleasantness involves comparison with feelings that are only represented in idea (I'll return to this shortly), the estimate is liable to be wrong because the representation may be wrong. But no-one is in a position to controvert someone's preference regarding the quality of his present feeling. Now, when we judge the preferableness of a state of consciousness on the grounds of some quality such as 'elevation' or 'refinement' rather than its pleasantness, we seem to be relying on some common standard that others can apply as well as the sentient individual. This leads me to think that when one kind of pleasure is judged to be qualitatively superior to another, although less pleasant, what is being preferred is not really the feeling itself but something in the mental or physical conditions or relations under which it arises, with these being regarded as things that anyone can know. If I in thought distinguish any feeling from all its conditions and concomitants—and also from all its effects on the subsequent feelings of that person or others—and contemplate it merely as the *passing feeling of a single person*, I can't find in it any preferable quality other than the one we call its pleasantness; and the degree of that is knowable directly only by the person who has the feeling.

If my definition of *pleasure* is accepted, and if 'ultimate good' is taken (as I have proposed) to be equivalent to 'what is ultimately desirable', the fundamental proposition of ethical hedonism has a chiefly negative significance; for the statement that

> 'pleasure is the ultimate good '

will only mean that

> nothing is ultimately desirable except desirable feeling, found to be desirable by the sentient individual at the time of feeling it.

It might be objected that this definition couldn't be accepted by a moralist with a stoical cast of mind who refused to recognise pleasure as in any degree ultimately desirable. But even this moralist ought to admit that •an implied judgment that a feeling is *per se* desirable is inseparably connected with •the recognition of that feeling as pleasure; while holding that sound philosophy shows that such judgments are illusory. This indeed seems to have been substantially the view of the Stoic school.

The preference that pure hedonism regards as ultimately rational should be defined as the preference for a feeling valued merely as feeling, according to the estimate of the sentient individual at the time of feeling it; without regard for the conditions and relations under which it arises. So we can state as the basic assumption of what I have [on page 43] called 'quantitative hedonism'—implied by adopting 'greatest surplus of pleasure over pain' as the ultimate end—that all pleasures and pains have for the sentient individual knowable degrees of desirability, positive or negative. The empirical method of hedonism can be applied only if we assume that these degrees of desirability are definitely given in experience. . . .

NOTE. It is sometimes thought that hedonists have to assume that human beings actually *can* achieve a surplus of pleasure over pain—a proposition that an extreme pessimist would deny. But the conclusion that life is always on the whole painful wouldn't make it unreasonable for a man to aim at *minimising* pain. . . ., though it would make immediate painless suicide the only reasonable course for a perfect egoist, unless he looked forward to another life.

Chapter 3: Empirical hedonism (continued)

1. Let *pleasure* be defined, then, as

> a feeling that the sentient individual at the time of feeling it implicitly or explicitly takes to be desirable—i.e. desirable considered merely as a feeling, and not because of •its objective conditions or consequences or of •any facts that can be known and judged by anyone else.

And let it be provisionally assumed that from this point of view feelings generally can be •compared definitely enough for practical purposes and •empirically known to be more or less pleasant in some definite degree. Then the **empirical-reflective method of egoistic hedonism** will be

- •to represent in advance the different series of feelings that our knowledge of physical and mental causes leads us to expect from the different lines of conduct that are open to us;
- •to judge which of the represented series appears to be over-all preferable, taking all probabilities into account; and
- •to adopt the corresponding line of conduct.

This calculation may seem to be too complex for practice: any complete forecast would involve vastly many contingencies with varying probabilities, and we would never get to the end of calculating the hedonistic value of each of these probabilities of feeling. But perhaps we can reduce the calculation to a manageable size without serious loss of accuracy, by •discarding all obviously imprudent conduct, and •neglecting the less probable and less important contingencies. Such discarding is common practice in some of the arts [see Glossary] that have more definite ends—e.g. military strategy and medicine. If the general in ordering a march, or the physician recommending a change of living conditions, took into consideration *all* the relevant circumstances their calculations would become impracticable; so they confine themselves to the most important, and we can do the same in the hedonistic art of life [Sidgwick's phrase].

Some objections against the hedonistic method go much deeper, and by some writers are taken to the point of rejecting the method altogether. Dealing carefully with these objections will be a convenient way of getting a clear view of the method itself and of the results we can reasonably expect from it.

What I'll be discussing are *intrinsic* objections to egoistic hedonism—i.e. arguments against the possibility of obtaining by it the results it aims at. I shan't consider here •whether it is reasonable to take one's own happiness as one's ultimate end; or •how far the moral output of this system will coincide with current opinions about what is right. I postpone these questions for future consideration—in chapter 5, III/14, and the Concluding Chapter of this work; my only concern here is with objections tending to show that hedonism isn't practicable as a rational method.

[The first three objections are by Sidgwick's contemporary T. H. Green. Sidgwick evidently has little respect for these objections, and nor should we. (That is not a judgment on Green's work in general, merely on the bits that Sidgwick has chosen to discuss.) **(1)** Green says that we have no concept of *pleasure as feeling*, only of pleasure as a component in a package that includes the conditions in which it arises. Sidgwick declares this to be •contrary to common sense, •contrary to assumptions made in empirical psychology, and •in conflict with several things that Green says in other

places. **(2)** Green declares hedonism's phrase 'greatest possible pleasure' to be meaningless. In one place he defends this on the ground that 'pleasant feelings are not quantities that can be added', apparently because 'each is over before the other begins'. The same is true of periods of time, Sidgwick points out, but 'it would be obviously absurd to say that hours etc. are not quantities that can be added'. He notes that Green elsewhere silently drops this •objection to adding pleasures in thought, in favour of an •objection to adding them 'in enjoyment or imagination of enjoyment'. Sidgwick: 'No hedonist ever supposed that the happiness he wanted to maximise was something to be enjoyed all at once, or ever wanted to imagine it as so enjoyed'. He rightly sees that as connected with this: **(3)** Green contends that 'an end that is to serve the purpose of a criterion' must 'enable us to distinguish actions that bring men nearer to it from those that do not'. Sidgwick replies:] This presupposes that 'end' has to mean a goal or consummation which, after gradually drawing nearer to it, we reach all at once. But I don't think this is the sense in which ethical writers ordinarily understand the word 'end'; and certainly all that I mean by it is something that it is reasonable to aim at for itself, whether or not attained in successive parts. And as long as my prospective balance of pleasure over pain can be affected by how I act, there seems to be no reason why 'maximum happiness' shouldn't provide a serviceable criterion of choice for conduct.

2. We get a relevant objection to the method of egoistic hedonism if experience confirms that this is true:

> •The consciousness of how transient pleasure is either •makes it less pleasant at the time or •causes a subsequent pain, and
> •the deliberate and systematic pursuit of pleasure tends to intensify this consciousness.

Green doesn't clearly say this, but it seems to be in his mind when he writes that it is 'impossible that self-satisfaction should be found in any series of pleasures' because 'satisfaction for a self that lasts and contemplates itself as doing so must be at least relatively permanent'. [That sentence's two quoted bits from Green's *Prolegomena to Ethics* come from widely separated sections of it.] I suppose the implication is that the disappointment of the hedonist who does't find self-satisfaction where he seeks for it is accompanied by pain or loss of pleasure. If this is so, and if the self-satisfaction thus missed can be obtained by firmly adopting some other principle of action, it seems to follow that the systematic pursuit of pleasure is in danger of defeating itself. So it is important to consider carefully how far this is really the case.

I don't find in my own experience that the mere transientness of pleasures is a serious source of discontent, as long as I have a fair chance of having future pleasures that are as valuable as those in the past, or even as long as my future life has any substantial amount of pleasure to offer. But I don't doubt that for all or most men an important element of happiness comes from the sense of having 'relatively permanent' sources of external pleasure (wealth, social position, family, friends) or internal pleasure (knowledge, culture, strong and lively interest in the well-being of innocently prosperous persons or institutions). But I don't see this as an objection to hedonism. Rather, it seems obvious from the hedonistic point of view that

> 'as soon as intelligence discovers that there are •fixed objects, •permanent sources of pleasure, and •large groups of enduring interests that deliver a variety of recurring enjoyments, the rational will—preferring the greater to the less—will unfailingly devote its energies to the pursuit of these'. [Quoted from *Pessimism* by Sidgwick's younger contemporary James Sully.]

It may be replied that if these permanent sources of pleasure are consciously sought merely as a means to the hedonistic end, they won't deliver the happiness they were sought for. To some extent I agree; but if the normal complexity of our impulses is properly taken into account, this statement won't count against the adoption of hedonism but will merely warn the hedonist of a danger that he has to guard against. In a previous chapter [page 24] I followed Butler in stressing the difference between •impulses that are strictly directed towards pleasure and •'extra-regarding' [see Glossary] impulses that don't aim at pleasure—though much and perhaps most of our pleasure consists in the gratification of the latter. . . . I argued there that in many cases the two kinds of impulse are so far incompatible that they do not easily coexist in the same moment of consciousness. I added, however, that in everyday life the incompatibility is only momentary, and doesn't prevent a real harmony from being attained by a sort of 'alternating rhythm of the two kinds of impulse in consciousness'. But this harmony is certainly liable to be disturbed; and while on the one hand

> individuals can and do sacrifice their greatest apparent happiness to the gratification of some imperious particular desire,

on the other hand,

> self-love is liable to absorb the mind enough to block a healthy and vigorous outflow of the 'disinterested' impulses towards particular objects—impulses that we must have if we are to achieve any high degree of the happiness that self-love aims at.

I don't infer from this that the pursuit of pleasure must be self-defeating and futile; but merely that when the principle of egoistic hedonism is applied with a due knowledge of the laws of human nature, it is practically self-limiting—i.e. that a rational method of reaching the desired end requires us to some extent to put it out of sight and not directly aim at it. I have before [page 22] called this the 'fundamental paradox of egoistic hedonism'; but though it looks like a paradox, it doesn't seem to present any practical difficulty once the danger has been clearly seen. We are *very* familiar with cases in which

> a man •lets the original goal of his efforts (whatever they may be) slip out of sight, and •comes to regard his means to this end as ends in themselves, so that eventually he even sacrifices the original end in order to attain what is only derivatively desirable.

If it's that easy and common to •overdo forgetting the end in favour of the means, there's no reason why it should be hard to do this •to the extent that rational egoism prescribes; and in fact this seems to be continually done by ordinary folk with amusements and pastimes of all kinds. . . .

But it is sometimes thought that there's an important class of refined and elevated impulses that are in a special way incompatible with the supremacy of self-love—such as the love of virtue, or personal affection, or the religious impulse to love and obey God. But the common view of these impulses doesn't seem to recognise this difficulty. Of all the moralists who followed Shaftesbury in contending that it is a man's true interest to acquire strictly disinterested social affections *none* have seen these affections as inherently incompatible with the supremacy of rational self-love. And Christian preachers who have commended the religious life as really the happiest haven't thought genuine religion to be irreconcilable with the conviction that each man's own happiness is his primary concern.

But there are others who seem to carry •religious consciousness and •the feeling of human affection to a higher stage of refinement at which a stricter disinterestedness is required. They maintain that the essence of each of •those

feelings in its best form is absolute self-renunciation and self-sacrifice. And these do seem incompatible with self-love, even in a cautiously self-limiting form. A man can't both wish to secure his own happiness and be willing to lose it. But what if willingness to lose it is the true means of securing it? Can self-love not merely reduce indirectly its prominence in consciousness, but directly and unreservedly annihilate itself?

This emotional feat doesn't seem to me possible; so I have to admit that a man who accepts the principle of rational egoism cuts himself off from the special pleasure that comes with this absolute sacrifice and abnegation of self. But however exquisite this pleasure may be, the pitch of emotional exaltation and refinement needed to attain it is comparatively so rare that it has no place in men's common estimate of happiness. So I don't think it's an important objection to rational egoism that it is incompatible with this particular state of consciousness. Nor do I think that the common experience of mankind really confirms that the desire for one's own happiness, if accepted as supreme and regulative, *inevitably* defeats its own aim by lessening and thinning the impulses and emotional capacities that are needed for great happiness; though it certainly shows a serious and subtle danger in this direction.

3. The habit of mind resulting from the continual practice of hedonistic comparison is sometimes thought to be unfavourable to achieving the hedonistic goal because •the habit of reflectively observing and examining pleasure is thought to be incompatible with •the capacity for experiencing pleasure in normal fullness and intensity. And it certainly seems important to consider what effect the continual attention to our pleasures. . . .is likely to have on these feelings themselves. This inquiry seems at first sight to reveal an incurable contradiction in our view of pleasure. ·**On one side**:· Pleasure exists only as it is *felt*, so the more conscious we are of it, the more pleasure we have; and it would seem that the more our attention is directed towards it, the more fully we'll be conscious of it. **On the other hand**, Hamilton's statement that 'knowledge and feeling' (cognition and pleasure or pain) are always 'in a certain inverse proportion to each other' seems at first to square with our common experience, because the purely cognitive element of consciousness seems to be neither pleasurable nor painful, so that the more our consciousness is occupied with cognition the less room there seems to be for feeling.

But this assumes that the total intensity of our consciousness is a constant quantity; so that when one element of it increases, the rest must diminish; and I can't see any empirical evidence for that. Rather, it seems that at certain times in our life intellect and feeling are simultaneously feeble, so that a single mental excitement could intensify both at once.

Still, it does seem that any very powerful feeling—as intense as we are normally capable of—is commonly diminished by a stroke of cognitive effort; so the exact observation of our emotions does face a general difficulty, namely that the observed item seems to shrink and dwindle in proportion as the study of it grows keen and eager. How, then, are we to reconcile this with the proposition that pleasure exists only as we are conscious of it? The answer seems to be this:

> Having a feeling *essentially* involves being conscious of it; so it can't be the case that the mere consciousness of a present feeling diminishes the feeling! But in introspection we go beyond the present feeling, comparing and classifying it with other feelings (remembered or imagined), and the effort of representing and comparing these other feelings tends to decrease the consciousness of the actual pleasure.

My conclusion is this: although there's a real danger of diminishing pleasure by the attempt to observe and estimate it, the danger seems to arise only for very intense pleasures and only if the attempt is made at the time when the pleasure is happening; and since the most delightful periods of life have frequently recurring intervals of nearly neutral feeling in which the pleasures immediately past may be compared and estimated without any such detriment, I don't attach much importance to the objection based on this danger.

4. More serious are the objections urged against the possibility of reliably comparing pleasures and pains in the way that the hedonistic standard requires. Of course we *do* habitually compare pleasures and pains in respect of their intensity—for example, we unhesitatingly declare our present state to be more or less pleasant than the one we have just left, or than a further-back one that we remember; and we declare some pleasant experiences to have been *worth the trouble* it took to obtain them or *worth the pain* that followed them. But despite this it may still be maintained **(1)** that this comparison is ordinarily haphazard and very rough, and that it can't be extended to meet systematic hedonism's demands or applied with any accuracy to all possible states, however differing in quality; and **(2)** that this kind of comparison as commonly practised is liable to illusion; we can't say exactly *how much* illusion, but we are continually forced to recognise that there is *some*. Plato adduced this as a ground for distrusting the apparent affirmation of consciousness in respect of present pleasure. . . .

I agree with critics who say that in estimating present pleasure there's no conceivable appeal from the immediate decision of consciousness—that here the phenomenal is the real. But error can come in, as follows. In any estimate of the intensity of our present pleasure we must be comparing it with some other state. And this other state must generally be a representation, not an actual feeling; for though we can sometimes experience two pleasures at once, we can't often compare them satisfactorily (either because •their causes interfere with one another, so that neither pleasure reaches its normal degree of intensity, or because •the two blend into a single state of pleasant consciousness whose ingredients can't be estimated separately). Now, if one of the compared items must be an imagined pleasure, that opens the door to a possibility of error, because the imagined feeling may not adequately represent the pleasantness of the corresponding actual feeling. And in the comparisons required by egoistic hedonism *all* the compared pleasures are represented rather than actual, for we are trying to choose between two or more *possible* courses of conduct.

Let us then look more closely at how this comparison is ordinarily made, so as to see what positive grounds we have for mistrusting it.

In estimating the values of different pleasures open to us, we trust mostly to our prospective imagination: we project ourselves into the future, and imagine what such and such a pleasure will amount to under hypothetical conditions. The conscious inference involved in this imagination is mainly based on our experience of past pleasures, which we usually recall •generically though sometimes we bring in definitely remembered •particular pleasures; and we are also influenced by the experiences of pleasure of others—sometimes •particular experiences we have been told about, and sometimes traditional •generalisations about the common experience of mankind.

A process such as this isn't likely to be free from error, and no-one claims that it is. In fact there's hardly anything that moralisers have emphasised more than the fact that forecasts of pleasure are continually erroneous. Each of us frequently recognises his own mistakes, and attributes to

others errors that they haven't seen for themselves—errors due to misinterpretation of their own experience or ignorance or neglect of that of others.

How are these errors to be eliminated? The obvious answer is that we must replace •the instinctive, largely implicit inference that I have just described by •a more scientific process of reasoning in which we deduce the probable degree of our future pleasure or pain in any given circumstances from generalisations based on a large enough number of careful observations of our own and others' experience. This raises three questions: **first:** How accurately can each of us estimate his own past experience of pleasures and pains? **second:** How far can this knowledge of the past enable him to make secure forecasts regarding the greatest happiness within his reach in the future? **third:** How much can he know about the past experience of others?

In tackling the **first question**, remember that it's not enough just to know generally that we derive pleasures from these sources and pains from those; we need also to know approximately how much pleasure or pain each source provides. If we can't form some quantitative estimate of them, it is futile to try to achieve our *greatest possible* happiness—at least by an empirical method. Our task with each pleasure as it occurs or is recalled in imagination, to compare it quantitatively with other imagined pleasures. And the question is: how trustworthy can such comparisons be?

When I reflect on my pleasures and pains, and try to compare their intensities, I can't get far in obtaining clear and definite results,. . . .whether the comparison is made •between two states of consciousness recalled in imagination or •between one such state and a pleasure I was having at the time of the comparison. This is true even when I compare feelings of the same kind; and as I move to feelings of different kinds, the vagueness and uncertainty of my results increases proportionately. Let us begin with sensual gratifications, which are thought to be especially definite and graspable. If when enjoying a good dinner I ask myself whether the fish (or the Chablis) gives me more pleasure than the beef (or the claret), sometimes I can decide but very often I can't. ·Another kind of example·: If I have undertaken two kinds of bodily exercise of which one and not the other was markedly pleasant (or tedious), I naturally take note of that difference between them; but I don't naturally go further than this in judging how pleasurable (or painful) each was, and when I try to do so I don't get any clear result. And similarly with intellectual exercises and predominantly emotional states of consciousness: even when the causes and quality of the compared feelings are similar, the hedonistic comparison doesn't yield any definite result except when the differences in pleasantness are enormous. When I try to get a *scale* for pleasures of different kinds, e.g. comparing

- •labour with rest,
- •excitement with tranquility,
- •intellectual exercise with emotional outflow,
- •the pleasure of scientific knowledge with the pleasure of beneficent action,
- •the delight of social expansion with the delight of aesthetic reception,

my judgment wavers and fluctuates far more, and in most cases I can't give any confident decision. And if this is the case with pain-free pleasures (Bentham calls them 'pure'), it's even more true of those commoner states of consciousness in which a predominant pleasure is mixed with a certain amount of pain or discomfort. If it's hard to say which of two states of contentment was the greater pleasure, its harder to compare •a state of placid satisfaction with •one of eager but hopeful suspense, or •triumphant conquest of painful obstacles. And it may be even *more* difficult to compare

pure pleasures with pure pains—to say how much of one we consider to be exactly balanced by a given amount of the other, when they don't occur simultaneously. . . .

5. Further trouble: these judgments aren't clear and definite, and still less are they consistent. . . . Each person's judgment of the comparative value of his own pleasures is apt to change through time, though it relates to the same past experiences; and this variation casts doubt on the validity of any particular comparison.

This variation seems to be caused partly by **(a)** the nature of the represented feeling and partly by **(b)** the general state of the mind at the time when the comparison is made. **(a)** We find that different kinds of past pleasures and pains are not equally revivable in imagination.

> Pains that hook into emotions and ones that somehow *mean* something are more easily revived than pains that are just unpleasant experiences. At this moment I can more easily get back in imagination the discomfort of the sense of 'I'm going to throw up' than the pain of the actual vomiting, although my memory of what I thought at the time tells me that the pain of expectancy was trifling compared with the nastiness of the vomiting.

·And the nature and context of a pleasure can also affect how, and how easily, it comes to mind later·. That seems to be why past hardships, toils, and anxieties often seem pleasurable when we look back on them later: the excitement, the heightened sense of life that accompanied the painful struggle, would have been pleasurable in itself, and it's *this* that we recall rather than the pain. **(b)** In estimating pleasures, the other cause of variation is more obvious: we're aware that our estimate of them varies with changes in our mental or bodily condition. Everyone knows that we can't adequately estimate the gratifications of appetite when we are in a state of satiety, and that we're apt to exaggerate them when we are very hungry. (I don't deny that a pleasure may be increased by the intensity of the antecedent desire for it; so that in these cases the pleasure doesn't merely *appear*, as Plato thought, but actually *is* greater because of to the strength of the preceding desire. But ·this isn't always the case·: we all know that intensely desired pleasures often turn out to be disappointing.)

There seems to be no analogue for this on the *pain* side of the ledger. . . . But the prospect of certain ·kinds of· pains throws most people into the state of passionate aversion that we call 'fear', leading them to estimate such pains as worse than they would be judged to be in a calmer mood.

Further, in the presence of any kind of pain or uneasiness we seem liable to underrate pains of very different kinds: in •danger we value •repose, overlooking its *ennui* [= 'boringness'], and the tedium of •security makes us imagine the mingled excitement of past •danger as almost purely pleasurable. And when we are absorbed in some pleasant activity, the pleasures of dissimilar activities are apt to be looked down on; they seem coarse or thin, as the case may be; and this is a basic objection to noting the exact degree of a pleasure at the time of experiencing it. [To ensure that Sidgwick's elegantly compressed 'coarse or thin, as the case may be' is understood: while I'm thrilling to a raft-trip down the Colorado, I might think of the pleasure of listening to Schubert as 'thin'; while listening to Schubert, I might think of the pleasure of the river-trip as 'coarse'.] The eager desire that often seems essential to a whole state of pleasurable activity usually involves a similar bias; indeed *any* strong excitement—whether from aversion, fear, hope, or suspense—in which our thought is concentrated on a single result. . . .tends to make us under-rate different pleasures and pains. More generally, at a time when we are incapable of experiencing a certain pleasure we can't imagine it as very

intense—the pleasures of intellectual or bodily exercise at the close of a wearying day, an emotional pleasure when our capacity for it is temporarily exhausted. Many philosophers have thought we could guard against error in this matter by making our estimate in a cool and passionless state, but that is wrong. Many pleasures can be experienced in their full intensity only if they are preceded by desire, and even by enthusiasm and high-pitched excitement; and we're not likely to evaluate these adequately when we're in a state of perfect tranquillity.

6. These considerations make clearer the extent of the assumptions of empirical quantitative hedonism, stated in the preceding chapter [page 60]: **(i)** that each of our pleasures and pains has a definite degree, and **(ii)** that this degree is empirically knowable. **(i)** If pleasure exists only in being felt, the belief that every pleasure and pain has a definite intensive quantity or degree must remain an *a priori* assumption that couldn't be given positive empirical verification. A pleasure can have such-and-such a degree only as compared with other feelings of the same or some different kind; but usually this comparison can be made only in imagination, and the best result it can come up with will be something of this hypothetical form:

> If feelings F_1 and F_2 were felt together precisely as they have been felt separately, one would be found more desirable than the other in ratio R.

If we're asked what grounds we have for *believing* this hypothetical, all we can say is that it is irresistibly suggested by reflection on experience, and at any rate uncontradicted by experience.

(ii) Granting that each of our pleasures and pains really does have a definite degree of intensity, do we have any means of accurately measuring it? Is there any evidence that the mind is ever in a state that makes it a perfectly neutral and colourless medium for imagining all kinds of pleasures? Experience certainly shows that we are often in moods in which we seem to be biased for or against a particular kind of feeling. Isn't it probable that there's *always* some bias of this kind? that we are always more in tune for some pleasures, more sensitive to some pains, than we are to others? There's no getting away from it: exact knowledge of the place of each kind of feeling in a scale of desirability—with its mid-point being a zero of perfect indifference—is at best an ideal, and we can never tell how close we come to it. Still, ·things could be worse·. The variations in our judgments and the disappointment of our expectations give us experience of errors whose causes we can trace and allow for, at least roughly, correcting in thought the defects of imagination. And what we need for practical guidance is only to estimate. . . .the value of a *kind* of pleasure or pain as obtained under certain circumstances or conditions; we can diminish somewhat the chance of error in this estimate by making several observations and imaginative comparisons, at different times and in different moods. To the extent that these agree, we're entitled to more confidence in the result; and to the extent that they differ, we can at least reduce our possible error by taking the average of the estimates. Obviously, though, a method like this can't be expected to do more than roughly approximate to the truth.

7. So we must conclude that our estimate of the hedonistic value of any past pleasure or pain is liable to error and we can't calculate how much, because the represented pleasantness of different feelings fluctuates, varying indefinitely with changes in the condition of the representing mind. And even if we *could* adequately allow for it, this source of error in our comparison of past pleasures is liable to intrude again when we argue from the past to the future. ·This brings

us to the **second question** of the three posed on page 65·. Here are some of the things that could interfere with the past-to-future inference:

- •Our capacity for particular pleasures has changed since the experiences that our calculation is based on.
- •We have reached satiety for some of our past pleasures, or become less susceptible to them because of changes in our constitution.
- •We have become more susceptible to pains connected with these pleasures.
- •Altered conditions of life have given us new desires and aversions, and brought new sources of happiness into prominence.

Or any or all of these changes are likely to occur before the completion of the course of behaviour that we are now deciding on. The hedonistic calculations of youth have to be adjusted as we become older; a careful estimate of a •girl's pleasures. . . .wouldn't be much use to a •young woman.

No-one when trying to estimate the probable effect on his happiness of •new circumstances and influences, •untried rules of conduct and •fashions of life, relies entirely on his own experience; such a person always argues partly from the experience of others. But by including inferences from other men's experience we inevitably introduce a new possibility of error; ·now we confront the difficulty raised in the **third question** of the three posed on page 65·. For any such inference ·from others to oneself· assumes a similarity of nature among human beings, and this is never exactly true. We can't tell exactly know how far short of the truth it falls; but we have enough evidence of the strikingly different feelings produced in different men by similar causes to convince us that the ·similarity-of-nature· assumption would often be wholly misleading.

(That is why Plato's argument that the philosopher's life has more pleasure than the life of the devotee of sensual pleasure is a total failure. He argues:

> The philosopher has tried both kinds of pleasure, sensual and intellectual, and prefers the delights of philosophic life; so the sensualist ought to trust his decision and follow his example.

But—who knows?—the philosopher's constitution may be such as to make his enjoyments of the senses comparatively feeble, while the sensualist's mind may be unable to achieve more than a thin shadow of the philosopher's delight.)

If we are to be guided by someone else's experience, therefore, we need to be convinced •that he is generally accurate in observing, analysing, and comparing his sensations, and •also that his relative susceptibility to the different kinds of pleasure and pain in question coincides with our own. . . . And however accurate he is about the causes of his feelings, there's the question of whether similar causes would produce similar effects in us; and this uncertainty is greater if our adviser has to rely on long-term memory to know about some of the pleasures or pains that are being compared. Thus in the perpetual controversy between Age and Youth, wisdom isn't as clearly on the side of maturer counsels as it seems to be at first sight. When a youth is warned by his senior to abstain from some pleasure because it's not worth the possible pleasures that must be sacrificed for it and the future pains that it will entail, he can't easily know how far the older man—even if he could once •feel the full rapture of the delight that he is asking the younger to renounce—can now •recall it.

And this source of error gets at us in a more extended and more subtle manner than has yet been noticed. Our sympathetic [see Glossary] sense of others' experiences of pleasure and pain has been continually and variously

exercised throughout our life, by actual observation and oral communication with other human beings, and through books, paintings, sculptures and so on, so that we can't tell how far it has unconsciously blended with •our own experience and affected how •it is represented in memory. . . .

Those considerations should seriously reduce our confidence in the 'empirical-reflective method of egoistic hedonism', as I have called it. I don't conclude that we should reject it altogether; I'm aware that despite all the difficulties that I have presented I continue to act on the basis of comparisons that I make between pleasures and pains. But I conclude that if one wants help with the systematic direction of conduct, it would be highly desirable to control and supplement the results of such comparisons by the assistance of some other method—if we can find one that we see any reason to rely on.

Chapter 4: Objective hedonism and common sense

1. Before I examine (in chapter 6) methods of seeking one's own happiness that are further remote from the empirical methods, I want to consider (in this chapter) how far we can escape the difficulties and uncertainties of the method of reflective comparison by relying on *current opinions and accepted estimates* of the value of different objects commonly sought as sources of pleasure.

At least in large-scale planning of their lives men seem to find it natural to seek and estimate •the objective conditions and sources of happiness, rather than •happiness itself; and it's plausible to suggest that by relying on estimates of the former we avoid the difficulties of the introspective method of comparing feelings. What makes this plausible is the thought that common opinions about the value of different sources of pleasure are the net result of the combined experience of mankind down through the generations, in which all the individual differences I have been writing about balance and neutralise one another and thus disappear.

I don't want to undervalue the guidance of common sense in our pursuit of happiness. But when we consider these common opinions as premises for the deductions of systematic egoism we find them to be open to ·at least seven· serious objections.

(i) At best common sense gives us only an estimate that is true for an average or typical human being—and we have already seen that any particular individual will probably diverge somewhat from this type. So each person will have to correct common opinion's estimate by the results of his own experience, in order to get from it trustworthy guidance for his own conduct; and it looks as though this process of correction must be infected by all the difficulties we are trying to escape. **(ii)** The experience of the mass of mankind has too narrow a range for its results to help much in the present inquiry. Most people spend most of their time working to avert starvation and severe bodily discomfort; and their brief periods of leisure, after supplying the bodily needs of food, sleep, etc. is spent in ways determined by impulse, routine, and habit rather than by a deliberate estimate of probable pleasure. So it seems that the 'common sense' we are to appeal to ·in our hedonistic inquiry· can only be

that of a minority of comparatively rich and leisured persons. **(iii)** For all we know to the contrary, the mass of mankind—or some section of it—may be *generally and normally* under the influence of some of the causes of mis-observation that I have discussed. We avoid the 'idols of the cave' by trusting common sense, but what is to guard us against the 'idols of the tribe'? [In his *New Organon* Bacon labels as 'idols of the cave' the sources of error in the individual person, and as 'the idols of the tribe' the sources of error in the human species as a whole.] **(iv)** The common estimate of different sources of happiness seems to involve all the confusion of ideas and points of view that we have worked hard to eliminate in describing the empirical method of hedonism. •It doesn't distinguish objects of natural desire from sources of experienced pleasure. We saw in I/4 that these two don't exactly coincide; indeed we often see men continuing not only to feel but to indulge desires, though they know from experience that they'll bring more pain than pleasure. So the current estimate of the desirability of various goals doesn't express simply men's experience of pleasure and pain; for men are apt to think desirable what they strongly desire, whether or not they have found it conducive to happiness on the whole; and so the common opinion will tend to represent a compromise between the average force of desires and the average experience of the consequences of gratifying them.

(v) We must allow for the intermingling of moral with purely hedonistic preferences in the estimate of common sense. For it often happens that a man chooses a course of conduct because he expects greater happiness from it, but this expectation comes from his thinking that the chosen conduct is the right or more excellent or more noble thing to do. He is here perhaps unconsciously assuming that the morally best action will turn out to be also the one that does most for the agent's happiness. (I'll explore this assumption in chapter 5.) And a similar assumption seems to be made, on no good evidence, regarding merely aesthetic preferences.

(vi) Are we to be guided by the preferences that men say they have, or by those that their actions would lead us to infer? On one hand, we can't doubt that men often, from weakness of character, fail to seek what they sincerely believe will give them most pleasure in the long run; on the other hand, because a genuine preference for virtuous or refined pleasure is a mark of genuine virtue or refined taste, men who don't actually *have* such a preference are (perhaps unconsciously) influenced by a desire to be credited with having it, which affects what they *say* about their estimates of pleasures.

2. **(vii)** Even if we had no doubt on general grounds that common sense would be our best guide in the pursuit of happiness, we would still be in difficulties because its utterances on this topic are so unclear and inconsistent. Quite apart from differences of time and place, serious conflicts and ambiguities are found if we consider only the current common sense of our own age and country. Let us list the sources of happiness that seem to be recommended by an overwhelming consensus of current opinion:

- •health,
- •wealth,
- •friendship and family affections,
- •fame and social position,
- •power,
- •interesting and congenial occupation and amusement, including
- •the gratification of the love of knowledge, and of the refined susceptibilities—partly sensual, partly emotional—that we call 'aesthetic'.

What are the relative values of these objects of common pursuit? We seem to get no clear answer from common

sense. A possible exception to that: it would be generally agreed that health ought to be outrank everything else; but even on this point we couldn't infer general agreement from observation of the actual conduct of mankind! Indeed, even as regards the positive [see Glossary] estimate of these sources of happiness, we find on closer examination that the supposed consensus is much less clear than it seemed at first. Not only are there many important groups of dissidents from the current opinions, but their paradoxical views are in a strange and unexpected way welcomed and approved by the very same majority—the same common sense of mankind—that maintains the opinions *from* which they are dissenting. Men show a really startling readiness to admit that •the estimates of happiness that guide them in their ordinary habits and pursuits are wrong and that •sometimes the veil is lifted, so to speak, and the error is displayed. [Sidgwick is being sarcastic here; read on!]

For example, men seem to put great value on the ample gratification of bodily appetites and needs; wealthy people spend a lot of money and forethought on the means of satisfying these appetites in a luxurious manner; and though they do not often deliberately sacrifice health to this gratification—common sense condemns that as irrational—still one may say that they are habitually *courageous* in pushing right up to the edge of this imprudence.

Yet the same people are fond of saying that 'hunger is the best sauce', and that 'temperance and labour will make plain food more delightful than the most exquisite products of the culinary art'. And they often argue with perfect sincerity that as regards these pleasures the rich really have little or no advantage over the comparatively poor, because habit soon makes the rich man's luxurious provision for the satisfaction of his •acquired needs no more pleasant to him than the satisfying of more •primitive appetites is to the poor man. [There is nothing condescending about 'primitive' here; it is contrasted with 'acquired', and means about the same as 'natural'.] And the same line of thought is often extended to all the material comforts wealth can purchase. It is often contended that habit makes us indifferent to these comforts while we have them, and yet we suffer when we have to do without them. . . . And it's only a short step to the conclusion that wealth—

- •in the pursuit of which most men agree in concentrating their efforts,
- •on the attainment of which all congratulate each other,
- •for which so many risk their health, shorten their lives, reduce their enjoyments of domestic life, and sacrifice the more refined pleasures of science and art

—is really a very doubtful gain for most people, for whom the cares and anxieties it involves cancel out the slight advantage of the luxuries that it purchases.[1]

In England social rank and status is an object of passionate pursuit, yet there's an often-expressed and generally accepted view that •it has no intrinsic value as a means of happiness; that •though the process of social ascent is perhaps generally agreeable, and descent is certainly painful, yet life *up there* is no more pleasant than life at the humbler level; that •happiness can be found as easily (if not more easily) in a cottage as in a palace; and so forth.

Even more routine are the commonplaces about the

[1] It is striking to find the author of *The Wealth of Nations*, the founder of a long line of economists who are commonly believed to exalt the material means of happiness above all the rest, declaring that 'wealth and greatness are only trivially useful, mere trinkets', and that 'in ease of body and peace of mind all the different ranks of life are nearly on a level; the beggar sitting in the sun beside the highway *has* the security that kings *fight* for'. Adam Smith, *Theory of the Moral Sentiments* IV/1.

emptiness and vanity of the satisfaction brought by fame and reputation. The case of posthumous fame, indeed, is a striking instance of my general thesis that the commonly accepted ends of action are determined partly by the average force of desires that are *not* directed towards pleasure or shaped by experiences of pleasure. Posthumous fame seems to rank pretty high among the objects that common opinion regards as good or desirable for the individual; and the pursuit of it isn't ordinarily criticised as imprudent, even if it leads a man to sacrifice other important sources of happiness to a result that he doesn't expect to have any consciousness of. Yet the slightest reflection shows such a pursuit to be *prima facie* irrational from an egoistic point of view;[1] and every moraliser has found this an obvious and popular topic. The actual consciousness of present fame is no doubt very delightful to most persons; but the moraliser has no trouble maintaining that even this is accompanied by disadvantages that make its hedonistic value very doubtful.

The current estimate of the desirability of *power* is pretty high, and it may be that

> the more closely and analytically we examine men's actual motives, the more widespread and predominant we'll find the pursuit of power to be;

because many men seem to seek wealth, knowledge, even reputation, as a means to getting power rather than for their own sakes or as a means to other pleasures. And yet men willingly agree when they're told that •the pursuit of power, as of fame, is prompted by an empty ambition that is never satisfied but only made more uneasy by such success as it manages to achieve; that •the anxieties that accompany not only the *pursuit* but also the *possession* of power, and the jealousies and dangers inseparable from power, far outweigh its pleasures. . . .

Moralisers broadly agree that •the exercise of the domestic affections is an important means to happiness; and •this certainly seems to be prominent most people's plan of life. But it may fairly be doubted whether men in general *do* value domestic life very highly, apart from the gratification of sexual passion. Certainly at any time and place where men could freely indulge their sexual urges while avoiding the burden of a family, without any serious fear of social disapproval, the unmarried state has tended to become common; it has even become so common as to arouse the grave anxiety of legislators. And though common sense has always disapproved of such conduct, that seems to have been because it is seen as anti-social rather than because it is seen as imprudent.

Thus we find great instability and uncertainty in the most decisive judgments of common sense concerning the things that common opinion seems most clearly and confidently to recommend as sources of pleasure—bodily comfort and luxury, wealth, fame, power, society. It's true that the pleasures derived from art and the contemplation of the beauties of nature, and the pleasures of scientific curiosity and the exercise of the intellect generally, are highly praised; but it's hard to formulate a 'common opinion' regarding them because the high estimates often given to them seem to express the real experience of only small minorities. These have persuaded leisured people to let culture be regarded as an important source of happiness; they haven't produced any generally accepted opinion as to its importance in comparison with the other sources I have mentioned. . . .

[1] It might be justified on a self-love basis by dwelling on the •pleasures of hope and anticipation that accompany the pursuit. But this is obviously an after-thought. It is not for the sake of •these that posthumous fame is sought by him whom it spurs 'To scorn delights and live laborious days' [quoted from Milton].

(viii) [This carries on from **(vii)** on page 70.] Even if the consensus regarding sources of happiness were far more complete and clear than they seem in fact to be, its value would be greatly lessened by the dissent of important minorities that I haven't so far talked about. For example, many religious folk regard *all* worldly pleasures as mean and trifling; so full of vanity and emptiness [Sidgwick's phrase] that the eager pursuit of them is possible only through ever-renewed illusion, leading to ever-repeated disappointment. And this view is shared by a good many reflective persons who have no religious bias, as you can see from the numerous adherents that pessimism has won in recent times. Indeed a somewhat similar opinion about the value of the ordinary objects of human pursuit has been expressed by many philosophers who weren't pessimists; and considering the fact that it's the philosopher's business to reflect with care and precision on the facts of consciousness we shouldn't rush to let them be *outvoted* by the mass of mankind. On the other hand, the philosopher's capacities of feeling aren't typical of those of humanity in general; so if he erects the results of his individual experience into a universal standard, he is likely to overrate some pleasures and underrate others. Convincing illustrations of this are provided by thinkers such as Epicurus and Hobbes—not of the idealist or transcendental type, but overt hedonists. We can't accept as fair expressions of normal human experience either •Epicurus's identification of painlessness with the highest degree of pleasure, or •Hobbes's assertion that the gratifications of ·scientific· curiosity 'far exceed in intensity all carnal delights'. So here is our problem: the mass of mankind, to whose common opinion we naturally look for universally authoritative beliefs about the conditions of happiness, are not good at or practised in observing and recording their experience; and usually the better a man is at observing ·his conscious processes·, the wider is the gap between the phenomena that he can observe and the ordinary type.

3. We have to accept that the hedonistic method can't be made exact and certain by appealing to what common sense says about the sources of happiness. But I don't want to exaggerate the difficulty of organising common sense into a fairly coherent body of *probable* doctrine that can provide some practical guidance. ·I have two main points to make about this·. **(a)** Commonly commended sources of happiness *can* compete with one another and present themselves as alternatives, but this doesn't happen often, and when it does the competition isn't severe. The pursuit of wealth often leads also to power (in addition to the power that is inherent in wealth) and to reputation; and these objects of desire can usually be best obtained—if we can obtain them at all—by activities that in themselves provide the pleasure that normally comes with the energetic use of one's best faculties; and these congenial activities are not incompatible with •adequate exercise of social and domestic affections. or with •cultivated entertainment (which must be carefully limited if it is to be really entertaining). . . .

(b) As for the philosophical or quasi-philosophical paradoxes regarding the illusoriness of sensual enjoyments, wealth, power, fame, etc., we can explain the widespread acceptance of these by admitting a certain general tendency to exaggeration in the common estimates of such objects of desire, which from time to time causes a reaction and an equally excessive temporary depreciation of them. As I pointed out in chapter 3, it is natural for men to value too highly the pleasures they hope and long for; power and fame, for example, bring anxieties and disgusts that aren't foreseen when they are represented in longing imagination; yet it may still be true that power and fame give most men ·who have them· a clear balance of happiness on the whole. And it

seems clear that luxury adds less to the ordinary enjoyment of life than most impoverished men suppose. . . ., so we can fairly conclude that increase of happiness is very far from keeping pace with increase of wealth. But when we take into account the pleasures and the security that wealth can bring, we can hardly doubt that increase of wealth normally brings some increase of happiness—at least until a man reaches an income beyond that of the great majority in any actual community. So we can reasonably conclude that although it is extravagant to say that happiness is 'equally distributed through all ranks and callings', it *is* distributed more equally than men's external circumstances might suggest, especially given the importance of the pleasures that accompany the exercise of the affections. Also, common sense recognises that •some people with unusual temperaments find the ordinary pleasures of life to be quite trifling compared with more refined enjoyments; and also that •men generally are liable occasionally to fall under the sway of absorbing impulses that take them out of the range within which the judgments of common sense are even broadly and generally valid. No-one expects a lover to care much for anything except the enjoyments of love!. . . .

Common sense, in fact, hardly claims to provide more than rather indefinite general rules that no prudent man should neglect without giving himself a reason for doing so. Such reasons may come from his knowledge of some special features of his own nature, or from the experience of others whom he believes to be more like himself than the average of mankind are. Still, we have seen that there's a considerable risk of error in relying on the special experience of others; and, to cut the story short, it seems that no process of this kind—appealing to the opinion of the many, or of cultivated persons, or of those whom we judge most to resemble ourselves—can solve with precision or certainty the problems of egoistic conduct.

So we still have our question:

> Can we have a general theory about the causes of pleasure and pain that is sufficiently certain and usable to enable us to rise above •the ambiguities and inconsistencies of common or sectarian opinion, and •the shortcomings of the empirical-reflective method, and establish the hedonistic art of life on a thoroughly scientific basis?

I shall consider this question in chapter 6; but first I shall examine a common belief about the way to happiness which, though it doesn't rest on a scientific basis, is thought by its adherents to be more certain than most of the current opinions that we have been examining. This is the belief that a man's *doing his duty* [see Glossary] will bring him the greatest happiness he can have. This means his duty as commonly recognised and prescribed, unless he deviates from this standard in obedience to a truer conception of how to achieve or promote universal good.[1] Because of how important this opinion is to a writer on morals, I give it a chapter of its own.

[1] In chapter 6 I shan't discuss the case where the person's conscience definitely conflicts with the general moral consciousness of his age and country. It is commonly held to be a man's duty always to obey his own conscience, even at the risk of error, but it isn't commonly held that this will always bring him the greatest happiness open to him.

Chapter 5: Happiness and duty

1. The belief that happiness is connected with duty tends to be widely accepted by civilised men, at least after a certain stage in civilisation has been reached. But it doesn't seem likely that we would affirm it as a generalisation from experience, rather as something known from direct divine revelation or by inference from •the belief that the world is governed by a perfectly good and omnipotent being. To examine •that belief thoroughly is one of the most important tasks that human reason can attempt; but because it involves delving into the evidence for natural and revealed religion, I can't include it here. (In my concluding chapter I'll say as much about it as seems desirable.) All I shall discuss here, then, is the coincidence of duty and happiness considered as something that we know about from experience and can expect to show up in our present earthly life. With that restriction the alignment of happiness with duty can hardly be said to be 'currently believed'; indeed the *opposite* belief may seem to be implied by the general admission that the moral government of the world can't be completely exhibited unless there are rewards and punishments in •a future state. But if you think about it you'll see that this implication is not necessary; for one might hold that even •here virtue is always rewarded and vice punished, making the virtuous course of action always the most prudent, while also holding that these ·earthly· rewards and punishments aren't sufficient to satisfy our sense of justice. Admitting that the circumstances of a virtuous man are often so adverse that his life is less happy than that of many less virtuous people, we might still maintain that virtue will give him the most happiness that *can* be had under these circumstances. . . . And this view has certainly been held by •reputable moralists on the evidence of actual experience of human life; and seems often to be confidently asserted on similar evidence by •popular preachers and moralisers. So we should carefully and impartially examine this opinion. In tackling this at this stage in my book, I'll have to use the common concept of *duty* without further definition or analysis; but the people whose view I'm going to discuss usually hold that the moral concepts of ordinary well-meaning folk are at least approximately valid, ·and approximations are all we can have anyway·. We have seen that hedonism's generalisations must be established, if at all, by broad considerations and decisive outweighings, and with a topic like this it's pointless to take account of slight differences, claiming to weigh small portions of happiness in our mental scales.[1]

2. The view I am examining isn't likely to provoke controversy with regard to 'duties towards oneself', because this ordinarily means 'acts that tend to promote one's own happiness'. (I'm here relying on the common division of duties into •self-regarding and •social. Any adjustments that turn out to be needed—see III/2.1 and 7.1—won't invalidate the conclusions of the present chapter.) So we can confine our attention to the social part of duty, and ask: If we obey the moral rules that tell us how to behave towards others, will we always tend to secure the greatest balance of happiness to ourselves?

[1] For a similar reason I shall here treat notions of *duty* and *virtuous action* as practically equivalent. Ordinary usage of the two terms appears to indicate that they diverge somewhat; I'll discuss that in III/2.

I'll adapt Bentham's terminology, and label the pleasures that come from conforming to moral rules, and the pains that come from violating them, the 'sanctions' of these rules. These 'sanctions' can be put into two classes. **External sanctions** are

- •legal sanctions, i.e. penalties inflicted by the authority of the state; and
- •social sanctions, which are either •the pleasures to be expected from the approval and goodwill of our fellow-men and the consequent good they'll be prompted to do for us. . . . or •the trouble and losses that are to be feared from their distrust and dislike.

Internal sanctions consist in

- •the pleasurable emotion that accompanies virtuous action, or •the absence of remorse, or •pleasure resulting indirectly from the effect on the agent's mind of his maintenance of virtuous dispositions and habits.

The main importance of this classification, for our present purpose, is that the systems of rules to which these sanctions are attached may be in conflict. A community's positive [see Glossary] morality *develops*, changing in ways that affect the consciences of the few before they are accepted by the many; so that at any time the rules backed by the strongest social sanctions may fall short of, or even clash with, the ·moral· intuitions of the members of the community who have most moral insight. For similar reasons, law and positive morality may be at variance in details. A law wouldn't last long if everyone thought it would be wrong to obey it; but there could easily be laws commanding conduct that is considered immoral by some fraction of the community, especially by some sect or party that has a public opinion of its own; and a person may be connected with this fraction so much more closely than with the rest of the community that in his case the social sanction practically operates *against* the legal one.

This conflict is of great importance when we are considering whether these sanctions, so far as we can foresee them, are always sufficient to get a rational egoist to perform his social duty; for. . . . we'll have trouble proving that duty coincides with self-interest in the exceptional cases where the sanctions oppose what the agent thinks to be his duty.

But even if we set these cases aside, it still seems clear that morality's external sanctions are not alone always enough to make •immoral conduct •imprudent as well. We must indeed admit •that in an even tolerably well-ordered society—i.e. in an ordinary civilised community in its normal condition—all serious open violation of law is imprudent unless it's a part of a successful violent revolution; and further •that violent revolutions would seldom if ever be made by people who were all perfectly under the control of enlightened self-love—because such disturbances always bring general and widespread destruction of security and of other means of happiness. Still, so long as actual human beings are not all rational egoists, such times of disorder will be liable to occur; and we can't say that rational self-love clearly directs everyone to 'seek peace and live in peace' [1 *Peter* 3:11]; because disturbing the political order may present openings to wealth, fame, and power for a cool and skillful person who knows how to fish in troubled waters—openings far wider than anything he could hope for in peaceful times. In short: though an organised society composed entirely of rational egoists would tend to be stable and orderly, it doesn't follow that any individual rational egoist will always be on the side of order in any existing community.[1]

[1] What about revolutionaries aiming sincerely at general well-being? The morality of such revolutions will generally be so dubious that these cases can't provide any clear argument on either side of the question here discussed.

Anyway, in the most orderly societies that we know the administration of law and justice is never so perfect as to make secret crimes *always* acts of folly because of the legal penalties attached to them. However much these penalties may outweigh the advantages of crime, there are bound to be cases where the risk of discovery is so small that on a sober calculation •the almost certain gain will more than compensate for •the slight chance of the penalty. And, finally, in no community is the law so perfect that *no* kinds of flagrantly anti-social conduct slip through its meshes and escape legal penalties altogether or incur only penalties that are outweighed by the profit of law-breaking.

3. Well, then, how far does the •social sanction in such cases make up for the defects of the •legal sanction? No doubt

- •the hope of praise and liking and services from one's fellow-men, and
- •the fear of forfeiting these and incurring instead aversion, refusal of aid, and social exclusion,

are often large enough to lead the rational egoist to lawfulness, even in the absence of adequate legal penalties. But where •legal penalties are defective, that's exactly where ·social· sanctions are liable to fail also: social penalties no less than legal ones are evaded by secret crimes; and in cases of criminal revolutionary violence, the social sanction is apt to be seriously weakened by the party spirit enlisted on the side of the criminal. The force of •the social sanction diminishes very rapidly in proportion to the number of dissidents from the common opinion that awards •it. Disapproval that is intense and truly universal would be a penalty severe enough, perhaps, to outweigh any imaginable advantages; for a human being couldn't live happily, whatever goods he enjoyed, if he wasn't looked on in a friendly way by some of his fellows; so the conventional portrait of a tyrant as necessarily suspicious of those nearest him, even of the members of his own family, makes us think that such a life must be extremely unhappy. But when we look at *actual* tyrannical usurpers—

- •wicked statesmen,
- •successful leaders of unjustified rebellions,
- •all the great criminals who have put themselves out of the reach of legal penalties

—it seems that in an egoistic calculation of the gain and loss from their conduct the moral odium they lie under needn't count for much. This lack of esteem is expressed by only a portion of the community, and is often drowned in the loud-voiced applause of the multitude, whose admiration is largely independent of moral considerations. And there's no shortage of philosophers and historians whose judgment shows a similar independence!

So we can't say that the external sanctions of men's legal duties will always make duty coincide with ·self·-interest. Still less can we say this about moral duties that aren't covered by the law. I'm well aware of the force of what we might call 'the principle of reciprocity', through which some utilitarians have tried to prove that each person's social duties coincide with his individual interests. It goes like this: Virtues are either •useful to others or •directly agreeable to others; so they either increase the market value of the virtuous man's services, causing others to purchase them at a higher price by giving him more dignified and interesting functions; or they dispose men to please him, out of gratitude and also in order to enjoy the pleasures of his society in return. And the display of these qualities naturally spreads to others through the mere influence of example (man is an imitative animal). I'm sure that the prospect of these advantages is an adequate motive for developing many virtues and avoiding much vice. For this reason a rational egoist will generally

- •be strict and punctual in fulfilling all his engagements,
- •be truthful in his assertions, in order to win the confidence of others,
- •be zealous and industrious in his work, in order to win promotion to more honourable and lucrative jobs,
- •control any of his passions and appetites that might interfere with his efficiency;
- •not exhibit violent anger or use unnecessary harshness even towards servants and subordinates; and
- •be polite and accepting and good-humoured towards his equals and superiors in rank, showing them kindness of the sort that costs little in proportion to the pleasure it gives.

But the conduct recommended by this line of reasoning doesn't really coincide with moral duty. **(a)** Social success requires us to *appear* to be useful to others; so this motive won't restrain one from doing secret harm to others, or even from acting openly in a way that really is harmful though it isn't seen to be so. **(b)** A man may be useful to others not through his virtue but rather through his vice—or through his good and useful qualities with some unscrupulousness mixed in. **(c)** Morality tells us to do our duty towards everyone, and to do our best not to harm anyone; but the principle of reciprocity tells us to exhibit our useful qualities chiefly towards the rich and powerful, and abstain from harming those who can retaliate. It leaves us free to omit our duties to •the poor and weak if we find a material advantage in doing so, unless •they can arouse the sympathy of persons who can harm us. **(d)** Some vices—e.g. some sensuality and extravagant luxury—don't harm anyone immediately or obviously, though they tend in the long run to impair the general happiness; so few persons are strongly motivated to check or punish this kind of mischief.

In cases **(b)**–**(d)** the mere disrepute attaching to open immorality is an important consideration. But this wouldn't always be enough to turn the scales of prudence against vice; if you think it would, perhaps you haven't properly analysed the muddy and fluctuating streams of social opinion on which the reputation of individuals mainly depends, and considered the conflicting and divergent elements that they contain. Many moralists have remarked on the discrepancy in modern Europe between •the law of honour (i.e. the more important rules maintained by the social sanction of well-bred persons) and •the morality professed in society at large. But this isn't the only example of a special code diverging from the moral rules generally accepted in the community where it exists. Most religious sects and parties, and probably the majority of trades and professions, show something of this sort. I don't mean merely that special rules of *behaviour* are imposed on members of each profession, corresponding to their special social functions and relations; I mean that a special moral *opinion* is apt to grow up, conflicting somewhat with the opinion of the general public. The most striking part of this divergence consists in the approval or allowing of practices that the current morality disapprove of—

- •wild behaviour by soldiers,
- •bribery among politicians in certain times and places,
- •untruthfulness of various degrees among priests and lawyers,
- •fraud in different forms among tradesmen,

—and so on. In such cases there are strong natural inducements to disobey the stricter rule (in fact the continual pressure of these inducements seems to be what relaxed the rule in the first place); while the social sanction is weakened to such an extent that it is sometimes hard to say whether it outweighs a similar force on the other side. When

a member of one of these groups conforms to the stricter rule, if he doesn't meet with outright contempt and aversion from the other members, is at least liable to be called eccentric and fantastic, especially if by such conformity he loses advantages not only to himself but to his relatives or friends or party. This professional or sectarian allowing of immorality is often not so clear and explicit as to amount to the establishing of a rule that conflicts with •the generally received rule, but even then it is sufficient to weaken the social sanction in favour of •the latter. And in addition to these special divergences, most civilised societies have *two* degrees of positive morality, each maintained somewhat by common consent: a stricter code that is publicly taught and avowed, and a laxer set of rules that is privately accepted as the only code that can be strongly supported by social sanctions. In most cases, a man can refuse to conform to the stricter code without being

- •excluded from social intercourse,
- •seriously hindered in professional advancement, or
- •seriously disliked by any of those whose society he will most naturally seek;

and in that case the mere loss of a certain amount of reputation isn't likely to be felt as a very grave evil, except by someone who is especially sensitive to the pleasures and pains of reputation. And there seem to be many men whose happiness doesn't depend on the approval of the moralist—and of people who support the moralist—to such an extent that it would be prudent for them to purchase this praise by any great sacrifice of other goods.

4. Thus, if the conduct prescribed to an individual by the openly accepted morality of his community coincides with what rational self-love would prompt, this must often be solely or chiefly because of the *internal* sanctions. In considering these I shall set aside the pleasures and pains involved in the anticipation of rewards and punishments in a future life: my topic is the calculations of rational egoism as performed without taking into account any feelings that are beyond the range of experience, and it will be more consistent with that to exclude also the pleasurable or painful anticipations of such feelings.

Let us start with the satisfaction that accompanies the performance of duty—meaning duty *as such*, leaving out any consequences—and the pain that follows on its violation. After the discussions of chapters 3 and 4 you won't expect me to try to weigh these pleasures and pains *exactly* against others; but ·inexactness can get us somewhere·. I see no empirical evidence that such feelings are always intense enough to turn the balance of prospective happiness in favour of morality. This will hardly be denied in application to isolated acts of duty.... The call of duty has often impelled a soldier or other public servant, or the adherent of a persecuted religion, to face certain and painful death under circumstances where it could be avoided with little if any loss of reputation. To prove this is reasonable from an egoistic point of view, we have to assume that in any such case the evasion of duty would bring so much pain[1] that the rest of the person's life would be hedonistically worthless. That assumption would be paradoxical and extravagant. Nothing

[1] I am here including in moral pain (pleasure) all pain (pleasure) that is due to sympathy [see Glossary] with the feelings of others. This is not the place for me to discuss fully the relation of sympathy to moral sensibility; but I am sure •on the one hand that these two emotional susceptibilities are actually distinct in most minds, whatever they may have been originally; and •on the other hand that sympathetic and strictly moral feelings are almost inextricably blended together in the ordinary moral consciousness; so that my present argument doesn't need to draw the line between them. But I shall look into sympathy, as the internal sanction that utilitarians specially emphasize, in the concluding chapter of this treatise [page 243.]

that we know about most people in any society suggests that their moral feelings taken alone form such a weighty element in their happiness. And a similar conclusion seems irresistible even in less extreme cases, where it's not •life but •a considerable share of ordinary sources of happiness that a man is called on to give up for virtue's sake. Can we say that all men—or even most men—are so constituted that

- •the satisfactions of a good conscience would certainly repay them for such sacrifices, or that
- •the pain and loss involved in them would certainly be outweighed by the remorse that would follow the refusal to make them?[1]

Few if any writers, however, have explicitly gone as far as this. What Plato in his *Republic* and other writers on his side have tried to prove is not •that for each person at each moment duty will produce more happiness than any alternative, but rather •that it's in each person's over-all interests to choose the life of the virtuous man. But it's hard to make this ·much weaker thesis· even probable. To see this, look at the lines of reasoning by which it is commonly supported.

Plato represents the soul of the virtuous man as a well-ordered polity—·i.e. an as-it-were-political structure·—of impulses, in which every passion and appetite obeys the rightful sovereignty of reason and operates only within the limits that reason lays down. He contrasts the tranquil peace of such a mind with the disorder of one where a series of lower impulses or a single ruling passion lords it over reason; and he asks which is the happier, even apart from external rewards and punishments. Well, we can grant all that Plato claims here and yet be no nearer answering the question before us. For the issue we are studying isn't

reason versus passion

but rather, in Butler's terminology,

conscience versus rational self-love.

We're supposing the egoist to have all his impulses under control, and are only asking how this control is to be exercised. We have seen that the way of life best calculated to achieve the end of self-interest appears *prima facie* to diverge at certain points from what men are prompted to by a sense of duty. To maintain Plato's position we would have to show that this appearance is false, and that a way of life which under certain circumstances leads us to pain, loss, and death is still what self-interest requires. Is our nature such that this anti-egoistic kind of regulation is the only one possible for us—i.e. that we have to choose between this and no regulation at all? Of course not! It is easy to imagine a rational egoist strictly controlling his passions and impulses, including his social sentiments, within such limits that indulging them doesn't involve the sacrifice of something that would please him more; and we seem to have encountered many people who approximate to this type at least as closely as anyone else approximates to the ideal of the orthodox moralist. Hence if the rules of conscience are to be demonstrably the best means to the individual's happiness, it must be because the over-all way of life maintained by self-love involves an over-all sacrifice of pleasure, as compared with the way of life maintained by

[1] A striking confirmation of this comes from Christian writers of the 18th century who treat the *moral* unbeliever as a fool who sacrifices his happiness both here and hereafter. Most of these writers were earnestly engaged in the practice of virtue, yet this practice hadn't made them love virtue so much that they would prefer it, even under ordinary circumstances, to the sensual and other enjoyments that it excludes. It seems absurd, then, to suppose that for people who *haven't* developed and strengthened their virtuous impulses by virtuous habits the pain that might afterwards result from resisting the call of duty would always be enough to neutralise all other sources of pleasure.

conscience. And if that is how things stand, it can only be because of the special emotional pleasure that comes with satisfying the moral sentiments or the special pain or loss of happiness that results from repressing and violating them.

By now you have probably noticed a fundamental difficulty:

> If a man thinks it reasonable to seek his own interest, he clearly can't disapprove of any conduct that comes under this principle, or approve any that goes against it. So the pleasures and pains of conscience •can't enter into the calculation of whether a certain line of behaviour is in accordance with rational egoism, because they •can't attach themselves in the egoist's mind to any way of behaving that hasn't already been decided, on other grounds, to be reasonable or the reverse.

There is some truth in this, but we must here recur to the distinction drawn in I/3.1 between •the general impulse to do what we believe to be reasonable and •special likings or aversions for special kinds of conduct independent of their reasonableness. In the moral sentiments of ordinary men these two kinds of feeling are blended together, because people generally think that the rules to which the common moral sentiments are attached are somehow reasonable. But we can conceive of the two as separated; and we actually observe such a separation when a man is led by a process of thought to a moral standpoint different from the one he has been trained in; for his mind will retain some quasi-moral likings and aversions that are no longer sustained by his deliberate judgment of right and wrong. So there's every reason to believe that most men, however firmly they might adopt the *principles* of egoistic hedonism, would still have *feelings* prompting them to perform duties commonly recognised in their society, without believing that the actions prompted by such feelings were reasonable and right. For such sentiments would always be powerfully supported by the sympathy of others, and their expressions of praise and blame, liking and aversion; and since it is agreed that the conduct commonly recognised as virtuous generally coincides with what enlightened self-love would dictate, a rational egoist's habits of conduct will naturally foster these (for him) 'quasi-moral' feelings. So our question is not: 'Should the egoist cherish and indulge these sentiments up to a certain point?', because everyone will answer Yes to that. Our question is this: 'Can the egoist consistently encourage these 'quasi-moral' sentiments to grow so much that they'll always prevail over the strongest opposing considerations—i.e. does prudence require him to give them their heads, letting them carry him where they will? We have already seen evidence that rational self-love will best achieve its end by limiting its conscious operation and allowing free play to disinterested impulses; can we accept the further paradox that it is reasonable for it to abdicate altogether its supremacy over some of these impulses?

When you think about it, I think you'll see that this abdication of self-love is not something that could happen in the mind of a sane person who still regards his own interest as the reasonable ultimate end of his actions. Such a man may decide to devote himself unreservedly to the practice of virtue, with no detailed thoughts about what seems to be in his interests; and by living up to this decision he may gradually acquire strong habitual tendencies to acts in that way. But these habits of virtue can't ever become strong enough to gain irresistible control over a sane and reasonable will. When virtue demands from such a man an extreme sacrifice—one that is too imprudent for him to ignore—he must always be able to move out of his habit of virtue and deliberate afresh, controlling his will in a way that

doesn't bring in his past actions. You may think:

> Although an egoist retaining his belief in rational egoism can't thus *abandon* his will to the sway of moral enthusiasm, it remains the case that if he were to change his conviction and *prefer* duty to interest he would find that this preference brings him an over-all gain in happiness. •The pleasurable emotions that accompany the kinds of virtuous or quasi-virtuous habits that are compatible with sticking to egoistic principles are so inferior to •the raptures that accompany the unreserved and passionate surrender of the soul to virtue—speaking only of raptures in this life, leaving the after-life out of this—that it really is in a man's interests to obtain, if he can, the convictions that make this surrender possible, even though it might sometimes lead him to act in a manner that is in itself undoubtedly imprudent.

[In other words: There is a rational-egoist case for living always virtuously, even if it would be psychologically impossible for a rational egoist to make the case and act on its conclusion.] This is certainly tenable, and I am quite disposed to think it true of persons with specially refined moral sensibilities. And I can't conclusively prove that it isn't true of everyone (the hedonistic calculus isn't good enough for that); but I do say that it seems to be opposed to the broad results of nearly everyone's experience. Observation convinces me that most men are so constituted as to feel the pleasures (and pains) arising from conscience far less keenly that pleasures and pains from some other sources—gratifications of the senses, the possession of power and fame, strong human affections, the pursuit of science, art, etc.—so that in many cases not even early training could have given the moral feelings the required predominance. . . .

To sum up; although the performance of duties towards others and the exercise of social virtue seem to be *generally* the best means to the individual's happiness, and it is easy to exhibit this alignment of virtue with happiness in speeches to a crowd, when we carefully analyse and estimate the consequences of virtue to the virtuous agent, it appears improbable that this alignment is complete and universal. We can *conceive* of its becoming perfect in a Utopia where men agreed as much on moral questions as they do now on mathematical questions, where law was in perfect harmony with moral opinion, and all offences were discovered and duly punished; or we can conceive achieving the same result by intensifying the moral feelings of all members of the community, without any external changes. . . . But just in proportion as existing societies and existing men fall short of this ideal, rules of conduct based on the principles of egoistic hedonism seem liable to diverge from those that most men are accustomed to recognise as prescribed by duty and virtue.

Chapter 6: Deductive hedonism

1. In chapter 5 we saw reason to conclude that although obedience to recognised rules of duty ordinarily tends to promote the agent's happiness, there's no good empirical evidence that the performance of duty is a universal or an infallible means to happiness. Even if this weren't so—even if it were demonstrably reasonable for the egoist to choose duty at all costs under all circumstances—the systematic attempt to act according to this principle, understanding 'duty' in terms of common notions of morality, would still bump into our problem of finding the right way to seek happiness. That is because common morality allows us to seek our own happiness (within limits) and even seems to regard it as morally prescribed;[1] and still more emphatically tells us to promote the happiness of others with whom we are in various ways specially connected; so that our questions about how to fix and measure the elements of happiness would still require some kind of answer. [In short: part of our duty involves seeking happiness. How are we to go about seeking happiness? The answer 'Seek it by doing your duty' is unhelpfully circular.]

The remaining question: How far can a scientific investigation of the causes of pleasure and pain help us to deal with this practical problem?

To decide on hedonistic grounds how to act, we obviously need not only to measure pains and pleasures but also to know how to produce or avert them. In most important prudential decisions, complex chains of consequences are expected to intervene between our initial volition and the feelings that we are ultimately aiming to produce; and how accurately we can predict each link in these chains obviously depends on what we know (implicitly or explicitly) about cause-effect relations among various natural phenomena. But the details of how to produce specific kinds of pleasure don't belong in a general treatise on the method of ethics; rather, they'll have to come from this or that special art [see Glossary] subordinate to the general art of conduct. Some of these subordinate arts have a more or less scientific basis, while others are still at the merely empirical stage [here = 'haven't gone beyond the accumulation of anecdotal data']; a detailed plan for seeking *health* belongs to the systematic art of hygiene, based on physiological science; but if we are aiming at *power* or *wealth* or *domestic happiness* the help we get from the experience of others will mainly be unsystematic—advice relative to our own special circumstances, or accounts of success and failure in situations like ours. Either way, the exposition of such special arts doesn't seem to come within the scope of the present treatise, and it couldn't help us in dealing with the measurement difficulties that we have considered in previous chapters.

You may think that a knowledge of the causes of pleasure and pain could carry us beyond the determination of the means of gaining particular kinds of pleasure and avoiding particular kinds of pain, and replace the empirical-reflective method whose defects we have been studying by some deductive method of evaluating the elements of happiness.[2]

[1] 'It would seem that an appropriate concern about our own interest or happiness and a reasonable attempt to secure and promote it. . . is virtuous, and the contrary behaviour faulty and blameworthy.' Butler (in 'The Nature of Virtue', appended to *The Analogy of Religion*).

[2] This view is suggested by Spencer's statement in a letter to Mill. . . .that 'it is the business of moral science to deduce, from the laws of life and the conditions of existence, what kinds of actions necessarily tend to produce happiness, and what kinds to produce unhappiness', and that when it has

A hedonistic method that entirely does without direct estimates of the pleasurable and painful consequences of actions? That is hardly more conceivable than a method of astronomy that does without observations of the stars! But it is conceivable that by induction from cases where empirical measurement is easy we may obtain generalisations that will give us more trustworthy guidance than such measurement can do in complicated cases; we may be able to discover some general mental or physical concomitant or antecedent of pleasure and pain—one that is easier to recognise, foresee, measure, and produce or avert than pleasure and pain themselves are in such cases. I'm willing to hope that this escape from the empirical hedonism's difficulties may one day be open to us; but I can't see that it is available yet. We don't have *now* any satisfactorily established general theory of the causes of pleasure and pain; and the theories that have gained some acceptance—as partly true or as probable—are manifestly not right for our present needs.

It's easy to explain why it is hard to find an all-purpose theory of the causes of pleasures and pains. Like other mental facts, pleasures and pains presumably occur along with certain cerebral nerve-processes, the details of which we don't know. So we can look for their causes either in prior physical or prior mental facts. But in one important class of cases the main knowable antecedents are obviously physical, while in another they are obviously mental; and the problem is to establish a theory that applies equally to both classes. . . . In the case of pleasures and pains—especially pains—connected with sensation, the most important knowable antecedents are clearly physical. . . . Under ordinary conditions the pains of sensation—probably the most intense in the experience of most persons—invade and interrupt our mental life from outside us; it would be idle to look for the main causes of their intensity or quality among antecedent mental facts. This is not so true of the most prominent pleasures of sense, because antecedent desire, if not absolutely required for such pleasures, seems to be required for them to reach a high degree of intensity. Still the main causes of these desires themselves are clearly physical states and processes—not merely neural ones—in the organism of the sentient individual; and this is also true of a more indefinite kind of pleasure that is an important element in ordinary human happiness, namely the 'well-feeling' that accompanies and is a sign of physical well-being.

But when we investigate the causes of •the pleasures and pains that belong to intellectual activities or the play of personal affections, or of •the pleasures (and to some extent pains) that belong to the contemplation of beauty (or ugliness) in art or nature, no physiological theory can take us far because we don't know what the neural processes are that accompany or precede these feelings.

That is my general conclusion, and I'll further illustrate and explain the grounds for it in the rest of this chapter. As for an exhaustive discussion of either psychological or physiological theories of the causes of pleasure and pain—I can't even *attempt* anything like that. I shall confine myself to certain leading generalisations that seem to have a special interest for students of ethics, either •because ethical

done this 'its deductions are to be recognised as laws of conduct that are to be conformed to irrespective of a direct estimate of happiness or misery'. [Sidgwick goes on to say that Spencer says he meant this only for 'an ideal society'; that he (Sidgwick) will consider such ideals in IV/4; and that at present he is] only concerned with the question how far any deductive ethics could furnish practical guidance to an individual seeking his own greatest happiness here and now.

motives help to cause their acceptance or •because though inadequately grounded as general theories they appear to have a partial and limited value for practical guidance.

[This chapter will refer to Sir William Hamilton—*not* the Irish Sir William Hamilton (1805–65), a distinguished physicist, astronomer and mathematician, but the Scottish Sir William Hamilton (1788–1856), an undistinguished philosopher. (Of a logical controversy that he got into with Augustus De Morgan, C. S. Peirce wrote that 'the reckless Hamilton flew like a dor-bug into the brilliant light of De Morgan's mind'.)]

2. Let us begin by considering a theory, primarily psychological, which. . . .is derived from Aristotle,[1] and is still current in one form or another. [Sidgwick cites two French writers, as well as G. F. Stout, 'to whom I will refer later'.] It's the thesis expressed by Sir William Hamilton (·in no. 42 of his *Lectures in Metaphysics*·) in the following propositions:

> •'pleasure is the reflex of the spontaneous and unimpeded exercise of a power of whose energy we are conscious;
>
> •pain is a reflex of the over-strained or repressed exercise of such a power'.

The phrases suggest active as distinct from passive states; but Hamilton explains that 'energy' and similar terms 'are to be understood to refer to *all* the conscious processes of our higher and lower life', because consciousness itself implies more than a mere passivity of the subject. But the theory is pretty clearly constructed primarily to fit the pleasures and pains of the intellectual life as such, and has to be *stretched* to cover an important class of the pleasures and pains of man's animal life. Hamilton explains his term **(a)** 'spontaneous' as implying the absence of 'forcible repression' or 'forcible stimulation' of the power that is exercised; and explains **(b)** 'unimpeded' in terms of the absence of obstacles or hindrances in the object that the faculty is dealing with. But these terms seem to have no clear mental import in application to organic sensations that are in the ordinary sense 'passive'. The feelings and vague representations of bodily processes that constitute consciousness of a toothache are as free from conscious repression or stimulation as those that constitute the consciousness that accompanies a warm bath. . . .

Indeed, the theory's one-sidedness seems to be exactly what gives it ethical interest and value. It tends to correct a commonplace error in the estimate of pleasure, by focusing on a class of pleasures that ordinary pleasure-seeking probably undervalues—the ones that especially belong to a life filled with strenuous activity, whether purely intellectual or practical and partly physical.[2] In the same way it effectively clears up the popular blunder of regarding labour as normally painful •because some labour is so and •because the pleasures of relief from toil are in most people's experience more striking than the pleasures of strenuous activity. But even if we limit the theory to the pleasures and pains immediately connected with voluntary activity—intellectual or physical—it strikes me as lacking in •definite guidance and in •adequate theoretical precision. It seems to imply that the exercise of our powers is always

[1] Aristotle's own theory is, briefly, •that every normal sense-perception or rational activity has its corresponding pleasure, the most perfect being the most pleasant; and •that the most perfect for any faculty is the exercise of the faculty in good condition on the best object. The pleasure follows the activity immediately, giving it a kind of finish, 'like the bloom of youth'. Pleasures vary in kind, as the activities that constitute life vary; the best pleasures are those of the philosophic life.

[2] In Aristotle's exposition of this theory—which for him is only a theory of pleasure—the ethical motive of exhibiting the philosophic life as preferable (in the pleasures it provides) to that of the sensualist is unmistakable.

made less pleasant by the presence of obstacles; but this is obviously not true either of mainly intellectual activities or of mainly physical ones. Some obstacles undeniably *increase* pleasure by inviting force and skill to overcome them, as is clearly shown in the case of games and sports. [Sidgwick discusses possible ways of making Hamilton's theory safer, less vulnerable to refutation, and faults them for reasons having to do with the fact that whether an impediment to my activity causes me pain depends on whether the impediment is stopping me from achieving my goal. He sums up:] It is a fundamental defect in Hamilton's theory, even in its more limited application, that it ignores the teleological [= 'goal-seeking'] character of normal human activity.

This defect is avoided in a variant on the theory that a recent writer has adopted. In his *Analytic Psychology* xii/2 Stout writes:

> 'The antithesis between pleasure and pain coincides with the antithesis between free and impeded progress towards an end. Unimpeded progress is pleasant in proportion to the intensity and complexity of mental excitement. An activity that is. . . thwarted and retarded. . . is painful in proportion to its intensity and complexity and to the degree of the hindrance.'

He admits that it is hard to apply this to the pleasures and pains of the senses; and unlike Hamilton he explicitly recognises that 'a struggle with difficulties that is not too prolonged or too intense may enhance the pleasure of success out of all proportion to its own painfulness'. But this admission makes the theory unimportant from our present practical point of view, whatever may be its theoretical value. Also, I think Stout should have recognised more explicitly the way in which •what pleasures and pains accompanied your activity depend on •what you wanted to achieve by them. When desire is strong, hopeful effort to overcome obstacles to success tends to be correspondingly pleasurable—apart from actual success—while disappointment or the fear of it tends to be painful; but when desire is not strong, the shock of thwarted activity and unfulfilled expectation may be actually agreeable. When I take a walk for pleasure, intending to reach a neighbouring village, and find an unexpected flood crossing my road, if I have no strong motive for reaching the village the surprise and consequent re-routing of my walk will probably be on the whole a pleasurable incident.

The importance of eager desire as a condition of pleasure is ethically significant, because it provides the psychological basis for •the familiar advice to repress desires for ends that are unattainable or incompatible with the course of life that prudence marks out; and for •the somewhat less trite advice to encourage and develop desires that push in the same direction as rational choice.

. . . .Spencer maintains that pains are the mental concomitants of •excessive or •deficient actions of organs, while pleasures are the concomitants of activities that are neither excessive nor deficient [*Psychology* ix/128]. In considering this theory I'll take pains and pleasures separately, because the theory is obviously based primarily on experiences of pain, especially of the pains of sense, which Hamilton's theory seemed obviously wrong about. We encounter many cases where pain is obviously caused by excessive stimulation of nerves: if we gradually increase the intensity of sensible heat, pressure, muscular effort, at some point we encounter pain; 'deafening' sounds are highly disagreeable; and to confront a tropical sun with unprotected eyeballs would soon become torture. And, as Spencer points out, some pains come from the excessive actions of organs whose normal actions don't produce any feelings—e.g. when the digestive system is overloaded. But in none of these cases is it clear that pain comes from a mere intensification in *degree* of the

action of the organ in question; and not rather through some change in the *kind* of action—some shapeless disintegration or disorganisation. Think for example of the pains due to wounds and diseases, and even of the digestive discomforts that arise from an improper kind rather than an improper quantity of food. [Sidgwick says that hunger as such isn't painful, and that when it is accompanied by pain one has a strong sense that something is not merely *too intense* but *wrong*, *disordered*. Also:] In the case of emotional pains and pleasures, the notion of *quantitative* difference between the corresponding cerebral nerve-processes seems entirely out of place. The pains of shame, disappointed ambition, wounded love, don't seem to be distinguishable from the pleasures of fame, success, reciprocated affection, by any difference of intensity in the impressions or ideas accompanied by the pleasures and pains respectively.

Anyway, empirical evidence supports 'excessive action' of an organ as a cause of pain far more clearly than 'deficient action'. This evidence, has led Wilhelm Wundt and some other psychologists to the view that

> no kind of sensation is absolutely pleasant or unpleasant; when a sensation of *any* kind grows in intensity it reaches a point at which it becomes pleasurable, and then further up the intensity scale it becomes painful (having rapidly passed through a neither-pleasurable-nor-painful stage).

My experience doesn't support this generalisation. I agree with Gurney [*Power of Sound* I/2] that 'of many tastes and odours the faintest possible suggestion is disagreeable', while other feelings resulting from stimulation of sense-organs appear to remain highly pleasurable at the highest possible degree of stimulation.

[Sidgwick remarks that neither of the two theories of pain—•that it comes from neural excess, •that it comes from neural disorder—gives us any useful practical guidance because we don't have the neurological facts. Also:] No-one doubts that wounds and diseases are to be avoided under all ordinary circumstances; and in an exceptional case where we have to choose them as the least of several evils, our choice wouldn't be helped knowledge of exactly *how* they cause pain.

Turning from pain to pleasure, you might think this:

> The generalisation that we have been considering at least gives us a psycho-physical basis for the ancient maxim that we should 'avoid excess' in the pursuit of pleasure.

Sidgwick's next sentence: But we have to observe that the practical need of this maxim is largely due to the qualifications which the psycho-physical generalisation requires to make it true.

apparently meaning: The cases where the 'avoid excess' maxim is needed are mostly ones where the psycho-physical generalisation is not true *as it stands*.

Thus the 'avoid excess' maxim is especially needed in the important cases where over-stimulation is followed by pain not •at once but •after an interval of varying length. For many people drinking alcohol remains pleasurable right up to the point of excess, where the brain can no longer do its job; it's on 'the morning after' that the pain comes; and perhaps with 'well-seasoned' drinkers it comes only after many years of habitual excess. And another point: when excess leads from pleasure to pain, the organ involved in the pain isn't always the one that first gave the pleasure. When we are tempted to eat too much, the seductive pleasure is mainly due to the nerves of taste, which are not over-worked; the pains come from the organs of digestion, whose faint, vague pleasures weren't enough on their own to tempt the high-living person

to over-eat. In the case of dangerous *mental* excitements the penalty for excess is usually even more indirect.

Let's grant that pleasure like virtue resides somewhere in the middle, this proposition gives no practical directions for getting pleasure. Granted that •excessive and •deficient activities of organs cause pain, the question still remains: In any given case, what fixes the lower and higher limits between which action is pleasurable? I'll come to Spencer's answer to this shortly, but first I want to discuss a question—equally obvious, though Spencer doesn't explicitly mention it— namely:

> Why is it that among the normal activities of our physical organs that have counterparts in consciousness, only some are pleasurable in any appreciable degree, while many if not most are nearly or quite indifferent [see Glossary].

It seems undeniable, for example, that while tastes and smells are mostly either agreeable or disagreeable, most sensations of touch and many of sight and sound are not appreciably[1] either, and that in the daily routine of healthy life, eating and drinking are ordinarily pleasant whereas dressing and undressing, walking and muscular movements generally, are practically indifferent.

Stout has suggested that the explanation is to be found in the operation of habit, but this seems to me wrong. Actions do through frequent repetition tend to become automatic and lose their conscious counterparts; and hedonic indifference certainly seems in some cases to be a stage through which such actions pass on the way to unconsciousness. A business walk in a strange town is normally pleasant because of the novelty of the sights; a similar walk in one's home-town is usually indifferent, or nearly so; and if one's attention is strongly absorbed by the business, the walk may be performed to a great extent unconsciously. But the operations of habit often have the opposite effect of making pleasant activities that were at first indifferent or even disagreeable—as with acquired tastes, physical or intellectual. . . . Spencer, indeed, regards such experiences as so important that he infers from them that 'pleasure will eventually accompany every mode of action demanded by social conditions'. This seems unduly optimistic, however, because of the cases I have mentioned where habit produces hedonic indifference, and also because a third effect of habit, which is to make gradually more irksome actions that were at first indifferent or even pleasant. Our intellect gradually wearies of monotonous activities, and the boredom may sometimes become intense; and the taste of a kind of food that was at first agreeable may become disgusting through monotony.

So we have to look for some quite different explanation for the varying degrees in which pleasure accompanies normal activities. [Sidgwick reports a theory according to which pleasure is greater if the relevant nerves are acting faster than they usually do. He has little trouble shooting it down, and proceeds to look elsewhere.]

Of the various theories that have been offered to explain the fact we are trying to explain, none has acquired anything like general acceptance as covering the whole ground. I select for discussion one of them that has special ethical interest. (It is in Stout, *Analytic Psychology* xii.4.)

According to this hypothesis, the organic process accompanied by pleasure is a 'restoration of equilibrium' after 'disturbance'; so that when certain normal activities aren't accompanied by appreciable pleasure, that is because there

[1] I say 'appreciably' because there's controversy among psychologists about whether any states of consciousness are *strictly* neutral or indifferent. The issue seems to me unimportant from a practical point of view.

was no prior disturbance. This is obviously right for the pleasure of relief after physical pain or after the strain of great anxiety, and the pleasure of rest after unusual exertions, intellectual or muscular. But these cases, though by no means rare, are not central in a normal life. When we try to apply this theory to sense-related pleasures generally, we are faced with the indefiniteness of the notion of *equilibrium*, as applied to the processes of a living organism. Our physical life consists of a series of changes most of which recur (with slight variations) at short intervals; and it's hard to see why we should attach the idea of *disturbance* or *restoration of equilibrium* to any one of these normal processes rather than any other—e.g. why the condition of •having expended energy should be regarded as a departure from equilibrium any more than the condition of •having just eaten food. The fact is that this hypothesis doesn't *at all* fit normal pleasures of sense unless we pass from the physiological to the psychological point of view, and bring into the story the mental state of *desire* as a consciously *unrestful* condition, the essence of which is a felt impulse to move from this state towards the attainment of the desired object. Our hypothesis can then take this unrestful consciousness as a sign of what from a physiological point of view is 'disturbance of equilibrium'; and the satisfaction of desire can be taken to be, physiologically, a restoration of equilibrium. On this interpretation of it, the hypothesis becomes clearly true of the gratifications of sensual appetite that form the most prominent element of the pleasures of the senses, as the man in the street thinks of them.

I have already noted that through a wide-spread confusion of thought desire has often been regarded as a sort of pain. In line with that. the theory we are now considering was originally launched with an ethical motive, namely to down-play the commonly overvalued pleasures of satisfied bodily appetite by emphasising their inseparable connection with antecedent pain. The attempt fails, however, because the appetite that must precede pleasure is, though unrestful, not appreciably painful.

In any case, even if we admit that the physical counterpart of conscious desire either •is or •comes from a 'disturbance of equilibrium', this theory obviously doesn't cover the whole range of the pleasures of sense. The simple pleasures of the special senses *don't* have to be preceded by conscious desire; normally no sense of want has preceded the experience of pleasant sights, sounds, odours, flavours, or of the more important pleasures. . . .that we call aesthetic. [In some special cases, Sidgwick adds, aesthetic pleasures may be preceded by a strong desire for them or sense of being deprived of them; you could call these 'disturbances'; but there's no basis for extending this special pattern to] the ordinary cases where pleasures of this kind are experienced without any antecedent consciousness of desire or deprivation.

I may have said enough to support my general conclusion that psychophysical theories about the causes of pleasure and pain don't give us a basis for a deductive method of practical hedonism. I'll just add that the difficulties facing any such theory seem especially great for the complex pleasures that we call 'aesthetic'. [High-level aesthetic pleasure, Sidgwick says, does involve a very 'complex state of consciousness', but no-one would accept that the complexity is enough for the pleasure. However subtly we describe the objective relations of elements in a delightful work of art, we must always feel that there could be something •answering exactly to that description while •providing no aesthetic delight. The 'touch' that leads to delight is an instinctive *sense* of how the elements work together in the art-work; it can't be replaced by an *inference* from a premise

describing the complexity to a conclusion about aesthetic value (and thereby about pleasure). Sidgwick adds:] This is true even is we set aside the wide divergences among the aesthetic sensibilities of individuals. So there is even less need to argue that for an individual seeking his own greatest happiness the only way to estimate aesthetic pleasures is by a mainly inductive and empirical method.

3. From discussing a •psychophysical theory of pleasures and pains I now turn to one that is •biological: still concerned with organic states or events that accompany or immediately precede pleasures and pains, it focuses not on •the actual present characteristics of those states and events but on •their relations to the life of the organism as a whole. I mean the theory that 'pains are the correlatives of events that are ·potentially· *destructive* of the life of the organism, while pleasures are the correlatives of events that are *preservative* of its welfare'. [Spencer starts little differently, but Sidgwick says that 'destructive' and 'preservative' adequately express what Spencer ends up with.]

Spencer's argument is as follows (in his own words):

> If we substitute for 'pleasure' the equivalent phrase 'feeling that we seek to bring into consciousness and retain there', and substitute for 'pain' the equivalent phrase 'feeling that we seek to get out of consciousness and to keep out', we see at once that
>
> > if the states of consciousness that a creature tries to maintain are the correlatives of injurious actions, and if the states of consciousness that it tries to expel are the correlatives of beneficial actions, the creature must quickly disappear because of its persistence in doing what harms it and avoiding what helps.
>
> In other words, the only species that can have survived are ones in which, on the average, agreeable or desired feelings generally accompanied activities conducive to the maintenance of life, while disagreeable and habitually-avoided feelings accompanied activities directly or indirectly destructive of life; and other things being equal there must always have been the most numerous and long-continued survivals by species in which these adjustments of feelings to actions were the best, tending ever to bring about perfect adjustment. [All quotations from Spencer's *Principles of Psychology* and *Data of Ethics*.]

This summary deduction may well have value for certain purposes; but it's easy to show that substituting 'preservation' for 'pleasure' as the end directly aimed at it doesn't provide an adequate basis for a deductive method of seeking maximum happiness for the individual. For one thing, Spencer only affirms the conclusion to be true, as he rather vaguely says, 'on the average'; and it's obvious that though •the tendency to find harmful acts pleasant or preservative acts painful must be a disadvantage to any species in the struggle for existence, if •it exists only to a limited extent it may be outweighed by advantages, so that the organism that has it may survive in spite of it. It is obvious *a priori* that this *can* happen, and we know from common experience that it often *does*, as Spencer admits. [He quotes Spencer to this effect and remarks:] This seems to be a sufficient objection to basing a deductive method of hedonism on Spencer's general conclusion. It's a notorious fact that civilised men take pleasure in various forms of unhealthy conduct and find conformity to the rules of health irritating. . . . And this it is easy to explain this on [he must mean: reconcile this with] the 'evolution hypothesis', because that hypothesis doesn't rule out the possibility that

> the development of the nervous system in human beings brings with it intense susceptibilities

> to pleasure from non-preservative processes, if the preservation of the individuals who have the susceptibilities is otherwise adequately provided for.

This latter condition is obviously satisfied for leisured people in civilised society, whose needs for food, clothing, shelter, etc. are abundantly supplied through the. . . .institution of private property; and I don't know any empirical evidence that •a cultivated man's keen and varied pleasures enable him to live longer than •a man who goes through a comparatively dull round of monotonous routine activity, interspersed by slightly pleasurable intervals of rest and play.

4. If the individual isn't likely to obtain a maximum of pleasure by aiming merely at preservation, perhaps he will do better by aiming at 'quantity of life'. [That odd phrase has only two occurrences in this work, both in the present section. What Sidgwick means by it has to be gathered from his uses of it.] It is of course true of neural events accompanied by conscious pleasure that the more of them there are the happier we'll be. But even if we assume that the more intense and full life is 'on the average' the happier, it doesn't follow that we'll get maximum pleasure by aiming merely at intensity of conscious states; for we experience intense pains even more indubitably than intense pleasures; and in the 'full tides of soul' in which we seem to be most alive, pain can be mixed in with pleasure in almost any proportion. Also, we often experience very intense excitement that isn't clearly pleasurable or painful—e.g. in laboriously struggling with difficulties and perplexing conflicts of which the issue is doubtful.

It may be replied that 'quantity of life' should imply not merely •intensity of consciousness but •multiplicity and variety—a harmonious and many-sided development of human nature. Experience does support the view that men lose happiness by allowing some of their faculties or capacities to wither and shrink from disuse, thus not leaving themselves sufficient variety of feelings or activities; and we know that due exercise of most—if not all—of the bodily organs is indispensable for the health of the organism, and that the health maintained by this balance of functions is a better source of happiness than the unhealthy over-exercise of any one organ can be. Still, the harmony of functions needed for health seems to be very elastic, allowing for a wide margin of variation, as far as the organs under voluntary control are concerned. For example, a man who exercises only his brain will probably be ill in consequence; but he can exercise his brain much and his legs little, or vice versa, without any unhealthy results. Also, if the proposition that *a varied and many-sided life is the happiest* were to serve as a basis for deductive hedonism, we would have to make it precise, which we can't. That's because there is also truth on the other side: the more we exercise any faculty with sustained and prolonged concentration, the more pleasure we derive from such exercise, up to the point where it becomes wearisome or turns into a semi-mechanical routine that makes the mind dull and slack. It is certainly important for our happiness that we keep within this limit; but we can't fix it precisely in any particular case without experience of that individual; especially as there seems always to be some weariness and tedium to be resisted and overcome on the way to our bringing our faculties into full play and having the full enjoyment of our labour. Similarly with passive emotional consciousness: if too much sameness of feeling results in slackness, too much variety inevitably involves shallowness. The point where concentration ought to stop, and where dissipation begins, varies from man to man, and has to be decided by the specific experience of individuals.

There's another and simpler way of understanding the maxim of 'giving free development to one's nature' [Sidgwick

writes as though he had already introduced that phrase, but he hasn't]. We could take it to mean yielding to spontaneous impulses rather than trying to govern them by elaborate forecasts of consequences. The injunction to do this gets scientific justification from the theory that spontaneous or instinctive impulses are really effects of previous experiences of pleasure and pain on the organism in which they appear or its ancestors. This has led to the thesis that in complicated problems of conduct experience will 'enable •the constitution to estimate the respective amounts of pleasure and pain consequent on each alternative', where it is 'impossible for •the intellect' to do this; and 'will further cause the organism instinctively to avoid the conduct that produces on the whole most suffering'.[1] There is an important element of truth in this; but nothing that we know or can plausibly conjecture regarding biological evolution supports any broad conclusion that non-rational inclination is a better guide than reason to individual happiness. Natural selection fosters impulses favouring the preservation of the species rather than the pleasure of the individual, but I'll set that aside. Granting that every sentient organism tends to adapt itself to its environment in such a way as to acquire instincts that help to guide it to pleasure and away from pain, it doesn't follow from this that in the human organism the kind of adaptation that **(a)** involves the unconscious development of instinct is to be preferred to the kind of adaptation that **(b)** comes from conscious comparison and inference. [Sidgwick goes on to say that an empirical comparison of the success-rates of **(a)** and **(b)** wouldn't show **(a)** as a clear winner.] However true it may be that in certain cases instinct is on the whole a safer guide than prudential calculation, it seems that the only way we can discover which cases these are is by careful reflection on experience; we can't determine the limits to which prudential calculation may prudently be carried, except by this very calculation!

We seem, then, forced to conclude that there is no scientific short-cut to the ascertainment of the right means to the individual's happiness; every attempt to find a 'high priori road' to this goal brings us back to the empirical method ['high priori' is a joking form of *a priori*; it was coined by the poet Alexander Pope]. Rather than a clear and universally valid principle, the best we get is a vague and general rule, based on considerations that shouldn't be overlooked but can't be evaluated except by careful observation and comparison of individual experience. Any uncertainty in these processes then carries through to all our reasonings about happiness. I don't want to exaggerate •these uncertainties, feeling that we should all continue to seek happiness for ourselves and for others, however much we have to grope for it in the dark; but there is nothing gained by underrating •them, and it is idle to argue as if •they did not exist.

[1] The quotations are from Spencer's *Social Statics* chapter 4. In the passage from which I have quoted he is not writing from the point of view of egoistic hedonism.

BOOK III: Intuitionism

Chapter 1: Intuitionism

1. The effort in Book II to examine closely the system of egoistic hedonism may well have given you a certain aversion to that principle and method, even if you (like myself) find it hard not to admit the 'authority' of self-love, or the 'rationality' of seeking one's own happiness. In considering 'enlightened self-interest' as supplying a *prima facie* tenable principle for systematically guiding conduct, I have kept my aversion out of sight, being anxious to learn with scientific impartiality the results to which this principle logically leads.

We recoil from egoism when we see its occasional practical conflict with common notions of duty [see Glossary]; but our sympathetic and social nature is more deeply offended when we discover—through a careful empirical examination of egoism—that the common precepts of duty that we are trained to regard as sacred must be regarded by the egoist as mere rules that it is usually reasonable to follow but under special circumstances must be decisively ignored and broken. Furthermore, we look to morality for clear and decisive precepts or counsels; and rules for seeking the individual's greatest happiness can't be either clear or decisive. The calculus of egoistic hedonism seems to offer nothing but a dubious guide to an ignoble end! Butler admits (in the passage quoted on page 55) that the claims of self-love have *theoretical* priority over those of conscience, but the dictates of conscience are more certain than those of self-love, which is why Butler gives them *practical* supremacy.[1] A man knows for sure, he says, what he ought to do; but he doesn't know for sure what will make him happy.

This seems to me to represent fairly mankind's common moral sense, in our time no less than in Butler's. The moral judgments that men habitually express in ordinary discourse mostly imply that it's not usually hard for an ordinary man to •know what his duty is, though seductive impulses may make it hard for him to •do it. And such maxims as that

- •duty should be performed, come what may,
- •truth should be spoken without regard to consequences,
- •justice should be done 'though the sky should fall',

imply that we can see clearly that certain kinds of actions are right and reasonable in themselves, apart from their consequences; or rather with a consideration from which *some* consequences—admitted to be possibly good or bad—are definitely excluded.[2] And most of the writers who have maintained the existence of moral intuitions have claimed

[1] It may seem, he admits, that 'since one's own happiness is an obvious obligation', whenever virtuous action seems not to be conducive to the agent's happiness he would 'be under two contrary obligations—i.e. under none. . . But the obligation on the side of ·self·-interest really doesn't remain; because the natural authority of the principle of reflection is. . . the most certain and best known obligation, whereas the contrary obligation can't seem more than probable. No man can be certain in any circumstances that vice is in his interests in •the present world; much less can he be certain that it is in his interests in •another world. So the certain obligation would entirely outrank and destroy the uncertain one.' (Preface to Butler's Sermons.)

[2] I noted in I/8.1 [page 44] that in the common notion of an act we include a certain portion of the whole series of changes partly caused by the volition that initiated the so-called act.

that the human mind can do this; which is why I think I am justified in treating this claim as characteristic of what I call the 'intuitional' method.

But there's a wider sense in which either egoistic or universalistic hedonism might be legitimately be called 'intuitional', if either system presents *Happiness is the only rational ultimate end of action* as a first principle that can't be known in any way except intuitively. I shall return to this ·wider· meaning in chapters 13–14, where I'll discuss more fully the intuitive character of these hedonistic principles. But adopting this wider meaning wouldn't lead us to a distinct ethical method, so I have thought it best in my detailed discussion of intuitionism in chapters 1–11 to confine myself as far as possible to moral 'intuition' taken in the narrower sense that I have defined.

2. Someone might object as follows:

> Your definition of *intuitionism* omits its most fundamental characteristic: the intuitionist, properly so-called, doesn't judge actions by any external standard as the utilitarian does; he sees true morality as concerned not with •outward actions as such but with •the state of mind in which acts are done—i.e. with 'intentions' and 'motives'.[1]

This objection is partly due to a misunderstanding. Moralists of all schools would agree that the moral judgments we pass on actions relate primarily to intentional actions regarded as intentional. In other words, what we judge to be 'wrong' in the strictest ethical sense is not any of the actual effects of the muscular movements caused by the agent's volition, but the effects that he foresaw in willing the act. . . .[2] So when I speak of 'acts', take me to mean—unless I say otherwise—acts presumed to be intentional and judged as such. I don't think there needs to be any dispute about this.

The case of motives is different and requires careful discussion. In ordinary language the distinction between 'motive' and 'intention' isn't very precise: we apply the term 'motive' either to •consequences of an act that the agent foresaw and desired or •to the agent's desire for them; and when we speak of the 'intention' of an act we are usually thinking of desired consequences. But for purposes of exact moral or jural [see Glossary] discussion it's best to include under the term 'intention' all the consequences of an act that are foreseen as certain or probable: you'll agree that we can't evade responsibility for any foreseen bad consequences of our acts by the plea that we didn't *want* them for themselves or as means to some further end;[3] such undesired accompaniments of the desired results of our volitions are clearly chosen or willed by us. So the •intention of an act can be judged to be wrong though the •motive is recognised as good; as when a man tells a lie to save a parent's or a benefactor's life. Such judgments are made all the time in ordinary moral discourse. But this may be said:

[1] Some would add 'character' and 'disposition'. But characters and disposition can't even be *conceived* except in terms of the volitions and feelings that manifest them, so they can't be primary ·or basic· objects of intuitive moral judgments. See chapter 2.2 .

[2] No doubt we hold a man responsible for unintended bad consequences of his acts or omissions, when they are ones that he might with ordinary care have foreseen; still, as I said on page 27, if we think about it we attach moral blame to careless acts or omissions only indirectly, and only if the carelessness results from some previous willful neglect of duty.

[3] Think carefully about common usage and you'll see that it fits this definition. Suppose a nihilist blows up a railway train containing an emperor and other people; it would be regarded as correct to say simply '*His intention was to kill the emperor*'; but it would be thought absurd to say '*He did not intend to kill the other people*', even if he had no desire to kill them, and regarded their death as a regrettable by-product of the carrying out of his revolutionary plans.

> 'An act can't be right, even when the intention is what duty would prescribe, if it is done from a bad motive. To take an example of Bentham's, a man who prosecutes from malice a person whom he believes to be guilty doesn't really act rightly: it may be his duty to prosecute, but he ought not to do it from malice.'

No doubt it is our duty to get rid of bad motives if we can; so that a man's intention can't be wholly right unless it includes the repression, so far as possible, of a motive known to be bad. But no-one will contend that we can always suppress entirely a strong emotion; and such suppression will be especially difficult if we are to do the act to which the wrong impulse prompts. And if the act is clearly a duty that no-one else can perform as well, it would be absurd to say that we ought to omit it because we can't entirely erase an objectionable motive. It is sometimes said that even if in doing our duty we can't exclude a bad motive altogether from our minds, it is still possible to refuse to act from it—·i.e. possible to perform the action without giving the bad motive any role in our doing so·. But this is possible only if the details of the action to which a right motive would prompt differ to some extent from those to which a wrong motive would prompt. No doubt this is often the case. In Bentham's example, a malevolent prosecutor may be prompted to cause his enemy needless pain by well-aimed insults; and obviously he can do his duty without doing *that*. But when precisely the same action is prompted by two motives that are both present in my consciousness, I'm not aware of having any power to cause this action to come from one of the two to the exclusion of the other. . . .[1]

From all this I conclude **(1)** that while we commonly judge many actions to be made *better* or *worse* by the presence or absence of certain motives, our judgments of *right* and *wrong* strictly speaking relate to intentions, as distinguished from motives;[2] and **(2)** that while intentions affecting the agent's own feelings and character are morally prescribed no less than intentions to produce certain external effects, common moral understanding holds that the *main* prescriptions of duty are addressed to external actions. How far this is true will become clearer in due course.

·**One extreme:**· Some influential moralists have maintained •that the moral value of our conduct depends on the extent to which we are actuated by the one motive that they regard as truly moral, namely *the desire to do what is right because it is right*, doing one's duty for duty's sake,

[1] A further source of confusion between 'intention' and 'motive' arises from the different points of view from which either may be judged. If an act is one of a series that the agent intends to do for the achievement of a certain end, we may have •one moral judgment on the intention of the particular act and •a different one on the intention of the series as a whole. Either point of view is legitimate, and often both are required, for we commonly recognise that of the series of acts that a man does to achieve some (e.g.) ambitious goal some are right or allowable while others are wrong; while the general intention to achieve the goal by wrong means if necessary—

> Get place and wealth, if possible with grace;
> If not, by any means get wealth and place [quoted from Alexander Pope]

—is clearly a wrong intention. Also, in judging a motive to be good or bad, we may consider it •simply in itself or •in connection with other balancing and controlling motives that are present or that ought to be present but aren't. We don't usually think that the desire for wealth or rank is bad in itself, but we think it bad as the sole motive of a statesman's public career. It's easy to see that either of these different distinctions is apt to blend with and confuse the simple distinction between intention and motive.

[2] The view that moral judgments relate primarily or most properly to motives will be more fully discussed in chapter 12.

being virtuous for virtue's sake;[1] and •that a perfectly good act must be done entirely from this motive. But it's hard to combine this view—which I'll label as 'Stoical'—with the belief that most modern orthodox moralists have maintained, that it is always a man's true interest to act virtuously. I don't mean that someone who holds this belief must be an egoist; but if he thinks that his own interests will be promoted by the act that he is undertaking, it seems impossible for him to keep a concern for his own interests out of the motives that are driving him. So if we hold

- •that this self-regard impairs the moral value of an act that is otherwise virtuous, and also
- •that virtue is always conducive to the virtuous agent's interest,

we're forced to conclude that

- •knowledge of the true relation between virtue and happiness is an insuperable obstacle to the achievement of moral perfection.

I can't accept this paradox; and in later chapters I'll try to show that the Stoical view of moral goodness doesn't stand up to a comprehensive survey of common moral judgments, because some acts seem to be even *more* strikingly virtuous when performed from some motive *other than* the love of virtue as such. For now I merely remark that the Stoical doctrine contradicts ·**the other extreme**, namely· the view that

- **(i)** the universal or normal motives of human action are either •particular desires for pleasure or aversions to pain for the agent himself, or •the agent's concern for his happiness on the whole ('self-love');

and that it also conflicts with the less extreme doctrine that

- **(ii)** duties can to some extent be properly done from such self-regarding motives;

to which I add that **(i)** or **(ii)** has frequently been held by writers who have explicitly adopted an intuitional method of ethics. We find Locke, for instance, stating without reserve or qualification that 'good and evil are nothing but pleasure and pain, or that which procures pleasure or pain to us'; so that 'it would be utterly pointless to set a rule for the free actions of man without annexing it to some reward or punishment to determine his will' (*Essay on Human Understanding* II/28.5,6). Yet he also, just as emphatically, expresses the conviction that 'from self-evident propositions, by valid inferences as incontestable as those in mathematics, the measures of right and wrong could be derived', 'so that 'morality might be placed among the sciences capable of demonstration' (*Essay* IV/3.18). The combination of these two doctrines gives us the view that moral rules are essentially laws of God that men are impelled to obey—solely or mainly—from fear or hope of divine punishments or rewards; and a view like this seems to be widely accepted by plain men without very refined moral sensibilities.

·**Between the extremes:**· For other examples of thinkers who •recognise in human nature a disinterested regard for duty or virtue as such, but still •think that self-love is a proper and legitimate motive to right conduct, let us look at Butler and his followers. Butler regards 'reasonable self-love' as not merely a normal motive to human action, but as being a 'chief or superior principle [see Glossary] in the nature of man' as much as conscience is, so that an action 'becomes unsuitable' to this nature if the principle of

[1] Many religious folk would probably say that *obedience or love to God* is the highest motive. But most of them would also say that obedience and love are due to God as a moral being, one who is infinitely wise and good, and not otherwise; and in that case •these religious motives seem to be virtually identical with •regard for duty and love of virtue, though complicated by the addition of emotions (·obedience, love·), belonging to relations between persons.

self-love is violated. Accordingly the aim of his teaching is not to •induce men to choose duty rather than interest, but to •convince them that there's no need to choose: self-love and conscience lead 'to one and the same course of life'.

This intermediate doctrine strikes me as more in harmony with the common sense of mankind than either of the extreme views I have contrasted. But each of the three positions is consistent with the basic assumptions of the intuitional method. Even those who hold that

> human beings can't reasonably be expected to conform to moral rules from any motive except what comes from the sanctions God has attached to them

usually think of God as supreme Reason, whose laws must be essentially reasonable; and if such laws are knowable by the 'light of nature'—so that morality may (as Locke says) be classified as a demonstrative science—the method of settling what they are will still be intuitional, and won't lose that status because the method is combined with the belief that God will reward the observance of the laws and punish their violation. As for those who hold that regard for duty as duty is essential to acting rightly would generally admit that •acting rightly is not adequately *defined* as acting from a pure desire to act rightly; that although a man who sincerely desires and intends to act rightly does in a certain sense completely fulfill duty, he may have a wrong judgment about the particulars of what his duty is, so that in another sense he acts wrongly. From this it follows that even if the desire or resolution to fulfill duty as such is essential to right action, two kinds of rightness must be recognised:

(i) an act is 'formally'[1] right if the agent is moved by pure desire to fulfill duty;

(ii) an act is 'materially' right, if the agent intends the right particular effects.

So there's no reason why the same method for determining material rightness shouldn't be adopted by thinkers who disagree widely about formal rightness; and obviously the work of the systematic moralist is mainly concerned with material rightness.

3. Formal rightness as explained here involves •a desire or choice of the act as right, and also •a belief that it is right. But you could have the belief without having the motive (though not vice versa); and there's more agreement among intuitional moralists about the moral indispensability of the belief than about the moral indispensability of the motive. I think they would all agree that no act can be absolutely right. . . .if the agent believes it is wrong.[2] Such an act could be called 'subjectively wrong' though 'objectively right'. A question arises. In a particular case, which of these is better?

- •The man does what he mistakenly believes to be his duty.
- •The man does what really is his duty except that he doesn't think so.

This question is rather subtle and perplexing to common sense, so it's as well to note that it can't have much practical application. It can't arise for anyone with respect to how he is going to act; we can only raise it in relation to someone else whom we might influence. If someone is poised to do something that we think wrong while he thinks it right, and we can't alter his belief but can bring other motives to bear

[1] I don't usually employ the obscure and ambiguous form/matter antithesis when I write philosophy. In the present case we can interpret 'formal rightness' as denoting both a *universal* and *essential* and also a *subjective* or *internal* condition of the rightness of actions.

[2] Not necessarily that the belief that it is right should be actually present in the agent's mind; it might be completely right although the agent never actually raised the question of its rightness or wrongness. See page 106.

on him that may outweigh his sense of duty, we must decide whether to do that. Ought we to tempt him to act against his own convictions by bringing about what *we* believe to be objectively right? The moral sense of mankind would say No, regarding the subjective rightness of an action as more important than its objective rightness—except ·in special cases where· the evil of the act prompted by a mistaken sense of duty appeared to be very grave.[1] But however essential it may be that a moral agent should do what he believes to be right, *this* subjective condition of right conduct is too simple to be the basis for any theories; so our investigation here must relate mainly to 'objective' rightness.

But one practical rule of some value can be obtained by reflecting on the general notion of rightness, as commonly conceived. In I/3.3 I tried to make this notion clearer by saying that 'what I judge to be right must—unless I am in error—be judged to be so by all rational beings who judge truly of the matter'. This doesn't imply that what is judged to be right for one man must necessarily be judged so for another; 'objective' rightness may vary from A to B just as objective facts vary. But there's a difference between our conceptions of ethical and physical objectivity, concerning how they relate to variations for which we can discover no rational explanation. Experience compels us to admit such variations in physical facts, but we commonly refuse to admit them in moral facts. Physical facts involve an accidental or arbitrary element that we just have to accept. . . . Why does this region of space contain more matter than that? Physical science's only answer brings in laws of change and facts about earlier positions of portions of matter, facts that equally cry out for explanation; and however far back we take our explanations, the fact at which we stop seems as arbitrary as the one we first asked about. But it's generally agreed that we can't admit a similar unexplained variation concerning right and wrong. We can't judge an action to be right for A and wrong for B unless we can find some difference between the two agents—in themselves or in their circumstances—that we can regard as a *reason* for the difference in their duties. So if I judge any action to be right for myself, I implicitly judge it to be right for anyone else whose nature and circumstances don't differ significantly from mine. Now, by making this latter judgment explicit we can protect ourselves against the danger of too easily thinking that we *ought* to do what we very much *want* to do. 'Do I think that anyone like me in my circumstances ought to do A?'—the answer may clearly be No, and that may disperse the false appearance of rightness that my strong desire has given to doing A. . . . Indeed this test of the rightness of our volitions is so generally effective that Kant seems to have held that all particular rules of duty can be deduced from

[1] The decision would usually be reached by weighing bad effects on the agent's character against bad consequences of a different kind. In extreme cases common sense would decide against the agent's character. A statesman crushes a dangerous rebellion by working on the fear or greed of a leading rebel who has been rebelling on conscientious grounds—most of us would approve of this. See IV/3.3 .

[1] [The key to this footnote is high on the next page.] *Groundwork for the Metaphysic of Morals*, page 24 of the version at www.earlymoderntexts.com. Kant says: 'There is only one categorical imperative, and this is it: Act only on that maxim through which you can at the same time will that it should become a universal law of nature. Now if all imperatives of duty can be derived from this one imperative as a principle, we'll at least be able to show what we understand by *duty*, what the concept means.' He applies the principle to four cases, selected as representative of 'the many actual duties'; and continues: 'If we attend to what happens in us when we act against duty, we find that we don't (because we *can't*) actually will that our maxim should become a universal law.' And he sums up thus: 'I have made clear—and ready for *every practical application*—the content that the categorical imperative must have if it is to contain the principle of all duty, if there is such a thing as duty.'

the one fundamental rule *Act only on that maxim through which you can at the same time will that it should become a universal law of nature.*[1] But this strikes me as an error analogous to that of supposing that formal logic supplies a complete criterion of truth. I agree that a volition that doesn't stand this test is to be condemned; but a volition that passes the test may still be wrong. Almost everyone who acts conscientiously could sincerely will the maxims on which he acts to be universally adopted; yet we find such people conscientiously disagreeing about what each ought to do in a given set of circumstances. If they all act rightly (in the objective sense) because their maxims all conform to Kant's rule, that obliterates the distinction between subjective and objective rightness; it implies that whatever anyone thinks to be right *is* so, unless he is in error about the non-moral facts of the case. That flagrantly conflicts with common sense. It would make it futile to ·try to· construct a scientific code of morality, because the purpose of such a code is to supply a standard for correcting men's divergent opinions.

So we can conclude that the moral judgments that the intuitional method tries to systematise are primarily intuitions of the rightness or goodness (or wrongness or badness) of particular kinds of external effects of human volition, presumed to be intended by the agent but considered independently of his view about the rightness or wrongness of his intention; though the quality of •motives, as distinct from •intentions, must also be taken into account.

4. You may want to ask: 'Has it been legitimate for you to take it for granted that there are such intuitions?' No doubt there are people who deny that reflection shows them any such phenomenon in their conscious experience as the •judgment or •apparent perception that an act is in itself right or good—except in the sense of being the right means to some chosen end. But such denials are commonly recognised as paradoxical and opposed to the common experience of civilised men—as long as we are careful to distinguish **(a)** the psychological question about the **existence** of such moral judgments from **(b)** the ethical question about their **validity**, and from **(c)** the psychogenetic [see Glossary] question as to their **origin**. Of these, **(a)** and **(b)** are sometimes run together because of an ambiguity in the term 'intuition ', which has sometimes been understood to mean a *true* judgment. Let me be clear about this: by calling an affirmation about the rightness or wrongness of an action 'intuitive' I am *not* prejudging the question of its ultimate validity. . . . All I mean is that its truth is apparently known immediately and not as the result of reasoning. I admit that any such intuition may turn out to contain an error that we may be able to correct by reflection and comparison, just as many apparent visual perceptions turn out to be partially illusory and misleading. Indeed, you'll see later that I hold this to be to an important truth about moral 'intuitions' commonly so called.

Having separated **(a)** the existence question from **(b)** the validity question, we can see that obviously **(a)** can be decided for each person only by introspection. But don't think this:

> Deciding **(a)** is a simple matter, because introspection is always infallible.

On the contrary, I find that men are often liable to confuse moral intuitions with other mental states or acts that are essentially different from them—

- •blind impulses to certain kinds of action,
- •vague preferences for such actions,
- •conclusions from fast semi-conscious inferences,
- •current opinions that familiarity has given an illusory air of self-evidentness.

But errors of this kind can only be cured by more careful introspection, aided by consulting with others, and perhaps by looking into the antecedents of the apparent intuition,

which may suggest possible sources of error. Still, the question of whether **(a)** a certain judgment presents itself to the reflective mind as intuitively known can't be decided by any inquiry into **(c)** its antecedents or causes. See I/3 at page 16.

But it's still possible to hold that an inquiry into **(c)** the origin of moral intuitions must be decisive in determining **(b)** their validity. And in fact intuitionists and their opponents have often assumed that if •our moral faculty can he shown to be 'derived' or 'developed' out of other pre-existent elements of mind or consciousness, that's a reason for distrusting •it; whereas if it can be shown to have existed in the human mind from the outset, that establishes its trustworthiness. Neither assumption has any foundation that I can see. . . . **On the one hand:** I'm sure that each of our cognitive faculties—i.e. the human mind as a whole—has been derived through a gradual process of physical change from some lower life in which cognition properly so-called had no place. So the distinction between 'original' and 'derived' comes down to that between 'earlier' and 'later'; and the fact that the moral faculty appears later in the process of evolution than other faculties can't be regarded as an argument against the validity of moral intuition! The discovery of the causes of certain apparently self-evident judgments can't be a reason for distrusting them. 'Well, those who affirm the truth of such judgments ought to show that their causes have some power to make them true'—I don't accept that either. Indeed, if that is where the onus of proof lies, philosophical certainty would be impossible: the premises of the required demonstration must consist of caused beliefs, which just because they are caused will equally stand in need of being proved true, and so on ad infinitum. The only escape would be •to find among the premises of our reasonings certain apparently self-evident judgments that don't have causes, and •to argue that because they don't have causes they should be accepted as valid without proof—an extravagant paradox! And if it's accepted that all beliefs are effects of prior causes, this characteristic clearly can't on its own invalidate any of them.

So I hold that the onus of proof goes the other way: those who dispute the validity of •moral or other intuitions because of their derivation should show not merely that •they are the effects of certain causes but also that those causes are likely to produce invalid beliefs. Now. I don't think it is possible to prove by any theory of the derivation of the moral faculty that the basic ethical conceptions 'right' (or 'what ought to be done') and 'good' (or 'what it is reasonable to desire and seek') are invalid, and that consequently all propositions of the form 'x is right' or 'x is good' are untrustworthy. Why not? Because such ethical propositions can't be inconsistent with any physical or psychological propositions, since their subject-matter is fundamentally different from anything that physical science or psychology deals with. The only way to show that they involve error is to show that they contradict *each other*; and such a demonstration couldn't validly lead us to the conclusion that they are *all* false. Perhaps, though, we can prove that *some* ethical beliefs have been caused in a way that makes them likely to be wholly or partly false; and later we'll have to consider whether any of the ethical intuitions that we are disposed to accept as valid are open to attack on such psychogenetic grounds. My present point is just that no general demonstration of the derived status of our moral faculty can give an adequate reason for distrusting it.

On the other hand: If we are led to distrust our moral faculty on other grounds. . . ., it seems to me equally clear that our confidence in our moral judgments can't properly be re-established by a demonstration that they are 'original'.

I see no reason to believe that the 'original' element of our moral cognition can be discovered; but if it could, I see no reason to hold that it would be especially free from error.

5. Then how *can* we eliminate error from our moral intuitions? In chapter I/8 I suggested that to settle the doubts arising from the uncertainties and discrepancies in our judgments on particular cases, reflective people naturally appeal to general rules or formulae; and it's those general formulae that intuitional moralists commonly regard as ultimately certain and valid. There are obvious sources of error in our judgments about concrete duty ·in particular cases· that seem to be absent when we consider the abstract notions of different *kinds* of conduct. That is because in any particular case the complexity of the facts increases the difficulty of judging, and our interests and sympathies are liable to cloud our moral discernment. And most of us feel the need for such formulae not only to •correct, but also to •supplement, our intuitions about particular duties. Only exceptionally confident people think that they always see clearly what ought to be done in any situation they find themselves in. The rest of us, sure as we are about what is right or wrong in ordinary matters of conduct, quite often meet with cases where our unreasoned judgment fails us; and where we can't decide the moral issue in question without appealing to some general formula—just as we couldn't decide a disputed legal claim without reference to the positive law [see Glossary] that deals with the matter.

And such ·general· formulae are easy enough to find. A little reflection and observation of men's moral discourse will enable us to make a collection of general rules that

- •would be generally accepted by moral persons of our own times and our own civilisation, and
- •would cover fairly completely the whole of human conduct.

Such a collection, regarded as a code imposed on an individual by the public opinion of the community to which he belongs, I have called *the positive morality of the community*; but when it is warranted as a body of moral truth by the consensus of mankind—or at least of the portion of mankind that combines adequate intellectual enlightenment with a serious concern for morality—it is more significantly termed *the morality of common sense*.

But when we try to apply these currently accepted principles, we find that the notions composing them are often unclear and imprecise. We all agree in recognising justice and veracity as important virtues; and probably we'll all accept the general maxim that 'we ought to give every man his own', but when we ask whether

- •primogeniture or
- •the disendowment of corporations [= depriving churches of their wealth] or
- •the fixing of the value of services by competition

is just, we don't get clear and unhesitating decisions from that or any other current maxim. Again, we all agree that 'we ought to speak the truth', but when there's a question about whether and to what extent false statements are permissible

- •in speeches of advocates,
- •in religious ceremonials,
- •when speaking to enemies or robbers, or
- •in defence of lawful secrets,

we again get no help from that or any other general maxim. And yet such particular questions are just the ones that we naturally expect the moralist to answer for us. As Aristotle says, we study ethics for the sake of practice; and in practice we are concerned with particulars.

So it seems that if the formulae of intuitive morality are really to serve as scientific axioms, and to be available in clear and compelling demonstrations, they must first be

raised—by an effort of reflection that ordinary folk won't make—to a higher degree of precision than they have in the common thought and discourse of mankind. We have in fact to tackle the task launched by Socrates, of defining satisfactorily the general notions of duty and virtue that we all use in approving or disapproving of conduct. This is the task I'll be engaged on in the next nine chapters. *Please* bear in mind that I shan't be trying to prove or disprove intuitionism, but merely to get as explicit, exact, and coherent a statement as possible of its basic rules, doing this by reflection on the common morality to which appeal is so often made in moral disputes—the one that you and I share.

Chapter 2: Virtue and duty

1. Before trying to define particular virtues or kinds of duty, we should look further into the notions of *duty* and *virtue* in general, and into the relations between them. . . . Until now I have taken *duty* to be roughly equivalent to *right conduct*; but I pointed out that 'duty'—like 'ought' and 'moral obligation'—implies at least the *potential* presence of motives going the other way, so that it isn't applicable to beings who don't have such conflict of motives. Thus God is not conceived as performing duties, though he is conceived as realising [see Glossary] justice and other kinds of rightness in action. And we don't commonly label as 'duties' right actions of our own that we are strongly impelled to by non-moral inclinations; we don't usually say that it is a 'duty' to eat and drink enough, though we might say this to invalids who have lost their appetite. So we'll get closer to ordinary usage if we defined *duties* in a way that brings in 'the need for a moral impulse'. But the line drawn in this paragraph is vague and shifting, and it won't be necessary to draw attention to it in the detailed discussion of duties. . . .

This may be said:

> You have overlooked another element in the meaning of 'duty'—one that its derivation and that of the equivalent term 'obligation' plainly indicates—namely that it is 'due' or owed to someone.

I agree that 'duty' comes from 'due' = 'owed', but this is a case where etymology doesn't governs ordinary usage. Most people would recognise that duties owed to persons. . . .are only one species, and that some duties—e.g. truth-speaking—fall outside that species. No doubt any duty can be seen as relative to whoever is immediately affected by it, as when truth-speaking causes a physically injurious shock to the person spoken to, but we don't even in these cases speak of the speaker's 'duty to' the other person. You *could* say that truth-speaking is ultimately good for—and therefore 'due to'—the community or to humanity at large; but that isn't how it is thought of in the intuitional view that 'truth should be spoken regardless of consequences'. Religious folk may think that the performance of duties is 'owed to' God as the author of the moral law. I wouldn't deny that our common

conception of duty involves an implied relation of •an individual will to a •perfectly rational universal will; but I'm not convinced that this implication is necessary—·i.e. that this is an aspect of the *concept* of duty·—and I shan't discuss it because that would lead to metaphysical controversies that I want to avoid. In what follows, therefore, I'm going to set aside relation of duty •generally to a divine will, and also •the particular 'duties to God' that intuitionists have often picked out and classified. If we regard the basic moral rules that we can know by moral intuition as ones that it is rational for all men to obey, then we see them as rules that a supreme Reason *would* impose, and it shouldn't make any difference whether we think that a supreme Reason *did* impose them. [That sentence is a rather free rendering of the complex thing that Sidgwick wrote.] So I shan't treat 'duty' as implying a relation either to a universal Ruler or to the individuals affected by the conduct in question, but will use it as equivalent to 'right conduct', while focusing on actions and inactions for which a moral impulse is thought to be required.

The notion of *virtue* is more complex and difficult, and needs to be discussed from several angles. Start by noticing that some particular virtues (such as generosity) can be realised in acts that are objectively (though not subjectively) wrong, through lack of insight into their consequences; and some (such as courage) can be exhibited in acts that the agent knows to be wrong. We have a quasi-moral admiration for such acts; but we wouldn't call the courageous act virtuous, and if we were speaking strictly we wouldn't call the generous act virtuous either. So I won't be significantly deviating from ordinary usage when, from now on, I apply 'virtue' only to qualities exhibited in right conduct.

How far are the spheres of •duty and •virtue co-extensive? To a large extent they undoubtedly are so in the ordinary use of the terms, but not altogether, because each term in its common use seems to include something excluded from the other. We would hardly say that it was *virtuous* under ordinary circumstances to

- •pay one's debts,
- •give one's children a decent education, or
- •keep one's aged parents from starving;

these being duties that most men perform and only bad men neglect. And there are acts of high and noble virtue that we commonly regard as going beyond the agent's *duty*, because although we praise their performance we don't condemn their non-performance. But now a problem seems to arise: we wouldn't deny that it is in some sense a man's strict duty to do whatever action he judges most excellent, so far as it is in his power.

But can we say that it is as much in a man's power to realise virtue as it is to fulfill duty?[1] To some extent we would say this. No quality is ever called a 'virtue' unless it is thought to be something that any ordinary person could *choose* to exhibit when an opportunity arises. In fact virtues are commonly distinguished from other excellences of behaviour by their voluntariness: an excellence that we think isn't significantly under the immediate command of the will is called a gift, a grace, or a talent, but not properly a virtue. Writers who obliterate this line—as Hume does in *Inquiry Concerning the Principles of Morals*, Appendix 4—are obviously diverging from common sense. But it's plainly wrong to maintain that anyone can at any time realise virtue in the highest form or degree. No-one would say that any ordinary man can *at will* exhibit the highest degree

[1] In I/5.3 I have explained the sense in which determinists as well as libertarians hold that it is in a man's power to do his duty.

of courage—in the sense in which courage is a virtue—when occasion arises. So it seems that we can distinguish a margin of virtuous conduct which may be beyond the power of any individual and therefore beyond his duty.

Can we then say that virtuous conduct, so far as it is in a man's power, coincides completely with his duty? Well, we would agree that a truly moral man can't say to himself 'This is the best thing on the whole for me to do, and I could do it, but it's not my duty to do it'; this would strike common sense as an immoral paradox.[1] And yet there seem to be actions and inactions that we praise as virtuous without imposing them as duties on all who are able to do them—e.g. a rich man lives very plainly and devotes his income to works of public beneficence.

Perhaps we could harmonise these inconsistent views by distinguishing the question 'What ought a man to do or not-do?' from the question 'What ought other men to blame him for not-doing or doing?', and recognising that the standards for answering them are different. The double standard can be partly explained by differences in our knowledge relating to the two. . . . I can easily assure myself that I ought to subscribe to a given hospital; but I can't judge whether my neighbour ought to subscribe–·and so I can't blame him for not subscribing·—because I don't know the details of his income and of claims that he is bound to satisfy. But that isn't the whole explanation for the double standard: there are plenty of cases where we don't blame others for not doing x although we are sure that we in their place would have thought it our duty to do x. In such cases the line seems drawn by •a sense of what counts as ordinary behaviour in such a case, and by •a belief about the practical effects of praising and blaming: we think it best for moral progress if we praise acts that are above the level of ordinary practice and restrict precisely targeted blame to acts that fall clearly below this standard. This standard has to be vague, and to vary along with variations in the community's average level of morality. . . . So it isn't convenient to use it in drawing a theoretical line between virtue and duty; and I prefer to employ the terms so that virtuous conduct includes not only good actions that would commonly be thought to go beyond duty but also the less conspicuously virtuous actions in which the agent does his duty.

2. I have been considering 'virtuous' as applied to conduct. But this general word and the names meaning particular virtues—'just', 'liberal', 'brave', etc.—are applied to persons as well as to their acts; and the question arises as to which application is most appropriate or basic. Reflection shows that we don't think of these attributes as belonging to acts considered apart from their agents; so •virtue is primarily a quality of the permanent soul or mind rather of the transient acts and feelings in which •it is manifested. (It is widely held to be a possession worth aiming at for its own sake—to be indeed a part of the perfection of man that some regard as the sole ultimate good. I shall consider this view in chapter 14.) But although we think of virtues (and other habits and dispositions) as comparatively •permanent attributes of the mind, our only way of forming definite notions of them is by conceiving the •transient events in which they are manifested. What events? Obviously *voluntary actions considered as intentional* or, more briefly, *volitions*. Many moralists, perhaps most, would give this as a complete answer. If they don't affirm with Kant that •a good will is the

[1] A moral person who sincerely wants to do his duty might say that 'paradoxical' thing either **(i)** half-ironically—a nod to a customary standard of virtuous conduct that the speaker doesn't really adopt as valid—or **(ii)** loosely, meaning that the conduct in question would be best if the speaker were differently constituted.

only absolute and unconditional good, they agree with Butler that •'the object of the moral faculty is *actions*, taking that word to cover active or practical principles—the principles from which a man would act if circumstances enabled him to'. And if it is urged that that isn't the whole story for the Christian conception of the virtue of charity, the 'love of our neighbour', those moralists will join Kant in saying that this 'love' is not the *affectionate emotion* that goes by that name, but merely the *resolve to benefit* others, which alone has 'true moral worth'.

But I don't think that the common sense of mankind really does exclude every emotional element from the conception of *virtue*. In our common moral judgments certain kinds of virtuous actions are held to be, at least, adorned and improved by the presence of certain emotions in the agent, though the element of volition is doubtless more important and indispensable. Thus the highest form of the virtue of chastity or purity includes more than a mere settled resolve to abstain from unlawful lust; it includes also some feeling of repugnance to impurity. And we recognise that benefits arising from affection and lovingly given are more acceptable to the recipients than benefits given without affection, in the taste of which there's something harsh and dry; and so the affection, if it is practical and steady, seems to be in a certain way *more* excellent than the mere beneficent disposition of the will. In the case of *gratitude*, even Kant's rigidity seems to relax and to include an element of emotion in the virtue; and various other notions such as loyalty and patriotism can't plausibly be •stripped bare of all emotional elements or •denied to be virtues.

We are now in a position to answer a question raised in chapter 1: 'Does common sense hold that an act is virtuous in proportion as it was motivated by a concern for duty or virtue?' The answer is No. A •courageous or •loyal or •patriotic action isn't made less praiseworthy by the fact that its main motive was natural affection rather than love of virtue as such. Quite often we attribute virtue to actions where there was no conscious thought of duty or virtue; think of a case of a heroic act of courage in saving a fellow-creature from death, done from an impulse of spontaneous sympathy. And when we praise a man as 'genuinely humble' we don't imply that he is conscious of exhibiting a virtue by being humble!

With many important virtues we ordinarily see the person as virtuous without giving any thought to where his action comes from—whether from some emotional impulse or from a rational choice of duty because it is duty—but only crediting him with having a settled resolve to aim at external states of affairs of a certain kind. We call a man *veracious* if we see him as consistently trying to produce true beliefs in others, whatever his motive may be for doing this—a regard for virtue, a sense that lying is low, a belief that truth-speaking is the best policy, or a sympathetic aversion to the troubles that misleading statements cause to other people. It's not that we regard these motives as of equal moral value; but the presence or absence of any of them is not implied when we attribute to someone the virtue of veracity. Similarly we attribute justice to a man who has a settled habit of weighing claims and fulfilling them in the ratio of their importance; we credit him with good faith if he has a settled habit of strictly keeping engagements; and so on. And when we clearly do take motives into account in judging how virtuous an action is, we are often thinking not of •what produced the action but of •what temptations had to be overcome for it to be performed: we see more virtue in just or veracious conduct when the agent had strong temptations to be unjust or unveracious. . . .

Admittedly common sense seems to be perplexed about how •virtue relates to •the moral effort required for resisting unvirtuous impulses. On the one hand:

> We would generally accept that virtue is especially exhibited in a successful conflict with natural inclination; and perhaps even (more extreme) that there is no virtue in doing what one likes.

On the other hand:

> We would surely agree with Aristotle that virtue is imperfect if the agent can't perform the virtuous action without a conflict of impulses. What makes it hard for us to do what is best is a wrongly directed natural impulse, and it seems absurd to say that the more we cure ourselves of this wrong direction the *less* virtuous we grow.

Perhaps we should recognise two distinct elements in our common idea of virtue: **(a)** the most perfect ideal of moral excellence that we can conceive for human beings, and **(b)** the effort of imperfect men to achieve this ideal. Then as a man comes to like some particular kind of good conduct and to do it without moral effort, we'll say not •that his conduct becomes less virtuous but •that it comes to be more in conformity with a true moral ideal. We'll recognise that in this department of his life he has less room to exhibit the other kind of virtue—the one shown in resistance to seductive impulses and the will's energetic striving to get nearer to ideal perfection.

Up to here I have been discussing the roles of •emotions and •volitions in virtuous acts, and haven't explicitly mentioned their •intellectual conditions of virtuous acts. When we call an act 'virtuous' we imply that it was voluntary, i.e. arose from a volition, and we imply that the volition was accompanied with a •thought about the particular goal of the action. But we don't imply that the agent also had the •thought that what he was doing was right or good. I don't think that common sense holds *that* to be essential for an act to count as virtuous; some kinds of virtuous acts can be done without any deliberation and with no moral judgment passed on them by the agent. . . . But an action can't count as virtuous if the agent has even a vague thought of its being bad. As I have already said, it's more doubtful how far common sense will go in judging an action to be virtuous if the agent thinks it is good but it isn't;[1] but if we restrict the term 'virtuous' to acts that we regard as right, it's obvious that the realisation of virtue may not be in the power of a given person at a given time because he can't satisfy the required intellectual conditions.[2]

To sum up the results of this rather complicated discussion: Virtue is a quality manifested in doing one's duty (or doing things that go beyond strict duty); it is primarily attributed to the mind or character of the agent, but is known to us only through how it shows up in feelings and acts. To get precise concepts of the particular virtues, therefore, we have to examine the states of consciousness they are manifested in. Examining these, we find that volition

[1] I said earlier that decidedly wrong acts often exhibit character-traits that would be regarded as particular virtues when exhibited in right acts—generosity, courage, patriotism, etc.—and this is especially true of acts that are bad through ignorance.

[2] Common sense on the whole accepts this conclusion, though the acceptance seems to be quite reluctant. The reluctance doesn't show in an inclination to regard as virtuous people who do clearly wrong acts, but rather in an effort to explain their ignorance as caused by some previous willful wrongdoing. We try to persuade ourselves that if, for example, Torquemada didn't know that it was wrong to torture heretics, he would have known if he hadn't *willfully* neglected some means of enlightenment; but in many cases this kind of explanation is unsupported by facts, and I see no ground for accepting it as generally true.

is primarily important, and in some cases almost of sole importance, but that common sense insists on bringing in the element of emotion. Focusing now on the volitional element: most of what we regard as manifestations of virtue are volitions to produce certain particular effects; the general resolve to do duty for duty's sake is indeed thought to be important as a **generally** necessary spring of virtuous action; but it's not thought to be indispensable for virtue in any **particular** case. Similarly with the emotional element: an ardent love of virtue or aversion to vice is in **general** a valuable stimulus to virtuous conduct, but it's not required for virtue in every **particular** case; and in some cases the presence of other emotions. . . .makes the acts better than if they were done from a purely moral motive. But such emotions can't be commanded at will, nor can the knowledge of what ought to be done in a particular case (knowledge that is obviously required to make conduct perfectly virtuous, if we are restricting the term 'virtuous' to right acts). From all this I conclude that although we distinguish virtue from other excellences by its voluntariness—it must be to some extent capable of being realised at will when occasion arises—this voluntariness is a matter of degree; and although a man can always do his duty if he knows it, he can't always realise virtue in the highest degree.

We recognise that even when we *can't* realise virtue immediately at will we have a duty to cultivate it and try to develop it; and similarly with all virtuous habits or dispositions that we are deficient in, so as to make ourselves more likely to perform the corresponding acts in future, however completely such acts are within the control of the will on particular occasions. Because acts of this latter kind are perfectly deliberate, you might think that they don't need any special virtuous habits as long as we know what is right and want strongly enough to do it.[1] But ·that is wrong—good habits do have a role·. In order to fulfill our duties thoroughly, we often have to act suddenly and without deliberation; there is *then* no room for moral reasoning, and sometimes none for explicit moral judgment; so that we need particular habits and dispositions that carry the names of the special virtues, and it's our duty to foster and develop these in any way we can.

The complicated relation I have laid out between virtue and duty must be borne in mind when I discuss the particular virtues in the following chapters. But we've seen that virtue is *mainly* manifested in voluntary actions, which any individual *can* do if he sees that they are right, and which therefore come within our definition of duty; so in most of the following discussion there will be no need to distinguish principles of virtuous conduct from principles of duty, because the definitions of the two will coincide.

3. I said in chapter 1 that the common notions of particular virtues—justice, etc.—are too vague to specify exactly the actions they cover. I was assuming there that rules of duty *ought* to be precisely statable in their full generality; and this assumption was right for the ordinary or jural [see Glossary] view of ethics as concerned with a moral code: if obligations are imposed on someone he ought at least to know what they are, and a law drafted indefinitely must be a bad law. But when we think of virtue as •going beyond strict duty and •not always achievable at will, this assumption is not so clearly appropriate, because from this point of view we find it natural to compare •excellence of conduct with •beauty in works of art. We commonly say that although rules and precise prescriptions can tell you a lot about a work of art,

[1] Hence the Socratic doctrine that 'all virtue is knowledge', on the assumption that a rational being must necessarily wish for what is good.

they can't tell you everything—that the highest excellence is always due to an instinct or tact [see Glossary] that can't be captured in precise formulas. We can describe the beautiful products and to some extent classify their beauties, but we can't prescribe any sure method for producing each kind of beauty. It may be said:

> '*That* is how things stand with virtues; so it's not possible to state an explicit maxim by applying which we can be sure of producing virtuous acts of any kind. All we can do is to give a general account of the virtue—a description, not a definition—and leave it to trained insight to find in any given situation the act that will best realise it.

This view might be called 'aesthetic intuitionism'; I'll have something to say about in chapter 14.1. But our primary task is to examine the larger claims of the rational or jural intuitionists who maintain this:

> Ethics is capable of exact and scientific treatment, having general rules as its first principles; so we have some hope of getting rid of the fluctuations and discrepancies of opinion that we accept calmly in aesthetic discussions but that tend to endanger the authority of ethical beliefs.

We can't evaluate this position without examining in detail the propositions that have been offered as ethical axioms, and seeing how far they prove to be clear and explicit and what clear and explicit rivals they have. The more judicious intuitionists wouldn't maintain that such axioms can always be found, with proper exactness of form, merely by observing men's common moral reasonings. Their view is rather that the axioms are at least *implied in* these reasonings, and that when they are made explicit their truth is self-evident and must be accepted at once by any intelligent and unbiased mind. (Similarly: some mathematical axioms aren't and can't be known to the multitude, because only a carefully prepared mind can see their certainty; but when their terms are properly understood the perception of their absolute truth is immediate and irresistible.) If we can't claim for a precisely formulated moral axiom that it has the explicit actual assent of *everyone*, it may still be a truth that men have already vaguely grasped and that they will now unhesitatingly accept.

In this inquiry it doesn't matter much what order we take the virtues in. [Sidgwick explains that the so-called morality of common sense doesn't provide any clear basis for classifying the virtues; he'll discuss classification in chapter 7. In the meant time he proposes to] approach the topic empirically, as we find it in the common thought expressed in the common language of mankind. . . .

So it seems best to take the virtues in ·descending· order of importance. Some virtues seem to include, in a way, all or most of the others; it will be convenient to begin with these. Of these *wisdom* is perhaps the most obvious; in the next chapter, therefore, I'll examine our common conceptions of wisdom, and of certain other virtues or excellences that are related to it.

Chapter 3: Wisdom and self-control

1. The Greek philosophers always put wisdom first in the list of virtues, and regarded it as including all the others, in a way. The post-Aristotelian schools, indeed, employed the notion of the *Sage* or *ideally wise man* to exhibit in a concrete form the rules of life laid down by each system. [Sidgwick distinguishes theoretical wisdom from practical wisdom, and suggests that in English when (say) a scientist or historian is said to be 'wise', this is usually because he has skills and habits that would also be of service in practical = moral matters, e.g. impartiality, breadth of view, etc. Anyway, Sidgwick's topic is solely practical wisdom.] How then shall we define practical wisdom? The most obvious part of its meaning is

> a tendency to discern in the conduct of life generally the best means to any ends that the natural play of human motives may lead us to seek;

as contrasted with technical skill,

> i.e. the ability to select the best means to given ends in a certain limited and special department of human action.

Such skill in the **special arts** [see Glossary] is partly a matter of grasping definite rules and partly a matter of tact [see Glossary] or instinct, depending somewhat on natural gifts and predispositions but to a large extent acquired through exercise and imitation. If practical wisdom were taken to be skill in the **art of life**, it would involve a certain amount of scientific knowledge, the pats of different sciences bearing directly on human action, together with empirical rules relating to the same subject-matter; and also the tact or trained instinct I have mentioned (it would even be more prominent here because the subject-matter is so complex). But this analysis doesn't show why this skill should be regarded as a virtue; and the fact is that we *don't* ordinarily mean by 'wisdom' merely the faculty of finding the best means to any ends. We wouldn't hesitate to credit an accomplished swindler with cleverness, ingenuity, and other purely intellectual excellences, but we wouldn't call him wise! We call a man who skilfully chooses the best means to his ambitious ends 'worldly-wise', but we wouldn't call him 'wise' without qualification. Wisdom appears to me to imply right judgment in respect of •ends as well as of •means.

Now a subtle question arises. I am assuming in this treatise that there are several ultimate ends of action, all claiming to be rational ends that everyone ought to adopt. So if wisdom implies right judgment about ends then someone who regards some one end as the sole right or rational ultimate end won't consider as *wise* anyone who adopts any other ultimate end. [Sidgwick embarks on an *extremely* difficult account of a problem that might be thought to arise. We can safely slip through that and come out where he does:] Common sense seems to mean by 'a wise man' a man who attains at once all the different rational ends; who by conduct in perfect conformity with the true moral code attains the greatest possible happiness for himself and for the part of mankind that he can affect. But if we find this harmony unattainable—e.g. if rational egoism leads to conduct opposed to the interests of mankind in general, and we ask whether we are to call 'wise' the man who seeks his private interests or the one who sacrifices them—common sense gives no clear reply.

2. If wisdom, as exhibited in right judgment about ends, is in any degree attainable at will, that makes it a 'virtue'

according to my definition. At first sight, the perception of the right end seems to be involuntary, like the knowledge of any other kind of truth. Learning the truth requires voluntary effort in most cases, but we don't think that such effort by anyone is on its own enough for him to reach—even approximately—the right solution of a difficult intellectual problem. It is often said, however, that knowledge of moral truth depends largely on the 'heart', i.e. on our desires and other emotions; this seems to be the basis for regarding wisdom as a virtue; so we might count it as a virtue to the extent that this condition of feeling is attainable at will. But look closer! There's little agreement about what are the right emotional conditions for the knowledge of ends: some would say that prayer or ardent aspiration produces the most favourable state; others would urge that emotional excitement is likely to perturb the judgment, and would say that what we is tranquillity of feeling; some would contend that the essential condition is a complete suppression of selfish impulses; others would regard this as chimerical and impossible (or, if possible, a plain misdirection of effort). Common sense won't settle this; but it would be generally agreed that •certain violent passions and sensual appetites pervert moral apprehensions, and that •these are to some extent under the control of the will; so that when a man who exercises moral effort to resist their influence when trying to decide on ends of action, he is to that extent voluntarily wise.

This also applies, to some extent, to the other function of wisdom, namely selecting the best means to the attainment of given ends. It seems that our insight in practical matters is liable to be perverted by desire and fear, and that this perversion can be prevented by an effort of self-control; so that even here unwisdom is not entirely involuntary. In a dispute that may lead to a quarrel, I may be unable to show any foresight and skill in standing my ground while avoiding needless anger, and in that respect I may be unable to conduct the dispute wisely. But I always *can*, before taking each important step in the dispute, reduce the influence on my decisions of anger or wounded vanity, and in this way I can avoid much unwisdom. Notice that volition plays a larger part in developing or protecting our insight into •the right conduct of life than it has in respect of •the technical skill that I have compared with practical wisdom—because the reasonings in which practical wisdom is exhibited are less clear and exact, and the conclusions are inevitably less certain. Desire and fear could hardly make one go wrong in an arithmetical calculation; but in estimating a balance of complicated practical probabilities it's harder to resist the influence of strong inclination; and it's our awareness of the continual need for such resistance that leads us to regard wisdom as a virtue.

We can say then that the virtue of practical wisdom involves a habit of resistance to desires and fears—what is commonly called 'self-control'. But if a man has determined with full insight what it is reasonable for him to do under any given circumstances, there's still the question of whether he will certainly do it. Now I don't think that common sense takes wisdom to include the •choice of right ends, as distinct from •knowledge of what they are; but if a man deliberately chose—as the modern mind admits to be possible[1]—to do what he knew to be contrary to reason we wouldn't call him wise! It seems that the notion of such a choice is less familiar than either •impulsive irrationality or •mistaken choice of bad for good. [Sidgwick goes on to say that the 'mistaken

[1] I have noted the difference between ancient and modern thought in this respect. I/5.1, note on page 26.

choice' case counts as 'culpable unwisdom' if, but only if, the mistake arises from desires or fears that the agent could have resisted but didn't. It still wouldn't be as bad a loss of self-control as that of the man who 'deliberately chooses to do what he knows to be bad for him'.]

The case of impulsive wrongdoing is different. If I deliberate and reach a decision in accordance with my view of what is right, I shouldn't abandon or change this on an impulse, without fresh deliberation; and the self-control needed for resist such impulses—we could call it 'firmness'—is an indispensable aid to wisdom. But sometimes gusts of impulse sweep in so fast that the person doesn't remember the decision that they run counter to; and when that happens the required self-control or firmness seems to be not attainable at will, just when it is most wanted. Still, we can cultivate this important habit by engraving our decisions more deeply ·into our minds· in moments of deliberation between the moments of impulsive action.

3. When we examine the functions of wisdom, other excellences come into view—partly included in our conception of wisdom, and partly auxiliary to it. Some of these are not what we would call *virtues*, for example

- •sagacity in selecting the really important points amid a crowd of others,
- •acuteness in seeing aids or obstacles that lie somewhat hidden,
- •ingenuity in devising subtle or complicated means to our ends,

and other related qualities more or less vaguely defined. We can't be acute, or ingenious, or sagacious when we please, though we can become more so by practice. That holds also for *caution*, i.e. taking into due account material [here = 'non-negligible'] circumstances that don't favour our purposes. We can't by any effort of will get ourselves to see what circumstances *are* material; we can only look steadily and comprehensively. But there's a sort of self-control that properly counts as a virtue and that can also be called 'caution', namely the tendency to go on deliberating for as long as we think is needed, although powerful impulses urge us to immediate action.[1]

Balanced against caution is another minor virtue, the quality of *decisiveness*, meaning the habit of resisting the irrational impulse that some men have of continuing to deliberate when they know that the time for that has passed and they ought to be acting. . . .

Why then should we classify such qualities as caution and decisiveness as •virtues and not merely as •intellectual excellences? It's because they are species of self-control—i.e. they involve voluntarily standing by rational judgments about conduct in spite of irrational motives pushing the other way. It may seem at first sight that

> given perfect correctness of judgment and perfect self-control, duty will be perfectly done in all parts of life, and virtue will be perfectly realised (except when it demands emotions that can't be commanded at will; see chapter 4.2.)

No doubt a perfectly wise and self-controlled man can't be conceived as breaking or neglecting any moral rule. But even sincere and single-minded efforts to do what we see to be right can vary in intensity; and the tendency to manifest high intensity in such efforts is properly praised as 'energy' (if the quality is purely volitional) or as 'zeal' or 'moral ardour' or the

[1] There's also a third meaning for 'caution'. Of the various means to our chosen end, some are more certain than others, and some are less dangerous (in one way or another) than others; and 'caution' is often used to name the temperament that inclines to the more certain and less dangerous means. . . .

like if the volitional energy is traced to intensity of emotion, though not connected with any emotion more special than the general love of what is right or good.
Note: This chapter hasn't yet brought us to the question at issue between intuitional and utilitarian ethics. Granted that we can by arrive at clear rules of duty concerning wisdom, caution and decisiveness, the rules are obviously not independent: they presuppose an intellectual judgment about what it is right or expedient [see Glossary] to do—a judgment that is or can be obtained in some other way.

* * * * * *

Note: Throughout chapters 4–10 I'll be primarily trying to pin down not •the true morality but •the morality of common sense. If any moral proposition is admitted to be paradoxical, that admission will exclude it—not necessarily as being false, but as not being what common sense holds.

Chapter 4: Benevolence

1. We have seen that the virtue of practical 'wisdom includes all the other virtues, in that virtuous conduct in each part of life results from a clear knowledge and choice of •the true ultimate ends of action and of •the best means to get them. So we can regard the names of the specific virtues as denoting specific departments of this knowledge, to which I now turn.

But when we look into the specific virtues, we find other virtues that can also—in various ways—be regarded as no less comprehensive than wisdom. Especially in modern times, since the revival of independent ethical theorising, there have always been thinkers who have maintained some form of the view that benevolence is a supreme and structural virtue, including and summing up all the others, and fitted to regulate them and settle their proper and inter-relations.[1] This widely supported claim to supremacy is my reason for taking benevolence, next after wisdom, in my examination of the commonly received maxims of duty and virtue.

The general maxim of benevolence would be commonly said to be 'We ought to love all our fellow-men' or '. . . all our fellow-creatures'; but we have seen that moralists don't all agree about the precise meaning of 'love' in this maxim. According to Kant and others, the duty of benevolence is not strictly the affection—with an emotional element—of love or kindness, but only the will's being set on seeking the good or happiness of others. I agree that it can't be a strict duty to feel an emotion that one can't call up at will.

[1] These days one usually meets this view in the form of utilitarianism, which I'll discuss more fully in Book IV; but some version of the view has been held by many who are nearer to the intuitional school.

Still (I repeat) it seems to me that this emotional element *is* included in our common notion of charity or philanthropy, regarded as a virtue; and it's paradoxical to deny that it raises the mere beneficent disposition of the will to a higher level of excellence, and makes its effects better. If this is so, it will be a duty to develop the affection as far as possible; and indeed this seems. . . .to be a normal effect of repeated beneficent decisions and actions. . . . This effect is admittedly less certain than the production of the benevolent disposition; and some men are so unattractive to others that the others can't feel affection for them though they may have benevolent dispositions towards them. Anyway, it seems to be a duty generally to do our best to cultivate kind feelings towards those whom we ought to benefit, not only by doing kind actions but by subjecting ourselves to influences that have been found to have a tendency to produce affection.

But we have still to discover in more detail what kinds of actions display this affection or this disposition of the will. They are described popularly as 'doing good'. Now, I noted in I/7 and I/9 that the ordinary notion of *good* includes the different conceptions—not harmonised with one another because not distinguished from one another in the first place !—that men form of the ultimate end of rational action. So there's a corresponding ambiguity in the phrase 'doing good': many would unhesitatingly take it to mean the promotion of happiness, whereas others, holding that perfection and not happiness is the true ultimate good, maintain that the real way to 'do good' to people is to increase their virtue or aid their progress towards perfection. But even among anti-Epicurean moralists there are some, such as Kant, who oppose this and contend that **(i)** my neighbour's virtue or perfection can't be a goal of mine because it depends on the free exercise of his volition, which I can't help or hinder. But that would let us argue that **(ii)** I can't cultivate virtue in myself, but can only practise it from moment to moment. [Sidgwick is here assuming that **(i)** is supposed to hold not because my neighbour's volition is *his* but because it is *free*.] Yet even Kant doesn't deny that we can cultivate virtuous dispositions in ourselves, doing this in ways other than by performing virtuous acts; and common sense assumes that we can do this and prescribes it as a duty. And it's equally undeniable that we can cultivate virtue in others: that is clearly the object of education and of a large part of social action—especially our expression of praise and blame. And if ·our· virtue is an ultimate end for ourselves, to be sought for its own sake, benevolence must lead us to do what we can to obtain it for our neighbour. . . .

So the common view of what benevolence tells us to promote for others seems not to include any clear selection between the different and possibly conflicting elements of good as commonly conceived. But the promotion of happiness seems to be in practice the chief part of what common sense takes to be prescribed as the duty of benevolence; and for clarity's sake I'll focus on this in the rest of the discussion.[1] By 'happiness' we're not to understand simply

> the gratification of the actual desires of others,

(for men too often desire things that would make them unhappy in the long run) but rather

> the greatest possible amount of pleasure or satisfaction for them on the whole

—in short, the happiness that egoistic hedonism takes to be the rational end for each individual. *This* is what rational benevolence tells us to provide for others. If x loves y, and is led by affectionate sympathy with y's longings to gratify them

[1] A further reason for this focus will appear in chapter 14, when I survey the •general relation of virtue to happiness, as the result of the detailed examination of •particular virtues that is the main subject of Book III.

while believing that this will bring y more pain than pleasure in the long run, we commonly say that such affection is weak and foolish.

2. Towards whom is this disposition or affection to be maintained? and to what extent? Well, it's not quite clear whether we owe benevolence only to men [here = 'human beings'] or to other animals also. It's generally agreed that we ought to treat all animals with kindness by avoiding causing them unnecessary pain; but there's disagreement about whether this is directly due to sentient beings as such, or merely prescribed as a means of cultivating kindly dispositions towards men. Reputable intuitional moralists have maintained this latter view; but I think that common sense •regards this as a hard-hearted paradox and •agrees with Bentham that the pain of animals is *in itself* to be avoided. As for the question of how our benevolence ought to be distributed among our fellow-men, I'll make the intuitional view clear by contrasting it with utilitarianism. Utilitarianism is sometimes said to analyse all virtue into universal and impartial benevolence; it doesn't prescribe that we should love all men equally—only that we should

- •aim at happiness generally as our ultimate end,
- •consider the happiness of any one individual as equally important with the equal happiness of any other, as an element of this total, and
- •distribute our kindness so as to make this total as great as possible,

in whatever way this result can be attained. [Regarding 'this total': *what* total? Sidgwick writes as though he had just used 'total' or something clearly equivalent to it, but he hasn't; that's a fact about the original, not an artifact of this version.] In practice, of course, any individual's services will be distributed unequally, even on this view, because obviously each man will promote the general happiness best by attending to a limited number of people, and to some more than others; but on the utilitarian theory, this inequality is secondary and derivative. Common sense, though, holds that that we owe special dues of kindness to those who stand in special relations to us, and seems to regard this as immediately certain and in no need of supporting argument. So we face this question:

> When there is doubt or apparent conflict of duties, what principles should we use to settle the nature and extent of the special claims. . . .that arise out of these special relations among human beings? Should we solve such problems by considering which course of conduct will do most for the general happiness, or can we find independent and self-evident principles that are clear and precise enough to guide us in such cases?

The different answers to this fundamental question constitute the main difference between the intuitional and utilitarian methods, in the context in which the 'good' that the benevolent man tries to give to others is understood to be happiness.

As we approach the question, we encounter a difficulty about how to *arrange* the topic. Like most problems of classification, this should be taken very seriously because it depends on important characteristics of the material that has to be arranged. Quite often 'benevolence' is taken in a narrower sense in which benevolence is distinguished from justice and even contrasted with it. We can of course be benevolent and just towards the same people, but we commonly assume that the special function of benevolence begins where justice ends. [In the rest of this difficult paragraph, Sidgwick sketches and criticises several attempts to demarcate benevolence from justice, and says that he will sort all this out in chapter 5.2. In the meantime, he says,] it seems proper ·in this chapter· to discuss separately all

duties that arise out of relations where affection normally exists and where it ought to be cultivated, and where its absence is deplored. Everyone agrees that there are such duties, the non-performance of which is a ground for censure, distinct from the obligations imposed by law or arising out of specific contracts. . . .

Beyond these duties, there's a region of performance—clearly belonging to benevolence and not to justice—where the services rendered can't be *claimed* as though they were owed, and where non-performance isn't felt to be *blameworthy*; and here again there's some difficulty in stating the view of common-sense morality. Two questions arise:

(a) Should we count as *virtuous* services that are driven by affection over and above what strict duty is thought to require?

(b) Should such an affection itself be considered worthy of admiration as a moral excellence and thus as a mental condition that we should try to achieve?

[The rest of this paragraph answers Yes to both questions in connection with **some ranges** of beneficiaries.] I think that common sense clearly regards as virtuous •the disposition to do good for men in general and to promote their well-being—whether •it arises from naturally kind feelings towards human beings generally or merely results from moral effort and resolve—provided it is accompanied by sufficient intellectual enlightenment.[1] This is true also of •the less comprehensive affection that leads men to promote the well-being of their community, and of •the affection that normally tends to accompany the recognition of rightful rule or leadership in others. In some ages and countries patriotism and loyalty have been regarded as almost the supreme virtues, and even today common sense ranks them high.

[Now we turn to the two questions in connection with **the remaining range** of beneficiaries.] But when we pass to more restricted and usually more intense affections, such as our feelings for relations and friends, it's harder to settle whether they should be considered as moral excellences and cultivated as such.

Love isn't merely a desire to do good to the loved one, though it always involves such a desire. It is primarily a pleasurable emotion that seems to depend on a certain sense of union with another person, and it includes a desire to be *with* the beloved; and this desire may predominate over the benevolent impulse and even conflict with it, so that the beloved's true interests may be sacrificed. In this case we call the affection selfish, and we blame it. Now, if we put this question to common sense:

> Is intense love for an individual, considered merely as a benevolent impulse, in itself a moral excellence?

it's hard to extract a very definite answer; but I think common sense inclines on the whole to answer No. We're generally inclined to admire any conspicuously 'altruistic' conduct and any form of intense love, however restricted in its scope; but we don't seem to regard the capacity for such individualised benevolent emotions as an essential element of moral perfection that we should try to have and cultivate. . . . Indeed, we seem to doubt whether such an effort is desirable in this case, except where the affection is required for the performance of recognised duties. [The next sentence is hard to sort out. Here it is, exactly as Sidgwick wrote it.] Again, we think it natural and desirable that—as generally speaking each person feels strong affection for only a few individuals—in his efforts to promote directly the well-being of others he should, to a great extent, follow the promptings of such restricted

[1] [Sidgwick has a footnote here saying that one is apt to get better results from thinking and planning than from acting on any impulse; but such thinking and planning tends to interfere with spontaneous kindly impulses; and common sense isn't sure which to prefer.]

affection: but we are hardly prepared to recommend that he should render services to special individuals beyond what he is bound to render, and such as are the natural expression of an eager and overflowing affection, without having any such affection to express; although, as was before said, in certain intimate relations we do not approve of the limits of duty being too exactly measured.

I conclude that—while we praise and admire enthusiastic benevolence and patriotism, and are touched and charmed by the spontaneous outflow of gratitude, friendship, and family affections—what chiefly concerns us as moralists, in this context of services and kind acts that we regard as morally obligatory, is to discover the right rules for distributing them. If a man fulfils these duties (and obeys the other recognised rules of morality), common sense isn't prepared to say *how far* it is right or good that he should sacrifice any other worthy aim—such as the cultivation of knowledge or any of the fine arts—to the claims of philanthropy or personal affection. There seem to be no generally accepted 'intuitional' principles for making such a choice of alternatives. I'll return to this question in chapter 14.

3. Then what *are* the duties that we owe to our fellow-men (apart from the ones that concern justice rather than benevolence)? It may not be hard to list them. We would all agree that everyone ought to be kind to

- his parents and spouse and children, and to other relatives in a lesser degree;
- to people who have rendered services to him,
- to any others whom he accepts as friends;
- and to neighbours and to fellow-countrymen more than others; and perhaps we may say
- to those of our own race more than to black or yellow men, and generally to human beings in proportion to their affinity [here = 'likeness'] to ourselves.

We think we owe the greatest sacrifices (when there's a need for them) to our country as a whole (but in a lower stage of civilisation this debt is thought to be due rather to one's king or chief); and a similar obligation seems to be recognised, though less definitely and in a lesser degree, in relation to guilds and societies etc. of which we are members. . . . These are generally recognised claims; but we run into difficulties and disagreements when we try to be more precise about how far each extends and how they compare in moral force; and the disagreements become indefinitely greater when we compare *our* customs and common opinions with those of other ages and countries. For example, the tie of hospitality used to be specially sacred, and claims arising from it were considered specially stringent; but this has changed as hospitality, in the progress of civilisation, has become a luxury rather than a necessity, and we don't think that we owe much to a man because we have asked him to dinner. Another example, where the change may be happening right now, concerns the claims of relatives on estates of someone who has died. These days we think •that a man ought usually to leave his property to his children and •that if he has no children he can do what he likes with it unless any of his brothers or sisters are in poverty. . . . But not so long ago a childless man was thought to be morally bound to leave his money to his collateral relatives; and it's natural to guess that in the fairly near future any similar obligation to children—unless they are in want or their education isn't completed—will have vanished out of men's minds. A similar change might be traced in what is commonly thought about the duty of children to parents.

It may be maintained that this variation in customs can be allowed for in the definition of duty, because we can lay down that

> The customs of any society ought to be obeyed so long as they are established, just as the laws ought, although both customs and laws can be changed from time to time.

Conformity to established customs is indeed usually a good idea, but on reflection we see that it can't be an absolute duty. . . . Every progressive community has a settled procedure for nullifying •laws that are found to be bad; but •customs can't be formally abolished—the only way we can get rid of them is through individuals' refusal to obey them. So it must sometimes be right to do this, if some customs are vexatious and pernicious, as we frequently judge the customs of ancient and alien communities to be. And if we ·weaken the above prescription by· saying that

> Customs should generally be obeyed, but may be disobeyed when they reach a certain level of inexpediency,

this seems to come down to utilitarianism; for

how Sidgwick completes this sentence: we cannot reasonably rest the general obligation upon one principle, and determine its limits and exceptions by another.

what he seems to be getting at: What is being proposed here amounts to 'Obey established customs unless they are really bad', so the most *basic* way morality comes into this is in judgments about which effects of custom are bad, 'inexpedient', and the only basis we have for judging *that* is utilitarianism.

If the duties I have listed are governed by independent and self-evident principles, the limits of each must be implicitly given in the intuition that reveals the principle.

4. Let us examine in more detail what common sense seems to say about these duties. They seem to fall into four classes:

(1) duties arising out of comparatively permanent relationships that haven't been voluntarily chosen, e.g. with relatives, fellow-citizens, neighbours;

(2) duties arising from relationships that are like those except for being voluntarily contracted, e.g. friendships;

(3) duties springing from special services received, i.e. duties of gratitude; and

(4) duties that seem to be due to special need, i.e. duties of pity.

This classification will be a convenient basis for discussion; but it doesn't clearly and completely avoid cross-divisions, because (example) **(3)** the principle of gratitude is often appealed to as supplying the rationale for **(1)** the duties children owe to parents. But does it? There's a substantive disagreement and difficulty about what maxim—what basic moral principle—underlies this species of duty. It would be agreed that children owe to their parents respect and kindness generally, and help in case of infirmity or any special need; but how far is this based on services rendered by the parents to the child during infancy, and how far is it based on the child/parent relationship alone, so that it is due even to cruel or neglectful parents? It is not clear how common sense answers this. Most people might say that mere nearness of blood creates a certain obligation, but they would find it hard to agree on its exact force.[1]

But when parents *have* done their duty ·towards their children·, there's great difference of opinion about other aspects of the children's duty to them. For example, when the offspring is no longer in his or her parents' guardianship

[1] It may be said that a child owes gratitude to the authors of its existence. But life alone, apart from any provision for making life happy, seems to be a gift of doubtful value, and one that hardly arouses gratitude when it wasn't given out of any regard for the recipient.

or dependent on them for support, how far does he or she owe them obedience? Is a son or a daughter morally obliged not to oppose a parent's wishes in marrying or choosing a profession? In practice, parental control is greater when the parents have wealth that they can bequeath to their children, but this is irrelevant to the basic moral question about the ideal of filial duty; for that is the question of whether the child is *absolutely* bound to obey his or her parents—absolutely and not as a quid pro quo for expected future benefits.. . . .

Now what about the duty of parents to their children? This too we might put partly into a different category, namely duties arising out of special needs; for the helplessness of children makes them natural objects of compassion, and not only for their parents. But they also have another kind of claim on their parents, arising from the duty—everyone recognises this—of not directly or indirectly causing pain or harm to other human beings, except in the way of deserved punishment; for the parent is a cause of the child's existing in a helpless condition, and would be indirectly the cause of the suffering and death that would result if the child were neglected. But this doesn't fully account for parental duty as recognised by common sense. We commonly blame a parent who leaves his children entirely to the care of others, even if he provides for their being nourished and trained up to the time when they can become independent. We think that the parent owes them •affection (as far as this can be said to be a duty) and •the tender care that naturally springs from affection and •something more than the necessary minimum of food, clothing, and education (if he can afford it). But it's not clear *how far* beyond the minimum he is bound to go. It's easy to say sweepingly that he ought to promote his children's happiness by all means in his power,. . . .but it seems unreasonable to require him to purchase a small increase in their happiness by a great sacrifice of his own; and also there are other worthy ends that can (and do) come into competition with this. Consider these two cases:

- •A parent is led to give up important and valuable work that no-one else can or will do, in order to leave his children a little more wealth;
- •A parent brings his children to the verge of starvation in order to conduct scientific research.

We condemn both extremes; but what clear and accepted principle can be stated for determining the true mean?

As I have pointed out. some people think that a parent has no right to bequeath his inheritance away from his children, unless they have been undutiful; and in some states even *that* is forbidden by law. Others hold that children don't have claims to their parents' wealth just by being their children; they have such claims only if •there's a tacit understanding that they will succeed to it, or •they have been brought up with habits of life and social relations that would make it difficult and painful for them to live without inherited wealth.

Our conception of the mutual duties of kinsmen becomes vaguer as the kinship becomes more remote, and it would tedious to go into all that in detail. Among children of the same parents, brought up together, some level of affection grows up so naturally that we regard those who feel no affection for their siblings with aversion and moral contempt, as somewhat inhuman; and we think that, whatever the circumstances, the services and kind acts that naturally spring from affection ought to be rendered to some extent; but the extent seems quite undefined. Even towards more distant relatives, we think, men of good dispositions will have a certain flow of kindly feeling. . . . Some people still think that *cousins* have a moral right to a man's inheritance in default of nearer heirs, and to assistance when they are in need; but it seems equally common to hold that cousins

can at most claim to be selected 'other things being equal' as the recipients of bounty, and that an unpromising cousin shouldn't be preferred to a promising stranger.

5. In speaking of those who are thought to have a certain claim for services, I included not only relatives but also neighbours. Probably no-one thinks that merely living-in-the-same-locality is a ground of duties; it seems rather that neighbours naturally feel more sympathy with one another than with strangers, because the tie of common humanity is strengthened even by such relations as living-close-together and meeting-on-the-street (without cooperation or friendship), and a man in whom this effect is not produced is thought to be somewhat inhuman. [Sidgwick goes on to say that this doesn't work in large towns where one has too many 'neighbours' to sympathise with them all. In that case, help for fellow-townsmen belongs to a different category. There are minor helps that it is thought all right for anyone to ask of anyone in time of need, and very slight factors may give *direction* to such appeals for help, making it seem more natural to appeal to folk who live in the same neighbourhood or town as oneself than to appeal to others; and preference is also given to those of the same age, sex, rank, or profession as oneself, the basis for all of this being a presumption that where any of those relations obtains there is likely to more sympathy. Sidgwick sums up:] The duty towards one's near neighbours seems to be only a particular application of the duty of general benevolence. The claim of fellow-countrymen is of the same kind; but only if they are taken •as individuals, because one's relation to one's country •as a whole is thought to involve much stricter obligations.

But the duties of patriotism are hard to formulate. Morality requires all a country's inhabitants to obey its laws, but that isn't patriotism because it applies equally to aliens living in the country in question. And in most of men's social functions [here = 'most of the things men do in the government service'] patriotism isn't involved, or anyway isn't a prominent or indispensable motive; for men do these things primarily for the sake of payment; and having undertaken them they are bound by justice and good faith—·not by patriotism·—to perform them adequately. If any of the functions of government are unpaid, we think that men exhibit patriotism in performing them. It's plausible to say that they do get paid, namely in social distinction, but on reflection that doesn't seem right: social distinction is meant to express honour and respect, and we can't come up with these feelings as part of a bargain—they're a tribute paid to virtue or excellence. But it's not clear how far anyone is obliged to do such work; and the question is usually decided on grounds of expediency, except when duties of this kind are legally or constitutionally imposed on all the citizens in a free country. . . . The duty of fighting the nation's enemies is a prominent example in many countries; and even in countries where the army is manned by paid professional volunteers, such service is often felt to be in a special sense the 'service of one's country', and we prefer it to be performed with feelings of patriotism, because we are repelled by the idea of a man's slaughtering his fellow-men for hire. And when a nation is in danger patriotic feelings are naturally intensified; and even in ordinary times we praise a man who serves his country beyond the common duties of citizenship. But whether a citizen is ever *morally* bound to more than certain legally determined duties isn't clear; and there's no general agreement on the question whether by voluntary expatriation he can clear himself of all moral obligations to the community he was born into.

Nor does there seem to be any consensus as to what each man owes to his fellow-men, just as fellow-men. The utilitarian doctrine. . . .is that each man ought to regard anyone else's happiness as being *theoretically* on a par with

his own, and only of less importance *practically* because he can be more efficient in making himself happy than in doing this for anyone else. And this ·stern utilitarian doctrine· *may* be the principle of general benevolence that common sense recognises. But admittedly there is also some acceptance of a lower and narrower estimate of the services that we are held to be strictly bound to render to our fellow-men generally. This lower view seems to recognise

(1) a negative duty to abstain from causing pain or harm to any fellow-man except in the course of deserved punishment; which immediately implies the duty of making reparation for any harm we have done anyone;[1] and

(2) a positive duty to give services that require little or no sacrifice on our part.

And a general obligation to be 'useful to society' by some kind of systematic work is vaguely recognised; rich persons who are obvious drones are criticised by most thoughtful people. Beyond this somewhat indefinite limit of **duty** extends the **virtue** of benevolence without limit; for there's no such thing as going too far in helping others unless it leads us to neglect definite duties.

The notion of benevolence as I have defined it covers the minor rules of gentleness, politeness, courtesy, etc. because these require •the expression of general goodwill and •avoiding anything that may cause pain to others in conversation and social demeanour. But one part of politeness deserves separate treatment—the duty of showing marks of reverence to those who are entitled to them.

We can define *reverence* as the feeling that accompanies the recognition of superiority or worth in others. It doesn't have to be benevolent, though it is often accompanied by some degree of love. But its ethical characteristics are like those of benevolent affection in this: it's a feeling that isn't directly under the control of the will, yet we expect it under certain circumstances and morally dislike its absence; and it seems that we sometimes think that a person has a duty to express reverence even if he doesn't have the actual feeling.

But there's a great divergence of opinion on this last point. The feeling seems to be naturally aroused by all kinds of superiority—not merely •moral and intellectual excellences, but also •superiorities of rank and position—and it's to •the latter that men more regularly and formally give this tribute. And yet it's often said that reverence is more properly due to •the former, because they are more real and intrinsic superiorities; and reverence towards men of rank and position is seen as servile and degrading; and some people even dislike the marks of respect that official superiors demand from their subordinates in most countries. . . .

A similar but more serious difficulty arises over the question: How far is it a duty to develop the affection of *loyalty*?—meaning by this ambiguous term the affection that a well-disposed servant or official subordinate normally feels towards a good master or official superior. It is widely thought that the duties of obedience that belong to these relations—like the duties of the family relations—will be better performed if affection enters into the motive. But it seems to be a tenable view that orderliness and good faith—ungrudging obedience to law and fulfilment of contract—is ordinarily enough, without personal affection; and a disposition to obey superiors beyond the limits of their legal or contractual rights can easily be harmful in its effects if

[1] How far are we bound to make reparation when the harm is involuntary and couldn't have been prevented by ordinary care on our part? It's not clear what the answer to this is; but I'll postpone discussion of that until chapter 5.5 , because the whole of this department of duty is usually classified as pertaining to justice.

the superiors have bad characters. It's clearly a good thing if inferiors are disposed to obey a wise and good superior beyond these limits; but it doesn't follow that we have a *duty* to cultivate this disposition in cases where it doesn't result from a sense of the superior's goodness and wisdom. And I don't think that common sense has anything firm and clear to say about this.

6. I now turn to the duties of affection that arise out of voluntarily acquired relationships [this is number **(2)** of the four on page 117]. The most important of these is the marriage relation. First question: Is it the duty of human beings generally to get married? It's normal to do so, of course, and most people are prompted to it by strong desires; but common sense doesn't seem to prescribe marriage as an independent duty but only derivative from and subordinate to the general maxims of prudence and benevolence.[1] And in all modern civilised societies, law and custom leave the marriage union perfectly optional; but there are carefully devised laws about the conditions under which it may be formed, and some of the mutual rights and duties arising out of it; and many people think that this branch of the law ought to be especially strongly governed by independent moral principles. . . . What are these principles? Common sense in modern European communities seems to answer that the marriage union ought to be

(i) exclusively monogamous,
(ii) at least *designed* to be permanent, and
(iii) between persons who are not too closely related.

But I don't think that any of these propositions can, on reflection, be claimed to be self-evident. Even against **(iii)** incest we seem to have an intense feeling rather than a clear intuition; and **(i)** it is generally recognised that the only *reasons* for insisting on monogamy are utilitarian ones.[2] As regards **(ii)** the permanence of the marriage-contract, no doubt everyone would agree that fidelity is admirable in all affectionate relationships, and especially in such a close and intimate one as the marriage relation; but we can't tell in advance how far love can be maintained in all cases; and it certainly isn't self-evident •that a marriage ought to be maintained when love has ceased, or •that if the spouses have separated by mutual consent they ought to be prohibited from forming fresh unions. . . .

And in considering what can go wrong in a marriage, people differ widely over what kind of feeling they think is morally indispensable to this relation. Some hold that marriage without intense and exclusive affection is degrading even though sanctioned by law; while others consider this a mere matter of taste, or at least of prudence, provided there's no mutual deception; and there are various intermediate views between these two.

Nor is there agreement about the external duties of the marriage relation. Everyone will agree on two things: •sexual fidelity, and •mutual assistance. . . . But beyond this opinions diverge: some say that 'the marriage contract binds each party, whenever individual gratification is concerned, to prefer the happiness of the other party to its own' (Wayland,

[1] I raise this question because if the rule of 'living according to nature' were really adopted as a first principle, in any ordinary sense of 'nature', it *would* seem to be the duty of all normal human beings to enter into marriage relations; but this very instance shows that the principle is not accepted by common sense. See I/6.2.

[2] The prohibition of polygamy is sometimes said to be implied by the numerical equality of the two sexes. But this presupposes that all men and women ought to marry; and who would explicitly affirm *that*? In fact, many remain unmarried; and there's no evidence that in countries where polygamy is allowed, a shortage of available women has ever made it practically difficult for any man to find a mate.

Elements of Moral Science); while others would say that. . . .as a mere matter of duty it is enough if each spouse considers the other's happiness equally with his or her own. As for the powers and liberties that ought to be allowed to the wife, and the obedience she owes her husband, I don't need now in 1874 [the date of the first edition] to waste space proving that there is no consensus of moral opinion!

7. The conjugal relation originates in free choice; but once it has been formed, the duties of affection arising out of it are commonly thought to be like those arising out of blood-relationships. So it has an intermediate position between these latter, and ordinary friendships, partnerships, and associations that men are free to make and equally free to dissolve. Most of men's associations are for certain definite ends, fixed by explicit contract or tacit understanding; so the duty arising out of them is merely that of fidelity to this contract or understanding, which I'll deal with later under the headings of justice [chapter 5] and good faith [chapter 6.5]. But this seems not to be the case with *friendships*;[1] for although •friendship frequently arises among people who have come together for other purposes, yet •the relation is always thought of as being its own purpose, and as being formed primarily for the development of mutual affection between the friends and the pleasure that comes from this. Still, it is thought that once such an affection *is* formed it creates mutual duties that didn't previously exist. Let us see how far this is the case, and on what the principles are that create the duties.

Now we encounter yet another difficulty in trying to formulate common sense. Some people say that it's essential to friendship that the mutual kindly feeling and the conduct arising from it should be spontaneous and unforced, neither the feeling nor the conduct should be required as a duty—and that this part of life should be fenced off from the intrusion of moral precepts, and left to the free play of natural instinct. Perhaps everyone would agree that there is something in this; and we have accepted it with regard to all the deeper flow and finer expression of feeling even in the domestic relations; for it seemed pedantic and futile to prescribe rules for this, or. . . .to delineate an ideal of excellence for everyone to aim at. Still, there seemed to be an important sphere of strict duty—however hard to define—in the relations of children to parents etc., and common sense seems to recognise some such sphere even in friendship, because it often enough occurs to us to judge that one friend has behaved wrongly to another, and to speak as if there were a knowable code of behaviour in such relations.

Perhaps we can say that all clear cases of wrong conduct towards friends are instances of *breach of understanding*. . . . All love is understood to include—among other things—a desire for the happiness of the loved one, so announcing one's friendship for someone seems to bind one to seek that person's happiness. . . . Now, common benevolence (see section 5 above) prescribes that we should at least give other men such services as we can give without any significant sacrifice on our part. And declaring someone to be one's friend—even with all the ambiguity of that term—at least implies a greater interest in his (or her) happiness than in that of men in general, so it must announce a willingness to make some real sacrifices for him (or her) if there's an opportunity for that. So if we refuse to make such sacrifices, we do wrong by failing to fulfill legitimate expectations. Up to here there's no difficulty except what comes from the great

[1] I'm using that word in a narrow sense in which it implies a mutual affection more intense than the kindly feeling that a moral man hopes to have towards everyone he is frequently in contact with through business or otherwise.

ambiguity of the term 'friendship'. But further questions arise because of changes of feeling: •is it our duty to resist such changes as much as we can? •and if this effort fails and love diminishes or departs, ought we to maintain a disposition to render services corresponding to our past affection. Moral and refined persons don't agree in their answers to these questions. **On one hand**, we naturally admire fidelity in friendship and stability of affections, and we commonly regard these as important excellences of character, which we try to imitate; so it seems strange not to aim at them. So many people would rule that we *ought not* to withdraw affection once it has been given, unless the friend behaves badly, and some would say that even then we oughtn't to break the friendship unless the friend's crime is very great. Yet **on the other hand**, we feel that affection produced by an effort of will is a poor substitute for the affection that arises spontaneously, and most refined persons would reject friendship on those terms such a boon. Also, ·to continue a friendship in that way you would have· to conceal the change in your feelings, and that seems insincere and hypocritical.

A refined person wouldn't accept help from a former friend who no longer loves him, unless he was in extreme need. . . . So *perhaps* there can't be a duty to offer such help when the need is not extreme; but it's not clear that this follows, because in relations of affection we often praise one person for offering something that we rather blame the other person for accepting. Delicate questions of this kind seem to have more to do with standards of good taste and refined feeling than with morality. . . .

Summing up: the chief difficulties in determining the moral obligations of friendship arise from •the indefiniteness of the tacit understanding implied in the relation, and •disagreements about how far fidelity is a positive duty. The latter difficulty is especially prominent in respect of those intimacies between persons of different sex that precede and prepare the way for marriage [that sentence is from Sidgwick verbatim].

8. I turn to the third item, *gratitude* [i.e. third of the four on page 117]. I have already noted that the obligation of children to parents is sometimes based on this; and in other affectionate relationships gratitude often blends with and strengthens the claims that are thought to arise out of the relations themselves; though none of the duties I have discussed seems to come down to nothing but gratitude. But where gratitude is due, the obligation is especially clear and simple. Indeed it seems that every community that has *any* morality recognises the duty of requiting [see Glossary] benefits, and intuitionists have reasonably taken this as an instance of a truly universal intuition. Still, though the general force of the obligation isn't open to doubt. . . .its nature and extent are not equally clear.

First question: If someone does something kind and helpful for me, how am I *obliged* to respond? Do I owe him only some repayment or requital, or do I owe him the special affection called *gratitude* which seems to combine •kindly feeling and eagerness to requite with •some sort of emotional recognition of superiority—as a benefactor is in a position of superiority to the beneficiary. On the one hand we seem to think that if any affection can be a duty then kindly feeling towards benefactors must be such; yet someone might find this feeling hard to attain because he dislikes the position of inferiority; and we somewhat approve of this feeling and call it 'independence' or 'proper pride'. But this feeling and the outpouring of gratitude don't easily mix, and the moralist can't recommend a proper combination of the two. Perhaps it makes a difference whether the service was lovingly done: if it was, it seems inhuman not to respond with affection; whereas if the benefit was coldly given, it seems enough

to recognise the obligation and have a settled disposition to repay. And 'independence' alone would prompt a man to repay the benefit in order to escape from the burden of obligation. But it seems doubtful whether we are morally satisfied with this as the sole motive for requital, whatever the spirit in which the original help was given.

It's partly this dislike of obligation that makes a man want to give more than he has received; for otherwise his benefactor has still the superiority of having taken the initiative. But the worthier motive of affection urges us in the same direction; and here we don't like too exact a measure of duty—a certain excess falling short of extravagance seems to be what we admire and praise. But when a conflict of claims creates a need to be exact, we think perhaps that an equal return is what the duty of gratitude requires, or rather willingness to make such a return, if it be required, and if it is in our power to make it without neglecting prior claims. [Sidgwick devotes a whole page to variations on the theme of 'equal return'—equal in the amount of sacrifice involved or equal in the amount of benefit conferred? One example: 'a poor man sees a rich one drowning and pulls him out of the water'—we approve of a requital somewhere *between* an amount commensurate with the size of the benefit (the beneficiary's life) and one commensurate with the benefactor's sacrifice (a few minutes of physical effort), but, Sidgwick says, there seems to be no clear accepted principle that settles *where* between those extremes the requital/repayment/reward should be.]

The last claim to be considered is that of *special need* [this is number **(4)** of the quartet on page 117]. I have substantially dealt with this already when investigating the obligation of •general benevolence or •common humanity: we owe to all men (I said back there) any help that we can give through a sacrifice that is small in comparison with the help; so the more urgent the needs of others are, the stricter our duty is to relieve them. But I bring this duty up again here because the specific emotion of pity or compassion prompt us to fulfill it. There seems to be a doubt concerning how far it is good to develop and encourage •this emotion as distinct from •the practical disposition to give prompt help to those in distress whenever such help is judged to be right. On the one hand, the emotional impulse tends to make the helpful action not only easier for the agent but more graceful and pleasing; on the other hand, it's generally recognised that mistaken pity is likely to do harm. It is in this respect worse than (for example) mistaken gratitude—because it's more likely •to interfere with the penal system required for the maintenance of social order, or •to weaken the motives to hard work and thrift that are necessary for economic well-being.

To guard against that last-mentioned danger, we try to define the duty of relieving want; and this brings us face-to-face with a serious *practical* perplexity—not a mere theoretical problem—for most moral persons at the present time. Many people ask:

> Isn't it our duty to refrain from all spending on things that we don't *need* until we have removed—as far as money can do this—the misery and want that exist around us?

In answering this, common sense is led to consider the economic consequences of trying to provide a sufficient income for all needy members of the community, either by •taxation and public expenditure or by •the voluntary gifts of private persons; and thus it comes to replace the intuitional method of dealing with such problems by a different procedure that is, to put it mildly, very like the utilitarian method. (See IV/3.3.)

So it must be admitted that although common sense unhesitatingly lays down some broad and fairly indefinite

rules regarding benevolence, it's hard or impossible to extract from them any clear and precise principles that pin down the extent of the duty in each case. And yet we *need* such particular principles telling us how to distribute the services that good-will prompts us to. It's not merely a matter of theoretical completeness; we need them also to get our *conduct* right, because the duties we have been considering are liable to come into apparent conflict with each other and with other prescriptions of the moral code.

Here's something that might be said in reply to this:

> If we want exactness about what our duty is, we have started in the wrong place: we ought to be examining *justice* rather than *benevolence*.

Well, perhaps the exactness that we sometimes need in practice *can* be found under the heading of *justice*. I'll examine this contention in the next chapter.

Chapter 5: Justice

1. In mapping the outline of duty as intuitively recognised, we have been forced to give some common words a definite and precise meaning. This process of definition always requires care, and is sometimes very hard to do. But there's no word for which the difficulty is greater, or the result more controversial, than the word 'justice'.

Remember what we have to do. We're not going to inquire into the derivation of the notion of justice: our topic is the actual state of our ethical thought, not its history. And we don't aim to provide a definition that will correspond to every part of the common usage of the term 'justice'—there's too much vagueness and looseness in current moral language for that to be feasible. But the intuitional method does assume[1] that the term 'justice' denotes a quality that it is ultimately desirable to realise in the conduct and social relations of men; and that this can be given a definition that all competent judges will accept as a •clear and •explicit version of what they have always •vaguely and •implicitly meant by 'justice'. In seeking such a definition we may trim the ragged edges of common usage but we mustn't cut out any considerable portion of it. [Sidgwick here has a footnote about Aristotle's point that the Greek word for 'justice' has two meanings, one covering much of the territory of virtue generally. This wider meaning is one that the English 'justice' doesn't have, Sidgwick says, and he will use the word only in its 'more precise' (he means 'narrower') meaning.]

When we think about justice, the first point that strikes us may be its connection with law. *Just* conduct is certainly to a great extent determined by *law*, and in some contexts the two terms seem interchangeable. We speak indifferently of 'law courts' and 'courts of justice'; and usually when a citizen demands 'justice' or his 'just rights' he is demanding that law should be carried into effect. Still reflection shows that by 'justice' we don't mean merely 'conformity to law'.

[1] I'll consider in IV/1 what need the utilitarian method has for an independent principle of justice.

(a) We don't call *all* the violators of law 'unjust'—e.g. we don't say this about duellists or gamblers. **(b)** We often judge that some actual law isn't completely just; our notion of justice provides a standard by which we judge actual laws and pronounce them just or unjust. **(c)** Thirdly, some aspects of just conduct lie outside the sphere even of law as it ought to be; for example, we think that a father may be just or unjust to his children in matters where the law rightly leaves him free.

So we must distinguish •justice from what has been called the virtue or duty of •law-observance. If we examine the ·three· points of divergence listed above, perhaps we'll be led to the true definition of *justice*. [In fact, this whole long chapter falls within the scope of point **(b)**. Material relevant to **(a)** and **(c)** comes in only incidentally.]

It is generally thought that the observance of *some* laws is a realisation of justice—what kind of laws are these? Most are laws that define and secure the interests of specific individuals; but that doesn't cover them all, because justice is involved in the assignment of punishments to offenders, but we wouldn't say that a man had an interest in the adequacy of his punishment! Well, then, let us try this:

> The laws in which justice is or ought to be realised are laws that distribute to individuals either
>
> •objects of desire, liberties and privileges, or
>
> •burdens and restraints, or even pains,
>
> with the burdens etc. being allotted by law only to persons who have broken other laws.

All law is enforced by penalties, so this definition lets us view the administration of law generally as the administration of justice—not because all laws are primarily intended as distributive, but because applying the law generally involves (or *should* involve) the proper assignment of pains and losses and restraints to those who violate it. What conditions, then, must laws fulfill to be just in their distributive effects?

You may think that I'm crossing the line that divides ethics from politics; because ethics primarily concerns the rules that ought to govern individual conduct, and it's commonly thought that individuals ought to obey all laws established by lawful authority, even ones they regard as unjust. But some people question this in the case of laws that seem extremely unjust—such as the Fugitive Slave law in the United States before the rebellion [= the civil war]. Anyway, I think I should here digress somewhat into political discussion—•to elucidate the notion of justice, which seems to be essentially the same in ethics and politics, and •because when individuals are regulating their conduct in areas that don't involve law-observance, they *need* to know whether the laws and established order of their society are just or unjust. [This political 'digression' runs to the end of the chapter, it seems.]

Now we all think of just laws as *equal*; and in at least some branches of legislation the common notion of •justice seems to be equivalent to that of •equality. It is commonly thought that a perfectly just system of taxation would impose exactly equal burdens on all;[1] and though this notion of 'equal burden' is hard to define precisely enough for practical application, we can say that justice is here being thought to come down to a kind of equality. But we can't say that *all* laws ought to affect all persons equally; because some laws assign special privileges and burdens to special classes of

[1] This, in my view, doesn't apply to •payments by individuals for services they have received from government. For •them, justice is held to lie in duly proportioning payment to the service received. Some writers have held that *all* payments made to government should be based on this principle; and this seems to be consistent with the individualistic ideal of political order, which I shall presently examine; but, as I have tried to show in *Principles of Political Economy* III/8, there's an important part of governmental expenditure to which this principle isn't applicable.

the community, and we don't think that such laws must be unjust. For example, we don't think it unjust that

- only persons who have been appointed in a certain way should share in legislation, or that
- men but not women should be forced to fight for their country.

So some have said that for a law to be just, all it needs—so far as 'equality' is concerned—is to affect *equally* all the members of any class specified in the law. This does indeed exclude a very real kind of injustice: it's vitally important that judges and administrators should never be persuaded—by money or otherwise—to show 'respect of persons' [i.e. 'to single out some individuals for special treatment']. But that much equality is automatically embodied in any law that is stated in general terms; and obviously laws can be equally applied and yet unjust—e.g. a law compelling only red-haired men to serve in the army and applied with strict impartiality to all red-haired men. We must therefore conclude that, in laying down the law no less than in carrying it out, all inequality[1] affecting the interests of individuals that appears arbitrary, and for which no sufficient reason can be given, is held to be unjust. But we have still to ask: What kind of reasons for inequality does justice allow, and from what general principle(s) can such reasons be deduced?

2. As an aid to answering this question, let us examine the notion of justice as applied to the part of private conduct that lies beyond the sphere of law. Here too the notion of *justice* always involves allotment of something regarded as

- advantageous: money or other material means of happiness; or praise or affection or other immaterial good, or
- disadvantageous: some deserved pain or loss.

In chapter 4.3 I raised the question of how to classify the duties there discussed under the heads of justice and benevolence respectively [see page 116]; and now I'm in a position to answer it. The notion of *justice* has no bearing on any carrying out of a duty of the affections; but it *is* applicable when we •compare the obligations arising from different affectionate relations, and •consider the right allotment of love and kind services. To make this allotment properly we have to find out what is just. Well, then, what do we mean by a 'just man' in contexts where law-observance has no place? It's natural to reply that we mean an *impartial* man, one who tries to satisfy all claims that he recognises as valid and isn't unduly influenced by personal preferences. And this seems a good enough account of the virtue of justice considered as a frame of mind, without bringing in the question of acting in a way that is objectively just: if we neglect to give due consideration to any claim that we **regard as** reasonable, our action can't be just in intention. The definition does exclude deliberate injustice, but it obviously doesn't give us a sufficient condition for an act to be just, any more than the absence of arbitrary inequality is a sufficient condition for a law to be just.[2] [Note that Sidgwick is here handling two distinctions at once, contrasting •a necessary condition for a •frame of mind to be

[1] It can happen that the words of a statute—through careless drafting or through the inevitable defects of even the most precise terminology—include (or exclude) persons and circumstances that are clearly not included in (or excluded from) the real intent and purpose of the law. When that happens, a strictly equal application of a law that is generally considered just can cause extreme injustice; and this sharply brings out the difference between actual law and justice. But thinking about this kind of case won't help us to find principles for judging generally concerning the justice of laws.

[2] We can't even say, in treating of the private conduct of individuals, that all arbitrary inequality is recognised as unjust: if a rich bachelor with no near relatives leaves his property to provide pensions for indigent red-haired men, this might strike us as unreasonable and capricious, but it wouldn't be commonly be thought to be unjust.

just with •a sufficient condition for an •action to be just.] We want to know what claims **are** reasonable.

Well, the most important claims—apart from the ones discussed in chapter 4—seem to be those that result from contracts. Some such claims are enforced by law; but it's clear to us that a just man will keep *all* his engagements, even when there's no legal penalty for violating them. In chapter 6 I'll discuss the exact definition of this duty, and its commonly admitted qualifications; but I can say now that common sense has no doubt that it is generally binding.

We count as *binding engagements* not merely •explicit promises but also 'implied contracts' or 'tacit understandings'. But this latter term is hard to keep precise. It is often used to include not only the case where

> x has somehow positively implied a pledge to y,

but also the case where

> y has certain expectations of which x is aware.

But in the latter cases the obligation is not so clear. We wouldn't say that a man is obliged to dispel all wrong expectations that he knows people have respecting his conduct. . . . But if an expectation is one that most people would form under those circumstances, there seems to be some sort of moral obligation to fulfill it if it doesn't conflict with other duties—though this obligation seems less definite and stringent than the obligation to keep a contract. I think we can make this general (though admittedly vague) statement:

> Justice requires us to fulfill all expectations (of services, etc.) that arise naturally and normally out of the relations—voluntary or involuntary—in which we stand towards other human beings.

But the discussions in chapter 4 have shown hard it is to •define some duties that seem certain and indisputable when looked at in an •indefinite form; while other duties are defined only by customs that appear arbitrary when looked at coolly. As long as these customs persist, the expectations arising from them are 'natural', so that a just man seems to be somewhat obliged to fulfill them; but this obligation can't be regarded as clear or complete, for two reasons that were given in chapter 4: **(a)** When a custom is changing—growing or decaying—the validity of a claim based on it is obviously doubtful. **(b)** It doesn't seem right that an irrational and harmful custom should last for ever, but a custom can be abolished only by being 'more honoured in the breach than in the observance'.

[He means '. . . by being disregarded **more often** than it is observed'. This reflects a misunderstanding—a very common one—of something in Shakespeare. Hamlet is talking about the Danish king's custom of holding drunken revels:

> But to my mind—though I am native here,
> And to the manner born—it is a custom
> More honoured in the breach than the observance.

He means that it would be **more honourable** to disregard the custom than to go along with it.]

So this line of thought has landed us in a real puzzle about the part of duty that we are now examining. We think of justice as something that is in itself perfectly •definite; a scrupulously just man, we think, must be very •exact and •precise in his conduct. But when we think about the justice of satisfying *natural and customary* claims that arise independently of contracts, it seems impossible to estimate these claims exactly. The attempt to map the territory of *justice* reveals a dim borderland inhabited by •expectations that aren't quite •claims, and we're not sure whether justice requires us to satisfy them. Men's ordinary actions reflect the expectation that the future will resemble the past; so it seems natural to expect that a man will do as others do in similar circumstances, and even more that he will continue to do whatever *he* has habitually done; and if he breaks away from some such habit in a way that causes loss or

inconvenience to others, they will be apt to think themselves wronged.[1] On the other hand, if a man hasn't promised to maintain a custom or habit, it seems hard that he should be bound by the unwarranted expectations of others. In this perplexity, common sense often appears to give different answers merely on the basis of how much disappointment is caused by the change. Suppose that customer x leaves tradesman y because y has become a quaker. If x is poor and y is affluent, we might think x's conduct is unreasonable but we would hardly call it unjust; but if x is rich and y is poor, many people would say that x's conduct is unjust persecution.

The same difficulty arises over the duties of kindness discussed in chapter 4—even the stringent and sacred duties of •the domestic affections and of •gratitude. If we are wrestling with a conflict among such duties, we won't get any help from asking 'What does justice require of us?' Bringing justice into it doesn't help us to *solve* our problem; it merely lets us see it from a new angle, as a problem about the right distribution of kind services. If we had clear and precise intuitive principles for settling the claims of parents on children, children on parents, benefactors on beneficiaries, and so on, we could pin-point exactly the place where one of these claims ought in justice to be subordinated to another;. . . .but the only methods I know of for settling a problem of this kind are either •implicitly utilitarian or •arbitrarily dogmatic and unsupported by common sense.

3. Coming back now to the political question: we see that the preceding discussion has provided one of the criteria of the justice of laws that we were seeking, namely that they must avoid running counter to natural and normal expectations; but we also see that the criterion can't be made definite •in application to individual conduct or—I now add—•in application to laws. Why not? Well, law itself is a main source of natural expectations; and alterations in law are usually very small in proportion to what is left unaltered, so we always expect that the existing laws will be maintained. This is an indefinite and uncertain expectation in a society like ours, where laws are frequently altered by lawful authority; but it's sufficient for people in general to rely on in arranging their concerns, investing their money, choosing where to live and what profession to follow, etc. When such expectations are disappointed by a change in the law, the disappointed people complain of injustice, and there's some recognition that justice requires that they be compensated for this loss. But such expectations vary greatly in how definite they are and how important; and in general the less value an expectation has, the more people who have it—like the ripples from a stone thrown into a pond. It's impossible to compensate them all; but I don't know any intuitive principle for separating valid claims from invalid ones, and distinguishing injustice from mere hardship.[2]

Even if this difficulty were overcome, the criterion given above (·namely that just laws must avoid running counter to natural and normal expectations·) is incomplete or imperfectly stated; for if it were complete (·i.e. if it were a *sufficient* condition for a law to be just·) it would imply that no old law could be unjust, because laws that have existed for a long

[1] Some claims generated in this way are legally valid: a ·legal· right of way can be established without the landowner's explicit permission, merely by his continued non-interference with it.

[2] This is the case not only when laws are altered lawfully but even more when there is some rupture of political order; for then the conflict between •legal claims arising out of the new order and •claims previously established has no theoretical solution; it can only be settled by a rough practical compromise. See chapter 6.3.

time create corresponding expectations. This is contrary to common sense, because we are continually becoming convinced that old laws are unjust (e.g. laws establishing slavery); indeed, this continually recurring conviction seems to be a principal source of change in the laws of a progressive society. . . .

This notion of *natural expectations* is worse than indefinite; the ambiguity of the phrase conceals a fundamental conflict of ideas—one that appears deeper and wider the more we examine it. The word 'natural', as used here, covers and conceals the whole chasm between the actual and the ideal—between what *is* and what *ought to be*. As I noted in I/6.2, the word in its ordinary use conveys the ideas of

(a) what is common rather than exceptional, and

(b) what is original or primitive rather than resulting from later conventions and institutions.

But it is also used to signify, in some combination with one or other of those meanings,

(c) what would exist in an ideal state of society.

And it's easy to see how these different meanings have been run together. By 'Nature' men have really meant *God*, or *God viewed from a certain angle*—God, we may say, as known to us in experience—so when they conceive a state of things better than the actual one they have regarded this **(c)** ideal state not only as being more 'natural' in exhibiting God's purposes more than the actual, but as being **(b)** what God originally created, so that the defects of the actual state of affairs must be due to damage done by men. But if we dismiss **(b)** this latter view as unsupported by historical evidence, it's easier for us to see the contrast and conflict between the other two meanings of 'natural', and the corresponding friction between the two elements of the common notion of justice. From one point of view **(a)** we are disposed to think that the **customary** distribution of rights, goods and privileges—as well as burdens and pains—is natural and just, and ought to be maintained by law (as it usually is); while from another point of view **(c)** we seem to recognise an **ideal** system of rules of distribution that ought to exist even if it never has, and we measure the justice of laws by their conformity to this ideal. The chief problem of political justice is to reconcile these two views.

On what principles is the ideal to be determined? This has been my question since the start of this chapter; but I couldn't satisfactorily discuss it without first distinguishing the two elements of in the common conception of justice—•one taking it to conserve law and custom, •the other tending to reform them. Let us now focus on the second of these. [Sidgwick says that there have been many views about the 'ideal constitution of society' that have embodied political ideals that go far beyond our common notion of *justice*. He won't be concerned with these, he says:] My present question is: Are there any clear principles from which we can work out an ideally just distribution of rights and privileges, burdens and pains, among human beings as such?

4. Many people have held that •a society can't be *just* unless certain 'natural rights' are conceded to all members of the community, and that •positive law, whatever else it does, should at least embody and protect these. But it's hard to find in common sense any definite agreement about what these 'natural rights' are, still less any clear principles from which they can be systematically deduced.

However, one way of systematising these rights and bringing them under one principle has been maintained by influential thinkers; it may now be a bit antiquated, but there's still enough of it around to make it deserve careful examination. I'm referring to this view:

> *Freedom from interference* is really the whole of what human beings can be strictly said to owe to each other,

> originally and apart from contracts. Or anyway the protection of this freedom (including the enforcement of free contracts) is the only proper aim of law. . . .

i.e. of the rules of mutual behaviour that are maintained by penalties inflicted under the authority of government. On this view, all natural rights come down to the right to freedom, so that completely establishing this right would be completely realising justice—interpreting the *equality* that justice is thought to aim as *equality of freedom*.

When I look at this as an abstract formula, though I don't see it as self-evidently the true basic principle of ideal law, I admit that it looks good to my mind; and I might persuade myself that my failure to see it as self-evident is due to some defect in my faculty of moral (or jural) intuition. But when I try to relate it more closely to actual human society, it soon starts to look different. ·I shall present four of its difficulties·.

(i) Obviously we would have to limit the extent of its application. It involves the negative principle that no-one should be coerced for his own good alone; but no-one would seriously maintain this in connection with children, or idiots, or insane persons. But in that case we can't know *a priori* that it ought to be applied to all sane adults. The acknowledged exceptions are usually justified on the ground that children etc. will obviously be better off if they're forced to behave as others think best for them; and it's not intuitively certain that the same argument doesn't hold for the majority of mankind in the present state of their intellectual progress. Indeed, many advocates of this principle concede that it doesn't hold for adults in a low state of civilisation. But then what's the criterion for where it applies? It can only be this: the freedom principle applies wherever human beings are intelligent enough to provide for themselves better than others would provide for them! And that presents the principle not as absolute ·and basic· but merely as a consequence of the wider ·and deeper· principle of aiming at the general happiness or well-being of mankind.

(ii) The term 'freedom' is ambiguous. If we interpret it strictly as meaning only freedom of action, the principle seems to allow any amount of mutual annoyance except constraint. But obviously no-one would be satisfied with such freedom as this. But if we include in the idea of *freedom* absence of pain and annoyance inflicted by others, it's immediately obvious that we can't prohibit all such annoyances without intolerably restraining freedom of action, because almost any gratification of a man's natural impulses might cause annoyance to others. So we would have to distinguish the mutual annoyances that ought to be allowed from those that must be prohibited, which would force us to balance •the evils of constraint against •pain and loss of other kinds; ·this is a utilitarian consideration·; and if we admit the utilitarian criterion far enough to do *this* work, it's hard to maintain that annoyance to individuals is to be permitted only to prevent more serious annoyance and never to attain any positive good result.

(iii) If a social construction is to be possible on this basis, the right to •freedom will have to include the right to •limit one's freedom by contract; and if such a contract is really voluntary and not obtained by fraud or force, and if it doesn't violate anyone else's freedom, it will have to be enforced by legal penalties. But I can't see that enforcement of contracts is strictly included in the notion of freedom; for a man seems to be most completely free when no one of his volitions is allowed to have any effect in causing the *external* coercion of any other. Again, if there are no limits to this right to limit one's own freedom, a man could freely contract himself out of freedom into slavery, so that the principle of freedom would be suicidal! But it seems clearly impossible to derive from the principle of freedom any limits to the right of limiting one's

freedom by contract.[1] [In this paragraph Sidgwick has identified two things that are needed in a régime of freedom. He points out that they can't be extracted from the notion of *freedom*, and the unstated conclusion he is drawing is that they'll have to be defended on a broadly utilitarian basis.]

(iv) So we see that it's hard to define freedom as an ideal to be realised in the personal relations between human beings, but it's even harder when we consider the relation of men to the material means of life and happiness. It is commonly thought that the individual's right to freedom includes the right to appropriate—·i.e. become owner of·—material things. But I can't see that 'freedom', taken strictly, implies more than a man's right to non-interference when actually using things that can be used only by one person at a time; x's right to prevent y from ever using something that x has once seized—·which is what it means for x to 'own' that thing·—seems to interfere with y's free action further than is needed to secure x's freedom, strictly speaking. You may say: 'When a man appropriates something he doesn't interfere with the freedom of others, because the rest of the world is still open to them.' But others may want exactly the thing he has appropriated; and they may be unable to find anything as good—or at least not without much labour and search—because many of the means and materials of comfortable living are in short supply. This argument applies especially to the ownership of *land*. ·It is sometimes said that· a man is allowed to own land of which he is the 'first occupier'; but how do we decide how much land a man 'occupies'? You might say 'He occupies whatever he is able to use'; but the use of land by any individual can vary almost indefinitely in extent, while diminishing in intensity. It would surely be paradoxical to derive from the principle of freedom that an individual has a right to exclude others from pasturing sheep on any part of the land over which his hunting expeditions *could* extend.[2] But is it *clear* that a shepherd has such a right against someone who wants to plough the land, or that someone who is using the surface has a right to exclude a would-be miner? I don't see how the derivation can be made out. Another point: if the right of property is derived from the principle of freedom, does it include the right to control what happens to one's possessions after death? Most people think of this as naturally tied to ownership: but it's paradoxical to say that anything we do with his possessions after his death could interfere with his freedom of action! Legal scholars have often treated this right ·of post-mortem control· as conventional and not as part of 'natural law'.

There are other difficulties too, but we needn't pursue them, because if freedom simply means that one man's actions are to be as little as possible restrained by others, there's obviously more freedom if there is no ownership. You might say:

> 'Freedom is more than mere absence of restraint; it also includes facility and security in the gratification of desires. That is the freedom that we think should be equally distributed, and it can't be achieved without ownership.'

Then I reply that in a society where most material things are already owned, this kind of freedom can't be equally distributed. A man born into such a society without inheritance is not only less free than those who have property but

[1] The question of how far the conception of freedom involves unlimited right to limit freedom by free contract will come up again in chapter 6, when I discuss the general duty of obedience to law.

[2] It has often been urged as a justification for dispossessing savages of the land of new colonies that tribes of hunters have no moral right to property in the soil over which they hunt.

also less free than he would have been if there had been no ownership. Optimistic political economists have said that

> •because this man has freedom of contract, he will give his services in exchange for the means of satisfying his wants, that
> •this exchange must give him more than he could have had if he had been alone in the world, and indeed that
> •every human society makes •the part of the earth that it inhabits better able to gratify the desires of all its posterity than •it would otherwise be.

However true this may be as a general rule, it's obviously not so in all cases—sometimes men can't sell their services at all, and often they can't sell them for enough to live on. And even if the optimists' view were true, it wouldn't show that ownership hasn't enabled society to interfere with its poorer members' natural freedom; all it would show is that society compensates them for such interference, and that the compensation is adequate; and it's surely *obvious* the achievement of freedom can't be the one ultimate end of distributive justice if encroachments on freedom can be justly compensated for by food or money!

5. So it seems that although freedom is a keenly desired and important source of happiness. . . ., the attempt to make it the basic notion of theoretical jurisprudence doesn't succeed; and even the 'natural rights' that it claims to cover can't be brought under it except in a very forced and arbitrary manner. And even if none of this were right, there is *more* to our notion of justice than an equal distribution of freedom. Ideal justice, as we commonly conceive it, seems to demand that not only freedom but all other benefits and burdens should be distributed equally (or at any rate *justly*: that's not the same as *equally*, but merely excludes arbitrary inequality).

How then shall we find the principle of this highest and most comprehensive ideal, ·the ideal of justice·?

We'll be led to it by returning to one of chapter 4's grounds of obligation to render services, namely the claim of gratitude. It seemed there that along with our natural •impulse to requite benefits we •think that such requital is a duty and its omission blameworthy—though we can't pin down *how* blameworthy. Now if we (so to speak) *universalise* this •impulse and •conviction, we get the element in the common view of justice that we're now trying to define. Let us take the proposition

> **(i)** 'Good done to any individual ought to be requited by him'

and leave out both references to the individual, we get the more general proposition

> **(ii)** 'Good deeds ought to be requited'

which we are equally sure is true.[1] And if we take into consideration all the different kinds and degrees of services that go into the workings of society is based, we get the proposition

> **(iii)** 'Men ought to be rewarded in proportion to their deserts'.

And this would be commonly held to be the true and simple principle of distribution in any case where there are no counter-claims arising from contracts or customs.

For example, it would be admitted that if no other arrangement has been made the profits of any enterprise should be divided among those who have contributed to it, in proportion to the worth of their contribution. Some thinkers maintain the section 4 proposition that

[1] This strikingly illustrates the difference between •natural instincts and •moral intuitions. The •instinctive impulse to requite a service is on its emotional side quite different from the •intuition that (for example) a fair day's work deserves a fair day's wages. Still, our sense of the *duty* of gratitude seems to fall under the more general intuition that desert ought to be requited.

(iv) 'Law should try to secure the greatest possible freedom for each individual'

not as an absolute axiom but as following from the principle that desert ought to be requited; because the best way of providing for the requital of desert is to leave men as free as possible to work for the satisfaction of their own desires, so that each wins his own requital. And it seems that this *is* the principle that is at work when people ·purport to· justify the right of property on the ground that

(v) 'Everyone has an exclusive right to whatever is produced by his labour'.

Think about it: no labour really 'produces' any material thing, but only adds to its value; and we don't think that x can acquire a right to a material thing that belongs to y by working on it, even if he genuinely believes that *he* owns it; all we think he is entitled to is adequate *compensation* for his labour; so this—namely the proposition

(vi) ·'Every man ought to receive adequate requital for his labour'·

—must be what proposition **(v)** means. The principle is sometimes *stretched* to cover the original right of property in materials, on the ground that someone who discovers something thereby 'produces' it;[1] but here again we can see that common sense doesn't grant this (as a *moral* right) absolutely, but only to the extent that it seems to be adequate—and not more than adequate—compensation for the discoverer's trouble. We wouldn't think that the first finder of a large uninhabited region became the rightful owner of the whole of it! So this justification of the right of property eventually takes us back to proposition **(vi)**. Also, when we speak of the world as 'justly governed' by God, we seem to mean that. . . .happiness is distributed among men according to their deserts; and divine justice is thought to be a pattern that human justice should imitate as far as it can.

This kind of justice (I repeat) seems like gratitude universalised; and the same principle applied to punishment can be seen as resentment universalised; although in the present state of our moral conceptions the parallel is incomplete. History shows us a time when it was thought that a man *ought* to requite injuries, just as he *ought* to repay benefits; but as moral reflection developed in Europe this view was rejected, so that Plato taught that it could never be right really to harm anyone, however he may have harmed us. And this is ·now· the accepted doctrine in Christian societies, as regards requital by individuals of personal wrongs. But the universalised form of the old conviction still lingers in the popular view of criminal justice: it seems still to be widely held that

> Justice requires •pain to be inflicted on a wrongdoer, even if •it brings no benefit to him or to anyone else.

I don't myself hold this view; indeed I have an instinctive and strong moral aversion to it. I hesitate to attribute it to common sense, because I think it is gradually passing away from the moral consciousness of educated folk in the most advanced communities; but I think it may be the more ordinary view.

So this is one element of what Aristotle calls 'corrective justice', which is embodied in criminal law. Don't confuse it with the principle of reparation, on which legal awards of damages are based. The duty of reparation, we have seen,

[1] It requires a lot of strain to bring the 'right of first discovery' under the notion of 'right to what is produced by one's labour'. Locke and others had to justify the right of first discovery by supposing that mankind have 'tacitly consented' to the principle that anything that isn't owned shall come to be owned by the first person who takes possession of it. But this is a rather desperate device of ethico-political construction, because it can so easily be used to justify almost any arbitrariness in positive law.

follows simply from the maxim of general benevolence that forbids us to harm our fellow-creatures: if we have harmed someone we can approximate to obeying the maxim by giving compensation for the harm. [Sidgwick here devotes a page to the question of whether the duty to make reparation depends on whether one was at fault in doing the harm. He thinks we would condemn a man who didn't offer some reparation for serious harm he has caused, even if he was in no way at fault for it; but suggests that this may be a matter of benevolence rather than justice. He also remarks that some kinds of harm can't be compensated, and that in other cases compensation is hard to calculate because the harm isn't measurable in terms that would apply to any possible compensation. This kicked off from the difference between punishment and compensation, to which Sidgwick now returns:] These days there's no danger of confusion or collision between the principles of reparative and of retributive justice, because obviously one is concerned with the claims of the injured person and the other with the deserts of the wrongdoer; though the obligation to paying compensation may sometimes in practice be treated as a sufficient punishment for the wrongdoer.

[Sidgwick will now use 'retributive' in a rare sense in which it covers reward as well as punishment. He will do this only once more, on page 169.] But when we turn back to the other branch of retributive justice, which is concerned with rewarding services, we find another notion—I'll call it 'fitness'—which needs to be carefully separated from desert just because they are so often run together. When they are properly distinguished, we'll see that fitness, ·so far from being equivalent to desert·, is liable to collide with it.[1] I'm not sure that the principle of 'distribution according to fitness' is contained in our ordinary notion of •justice, but it certainly enters into our common conception of •the ideal or perfectly rational order of society; I'm talking about the distribution of •instruments and •functions and (to some extent at least) of •other sources of happiness. We think it reasonable for

- •instruments to be given to those who can use them best,
- •functions to be assigned to those who are most competent to perform them, and
- •particular material means of enjoyment to go to those who are susceptible to the relevant kinds of pleasure.

Regarding the third of these, no-one would think of allotting pictures to a blind man, or rare wines to someone who had no taste; so we would probably think it fitting that artists should do better in the social distribution of wealth than mechanics. In none of these cases is it guaranteed that the recipients are the most *deserving*. So the notions of desert and fitness at least sometimes conflict; but perhaps (I repeat) fitness doesn't come into the interpretation of justice proper, and is really just a utilitarian principle of distribution that inevitably limits how far actual arrangements can satisfy the demands of abstract justic e. And our present concern is with abstract justice. Anyway, taking ideal justice to involve more than mere equality and impartiality, its chief other ingredient concerns the requital of desert. Let us then examine more closely what desert consists in, beginning with •good desert, i.e. •merit, because this is more basic and permanently important ·than bad desert·. Why more permanent? Because we can hope that as the world improves crime and punishment will decrease and gradually disappear; but the right or best distribution of the means of well-being is something we will always be trying to achieve.

[1] I think the term 'merit' often blends the two notions, as when we speak of 'promotion by merit'. But moralists generally use 'merit' as exactly equivalent to what I have called 'desert'.

6. The first question is one that also arose over gratitude: should a reward be proportional to the •effort made or to the •results achieved? ·In favour of the former·, it may be said that the actual value of any service will largely depend on favourable circumstances and lucky events that owe nothing to the agent's desert; or on powers and skills that the agent was born with or have been developed by favourable conditions of life or by good education; and why should we reward him for these?. . . . And certainly it is commonly thought that God will reward only the *moral* excellences that are exhibited in human actions. But that doesn't yet get rid of the difficulty; for it can still be said that ·morally· good *actions* are largely or entirely due to good dispositions and habits that have been inherited or are upshots of the care of parents and teachers, so that in rewarding them we are rewarding the results of natural and accidental advantages; and it's unreasonable to distinguish these from skill and knowledge etc. and to say that it is just to reward the one and not the other. Shall we say, then, that the reward should be proportionate to the amount of voluntary effort for a good end? But determinists will say that even *effort* results from causes extraneous to the man's self. On the determinist view, it would seem to be ideally just—if anything is just—that all men should enjoy equal amounts of happiness; for there seems to be no justice in making A •happier than B merely because circumstances beyond his own control have first made him •better. But why shouldn't we instead of 'all men' say 'all sentient beings'? Why should men have more happiness than any other animal? But now •the pursuit of ideal justice seems to be leading us to such a precipice of paradox that common sense is likely to abandon •it. At any rate the ordinary idea of desert has vanished.[1] So we seem to be led to the conclusion that I anticipated in I/5, namely that in this part of our moral consciousness the idea of free will seems to be involved in a special way in the moral ideas of common sense, because if free will is eliminated the important notions of desert (or merit) and justice require substantial modification.[2] And the difference between determinist and libertarian justice can hardly have any practical effect, because we can never separate the part of a man's achievement that is due strictly to his free choice from the part that is due to the original gift of nature and to favouring circumstances;[3] so that we can only •leave it to providence to bring about what we conceive as the theoretical ideal of justice, and •settle for trying to reward voluntary actions in proportion to the worth of the services

[1] The only tenable determinist interpretation of desert, I think, is the utilitarian one, according to which 'He deserves a reward for his services to society' means merely that it is expedient to reward him so that he and others may be induced to act similarly in the expectation of similar rewards. See IV/3.4.

[2] Perhaps it's partly because of these difficulties that some of the utopian reconstructors of society have dropped the notion of *desert* from their ideal, leaving 'equality of happiness' as the only end. Justice, they think, merely requires that each person should have an equal share of happiness, as far as happiness depends on how others act. But it's hard to work with this idea: apart from the issues about fitness mentioned above, equal happiness won't be achieved by equal distribution of objects of desire. It takes more to make some people happy than to make others happy, so we would have to take differences of needs into consideration. But if merely mental needs are included (as seems reasonable) we would have to give less to cheerful, contented, self-sacrificing people than to those who are naturally moody and demanding, because the former can be made happy with less! This is too paradoxical to recommend itself to common sense.

[3] We could lessen the inequalities that are due to circumstances by bringing the best education within the reach of all classes. . . ., and this seems to be prescribed by ideal justice as a way of mitigating arbitrary inequality. . . . But even then there will be much natural inequality that we can't remove or even estimate.

intentionally rendered by them.

The next question is: On what principle or principles can we rationally estimate the comparative worth of different services? We do commonly assume that such estimates can be made, for we speak of the 'fair' or 'proper' price of any kind of services as though this were generally known, and condemn the demand for more than this as extortionate. This may be said:

> The notion of fairness or equity that we ordinarily apply in such judgments is distinct from the notion of justice. Equity is often contrasted with strict justice, and thought of as capable of colliding with it.

This is partly true; but the wider and equally usual sense of 'justice', in which it *includes* equity or fairness, is the right one to adopt in an ethical treatise; for in any case where equity comes into conflict with strict justice, the latter's dictates are held to be •'just' in a higher sense and to be •what ought to be followed in the case in question. . . . So I treat equity—in a slight departure from ordinary usage—as a species of justice. Well then, on what principle can we determine the 'fair' or 'equitable' price of services? In the common judgments of practical persons what is 'fair' is settled by reference to analogy and custom: a service is considered to be 'fairly worth' what is usually given for services of that kind. . . . And in some states of society the payment given for services seems to be as completely fixed by usage as any other customary duty, so that deviating from this pay-scale would be a clear disappointment of normal expectation. But probably no-one in a modern civilised community would maintain a strict equality of 'just price' of services with 'usual price'; and when the judgments of practical persons seem to imply this, they are being superficial or merely inadvertent, and ignoring the established mode of fixing the market prices of commodities by free competition of producers and traders. For where such competition operates the market value varies at different places and times, so that no informed person will. . . .complain of injustice merely because of the variations in it.

Can we then say that 'market value' (as determined by free competition) corresponds to our notion of what is ideally just?

This is a question of much interest, because this is obviously how all payment for services would be made in a society constructed on the principle—previously discussed—of providing the greatest possible freedom for all members of the community. Modern civilised communities have tended to approximate (until recently) to this individualistic ideal (as I call it), so it's important for us to know whether it completely satisfies the demands of morality, and whether •freedom, if not an absolute end or first principle of abstract justice, is still to be sought as the best means to producing a just social order by the general requital of desert.

It seems initially plausible to suppose that x's 'market value' represents the estimate set on x by mankind generally, and therefore pins down for us the 'common sense' judgment about values that we are now trying to find. But on reflection it seems likely that most men don't know enough about the nature and effects of many important kinds of services to be qualified to judge their real value; so that for these services the true value won't be represented in the market-place. Even with things that a man can usually estimate, he may be ignorant of such a thing's utility in a particular case; and then the 'free' contract hardly seems fair; though if •the ignorance wasn't caused by the purchaser common sense doesn't condemn him for taking advantage of •it. For instance: a man legitimately using geological knowledge and skill discovers that there's probably a valuable mine on land owned by a stranger; reasonable people wouldn't

blame him for concealing his discovery until he had bought the mine at its market value; but the seller clearly didn't get what the land was really worth. In fact common sense is perplexed on this point; and the conclusion it arrives at must be based on economic considerations that take us well beyond the analysis of the common notion of justice. (See IV/3.4.)

Another point: Some highly important social services have no price in any market, because of the indirectness and uncertainty of their practical utility—scientific discoveries, for example. The extent to which any given discovery will aid industrial invention is so uncertain that even if the secret of it could be conveniently kept, it usually wouldn't be profitable to buy it.

But even with products and services that are generally marketable, and with bargains thoroughly understood on both sides, there are still problems about the thesis that a 'fair' price is simply the one set by the item's market value. When someone has a monopoly of a certain kind of services, the market-price of the aggregate of such services can sometimes be increased by diminishing their total amount; but it would seem absurd to say that those who provide the services thereby come to *deserve* more, and a plain man wouldn't accept that the price fixed in this way is fair. Still less is it thought fair to take advantage of the temporary monopoly produced by an emergency: if I saw Croesus drowning and no-one near, it wouldn't be regarded as fair for me to refuse to save him except at the price of half his wealth. But then can it be fair for *any* group of people to get a competitive advantage from the unfavourable economic situation of another group? And if we say No, where should we draw the line? Any increase in the numbers of a group makes its bargaining position less favourable, because the market price of a service depends partly on how easy or hard it is to get it—as economists say, 'on the relation between the supply of services and the demand for them'—and it doesn't seem that any individual's social *desert* can be lessened merely by an increase in the number or willingness of others offering the same services. Nor indeed does it seem that his desert can be decreased by his own willingness, for it would be strange to reward a man less because he is eager to do his work. Yet in bargaining, the less willing seller always has the advantage. And, finally, the social worth of a man's service isn't automatically increased by the fact that his clients can pay lavishly; but his reward is likely to be greater from this cause.

Considerations like these have led some political thinkers to hold •that a *just* system of paying for services would be nothing like the present system of free competition, and •that all labourers should be paid according to the intrinsic value of their labour as estimated by enlightened and competent judges. We might call this the 'socialistic ideal'; if it could be achieved without counter-balancing evils it would certainly seem to come closer than the present state of society does to what we conceive as divine justice. But it requires a rational method of determining value, and we are still looking for that. . . . For one thing, how are we to compare the values of different services that have to be combined to produce a happy life? How, for example, should we compare the respective values of necessities and luxuries? We may be more aware of the enjoyment we get from luxuries, but we couldn't have any enjoyment without the necessities. Also, when different kinds of labour go into producing something, how should we estimate their relative values? Even if all mere •unskilled labour is paid at a single standard rate, we can't do this with the different kinds of •skill. How are we to compare the labour of design with the labour of production? or the supervision of the whole job with the carrying out

of the details? or the labour of actually producing with that of training producers? or the service of the scientist who discovers a new principle with that of the inventor who applies it?

I don't see how any of these difficulties can be solved by any analysis of our common notion of justice. To deal with such matters satisfactorily we have to come at them differently, not asking •what services of this or that kind are intrinsically worth, but •what reward can procure them and •whether the rest of society gains more by the services than it loses by the reward. We have, in short, to give up as impracticable the construction of an ideally just social order in which all services are rewarded in exact proportion to their intrinsic value. And for similar reasons we're forced to the more general conclusion that it's impossible to devise a clear and reasoned method of exactly determining different amounts of good desert; and common sense may go along with this. Although common sense thinks that ideal justice consist in rewarding desert, it regards as utopian any general attempt to achieve this ideal in the social distribution of the means of happiness. In the actual state of society attempts to reward good desert are made only within a very limited range—parents with their children, the state with deserving statesmen, soldiers, etc.—and if you think about these cases you'll see how rough and imperfect are the standards used in deciding how much reward is due. The only kind of justice that we try to realise [see Glossary] most of the time is concerned with the fulfilment of contracts and definite expectations. We leave the general fairness of *distribution by bargaining* to take care of itself.

7. Criminal justice presents difficulties corresponding to the ones I have been discussing. One similarity: the partial confusion of the ideas of •law and •justice: by 'bringing a man to justice' we commonly mean 'inflicting legal punishment' on him; and we think that the penalty prescribed by law should be inflicted—neither more nor less—even if we regard the legal scale of punishment as unjust. One dissimilarity: there's no criminal-justice analogue of the civil-justice perplexity about changes in the law. We don't think that a man can acquire, by custom, prescriptive rights to over-lenient punishment, as he is thought to do to an unequal distribution of liberties and privileges. If now we investigate the ideal of criminal justice as intuitively determined, we certainly find that in so far as punishment isn't regarded as merely preventive it is commonly thought that it ought to be proportioned to the gravity of crime.[1] Still, when we try to make the method of apportionment perfectly rational and precise, the difficulties seem at least as great as in the case of good desert. **(i)** The assumption of free will seems to force its way in again: if a man's bad deeds are entirely caused by nature and circumstances, it certainly seems that Robert Owen was right: the man doesn't properly deserve to be punished for those actions, and justice requires us rather to try to alter the conditions under which he acts. And we actually *do* punish deliberate offences more than impulsive ones, perhaps as implying a more *free* choice of evil. **(ii)** We think that if an offender has had no moral training, or a perverted training, that makes him less criminal—but it's commonly agreed that he can't be let off from all punishment on this account. **(ii)** The moral gravity of a crime seems to be

[1] Those who hold that the essence of justice consists in securing external freedom among the members of a community, and that punishment is justified only as a means to this end, naturally think that in awarding punishment we ought to consider merely its efficacy as such means. But this isn't an interpretation of the common notion of just punishment. The utilitarian view of punishment is becoming more prevalent, but it hasn't yet prevailed.

reduced if the motive is praiseworthy, as when a man •kills a villain whose crimes elude legal punishment, or •heads a hopeless rebellion for the good of his country. But it would be paradoxical to say that we ought to reduce punishment proportionally; common sense would hold that—whatever God may do—*men* must generally inflict severe punishment for any gravely harmful intentional act that is forbidden by law, even if it was prompted by a good motive.

Set aside the motive and look only at the intention: it's still hard to state clear principles for determining the gravity of crimes. [For the distinction between motive and intention see chapter 1.2 .] Consider the case of the patriotic rebel: the intention of this criminal is to do what is right and good. And in many cases although an offender knows he is doing wrong, he may not intend to harm any sentient being—e.g. he steals something that he thinks won't be missed. Also, we don't commonly think that a crime is made less serious by being kept perfectly secret, yet much of the harm done by a crime is (in Bentham's phrase) the 'secondary evil' of the alarm and insecurity that it causes, and this part is cut off by complete secrecy. You may reply: 'This last difficulty isn't a *practical* one, because we aren't called on to punish a crime until it has been discovered, and by then the secondary evil has happened and is all the greater because of the previous secrecy.' But it remains true that the criminal didn't intend his crime to be discovered, so he didn't intend that part of the evil that was caused by the crime. And if we say that

> the awfulness of the crime depends on the loss of happiness that such acts would generally cause if they went unpunished, and we must suppose the criminal to be aware of this,

we seem to be trying to force a utilitarian theory into an intuitional form by means of a legal fiction.

I have been discussing intentional wrong-doing; but positive [see Glossary] law awards punishment also for harm that is due to rashness or negligence; and we run into further difficulties when we try to justify this. Some legal theorists seem to regard rashness and negligence as positive states of mind, in which the agent consciously refuses the attention or reflection that he knows he ought to give; and no doubt this sort of willful recklessness does sometimes occur, and seems as properly punishable as if the resulting harm had been positively intended. But in practice the law doesn't require evidence that this was the agent's state of mind (which indeed usually couldn't be given); it settles for proof that the harm would have been prevented if the agent had shown as much care as an average man would have shown in those circumstances. And by 'carelessness' we usually mean merely a purely negative psychological fact, i.e. that the agent *did not* go through certain processes of observation or reflection; so that the action was strictly involuntary, and so hardly seems to involve ill-desert. You might say that although the present carelessness isn't blameworthy the past neglect to develop habits of care is so. But often we can't infer even this past neglect; and in such cases the only basis for punishment involves the utilitarian theory of punishment, which regards it as a means of preventing similar harmful acts in the future. Similar difficulties arise—as I hinted on page 134—in fixing the limits within which reparation is due—i.e. on the view that we aren't obliged to compensate for all harm caused by our bodily movements but only for harm that is intentional or due to our rashness or negligence.

The results of this examination of justice can be summed up as follows. The prominent element in *justice* as ordinarily conceived is a kind of equality—i.e. impartiality in the application of certain general rules allotting good or evil to individuals. But when we have clearly picked out this element, we see that more is needed if we're to have an

account of the virtue of *justice* that can guide us in our conduct. Still looking for the right general principles of distribution, we find that our common notion of justice includes besides the principle of reparation for injury two quite distinct and divergent elements.

Conservative justice is realised **(1)** in the observance of law and contracts and definite understandings, and in the enforcement of legal penalties for the violation of these; and **(2)** in the fulfilment of natural and normal expectations. This latter obligation is somewhat indefinite.

Ideal justice is even harder to define, for there seem to be two quite distinct conceptions of it, embodied in two distinct ideals of a political community:

•The individualistic ideal takes the realisation of freedom as the ultimate end and standard of right social relations; but it turns out that the notion of freedom won't give a practicable basis for social construction without certain arbitrary[1] definitions and limitations; and even with these, a society in which freedom is realised as far as is feasible doesn't completely suit our sense of justice.

•The socialistic ideal of distribution—based on the principle of requiting desert—is *prima facie* closer to our sense of justice; but when we try to make this principle precise we again find ourselves in grave difficulties; and similar perplexities beset the working out of rules of criminal justice on the same principle.

Chapter 6: Laws and promises

1. The moral obligations of •obedience to law and •observance of contracts have appeared to be the most definite part of the complex system of private duties commonly labelled as 'justice'. But we have also seen that there are some laws the violation of which •doesn't interfere with the rights of others and therefore •doesn't look like injustice. Also, the duty to keep one's promises is also commonly conceived as independent of any issue of harm to the promisee: men ordinarily judge that promises to the dead ought to be kept, though the dead are out of harm's way. . . . So it seems desirable to examine the propositions

•Law ought to be obeyed, and
•Promises ought to be kept,

considered as independent principles.

How are we to ascertain what the Law is that we are commonly thought to be morally bound to obey? In II/5.2 I distinguished legal rules from others in terms of the punishments inflicted on violators of them; but that won't help us here, because commands issued by rebels and usurpers are not regarded as binding though they may be enforced by judicial penalties. It would be generally agreed that any duty we have to obey *these* commands come from the harm that

[1] By 'arbitrary' I mean definitions and limitations that destroy the principle's self-evidence, and when closely examined lead us to regard it ·not as basic but· as subordinate.

may come to us or others if we disobey, so that the extent of such a duty is determined by considerations of expediency. And even the commands of a legitimate sovereign don't count as 'laws' in the sense of the word involved in the proposition that *laws ought to be obeyed*, because we all recognise that a rightful sovereign may command his subjects to do something wrong, and their duty then is to disobey him. For our present purpose, then, we must define *laws* to be *rules of conduct laid down by a rightful authority which is commanding within the limits of its authority*.

If we are to be practically guided by the proposition that laws ought to be obeyed, we need to know **(i)** how to distinguish a rightful lawmaker, whether individual or corporate, and **(ii)** how to ascertain the limits of this lawmaker's authority. They are distinct questions, but we'll see that they can only partially be separated. Starting with **(i)**: we can assume that the authority to make laws is held by some living man or men. [This may not have held for some primitive societies, Sidgwick remarks, but it is accepted by] the common sense of civilised Europe, which is our present topic. We don't think that any of the definite prescriptions of positive law have an origin that puts them beyond the reach of alteration by any living authority.

Then where is this authority to be found?

In the usual answers to this question, the conflict between the •ideal and the •traditional or customary, which perplexed us in our attempt to define *justice*, now appears in an even more complicated form. Some say that **(b)** we ought always to obey the **traditionally legitimate** authority in our country. Others maintain that **(a)** a nation is entitled to demand—even at the risk of civil strife and bloodshed—that an **ideally legitimate** authority be established, i.e. one constituted in accordance with certain abstract principles. And often **(c)** the **actually established** authority doesn't have either kind of legitimacy. So that we have to distinguish three claims to authority: **(a)** that of the government held to be ideally or abstractly right, one that *ought* to be established; **(b)** that of *de jure* government, ·i.e. government legitimised· by the constitutional traditions of the country in question; and **(c)** that of the *de facto* government.

2. Let us start by considering **(a)** the ideal. Of all the views that theorists have put forward regarding the right constitution of supreme authority, I'll consider only those that have a *prima facie* claim to express mankind's common sense about this. The most important and most widely accepted of these is the principle that

> the sovereign in any community can only be rightly constituted by the subjects' consent.

I noted on page 131 that this is involved in the adoption of freedom as the ultimate end of political order; if a man basically owes nothing to anyone else except non-interference, he clearly ought to become a subject of a sovereign only by his own consent. Thus, to reconcile the basic right of freedom with the actual duty of law-observance, we have to suppose a social contract so that obedience to law becomes merely a special case of the duty of keeping compacts.

How are we to know the terms of this basic compact? No-one now accepts this old view:

> The transition from the 'natural' to the 'political' state occurred by means of a 'basic contract' that made some particular form of social organisation indelibly legitimate.

Well, then, we might try this:

> By remaining a member of a community a man 'tacitly undertakes' to obey the laws and commands of the authority recognised as lawful in that community.

But this reduces **(a)** the ideal to **(b)** the customary: it would allow the most unlimited despotism, if established

and traditional, to claim to rest on 'free consent'; so that the principle of •abstract freedom would support the most absolute •concrete tyranny and servitude; and thus the theory would strengthen men's chains under pretence of increasing their freedom. To avoid that result, we might try this:

> Some natural rights are inalienable—or are tacitly reserved in the tacit compact—and no law is legitimate if it deprives a man of these;

but we again have the problem deducing these 'inalienable rights' from any clear and generally accepted principles. Many think that all such rights can be summed up in the notion of *freedom*; but we've seen that 'freedom' is ambiguous, and especially that the right of private property as commonly recognised can't be clearly deduced from it. And it would certainly be most paradoxical to maintain that the only commands for which a government can legitimately claim obedience are ones protecting the freedom of the governed individuals from interference! A proposed way to avoid this difficulty:

> Constitute the supreme organ of government in such a way that each of its laws will have been consented to personally (or through representatives) by each person who is called on to obey it.

The idea is that a government so constituted—in which everyone 'obeys himself alone', as Rousseau put it—will completely reconcile freedom with order. But how is this result to be achieved? Rousseau thought it could be attained by pure direct democracy, where each individual subordinates his private will to the 'general will' of the sovereign people of which all are equally members. But in practice this 'general will' has to be the will of the majority; and it is paradoxical to say that the freedom and natural rights of a dissentient minority are effectively protected by rule that the oppressors must outnumber the oppressed! Also, if the principle is absolute it ought to apply to all human beings; and if to avoid this absurdity we exclude children, we have to *choose* where to drawn the line; and the exclusion of women, which is often supported even by those who regard the suffrage as a natural right, seems indefensible. And to suppose as some have done that the ideal of 'obeying oneself alone' can be even approximately realised by representative democracy is even more obviously absurd. For a representative assembly is normally chosen only by a part of the nation, and each law is approved only by a part of the assembly; so if a man's only involvement in all this has been to vote *against* one member of that assembly, it would be ridiculous to say that he has assented to a law passed by a majority of the assembly.

Anyway, to lay down absolutely that the laws of any community ought to express the will of the majority of its members seems incompatible with the view—vigorously maintained by Socrates and his most famous disciples—that laws ought to be made by people who understand law-making. Will the majority of a representative assembly be more fit to make laws for their country than any set of experts otherwise selected? Perhaps Yes for some countries at some times, but it isn't self-evident that this will always be the case. Yet surely the Socratic proposition (which is merely an application of the principle (see page 135 'that function should be allotted to the most competent') has as much claim to be considered a primary intuition as the one we have been discussing. Indeed, the age-old controversy between aristocracy and democracy seems to come down to a conflict between those two principles—a conflict that can't be solved as long it remains in the *a priori* region.

3. . . . .Common sense refuses to deliver any clear and certain intuitions about the principles on which an ideal constitution should be constructed. And there's no agreement, either, on

the intrinsic lawfulness of violating of the established order of a community by introducing an ideal constitution. Some think that a nation has a natural right to an approximately ideal government, a right that may be maintained by force. Others hold that although the ideal political structure may rightly be proposed, commended, and worked for by any means the established government permits, rebellion for this purpose is never justifiable. And yet others—perhaps the majority—would decide the question by weighing the advantages of improvement against the evils of disorder.

Furthermore, we have seen that it's not so easy to say what the established government *is*. When a legally illegitimate authority issues ordinances and controls the administration of justice, how far is obedience due to it? Everyone agrees that usurpation ought to be resisted; but there's no agreement about the right behaviour towards an established government born of a successful usurpation. Views that have been held about such a government:

(1) It should be regarded as legitimate as soon as it is firmly established.

(2) It ought to be obeyed at once, but under protest, with the purpose of renewing the conflict when conditions are right.

(3) The right attitude at first is that of **(2)**, but as a usurping government becomes firmly established it gradually loses its illegitimacy, so that eventually it's as criminal •to rebel against it as it was originally •to establish it.

Of these, **(3)** seems to be the view of common sense; but the question of *where* the metamorphosis happens has to be answered on the basis of expediency.

The fundamental problems about the legitimacy of authority are relatively simple—as they are in the foregoing discussion—in the case of an *absolute* government where customary obedience is unconditionally due to one or more persons. In a *constitutionally* governed state other moral disagreements arise. It's agreed that in such a state the sovereign is morally bound to conform to the constitution, there's dispute about whether the subjects' obligation to obedience should be conditional on this conformity—whether they have the moral right •to refuse to obey an unconstitutional command and even •to punish the sovereign for violating the constitution by rebelling against him. And there's much perplexity and disagreement about what the constitutional obligations really are—about what the relevant historical facts are and about how they should be understood. Quite often the limitations of sovereign authority embodied in a constitution began as concessions extorted by fear from a sovereign who was previously absolute; and there's a question about •how far such concessions are morally binding on that sovereign and still more about •how far they are binding on succeeding sovereigns. Or *vice versa*: a people may have allowed some of their liberties to fall into disuse, and there's a question about whether it retains the right of reclaiming them. And generally when a constitutional rule has to be gathered from a comparison of precedents, there can be dispute over whether a particular act of either party is a •constitutive precedent or an •illegitimate encroachment. That is why in constitutional countries men's view of what their constitution traditionally *is* has often been influenced by their view of what it ideally *ought to be*; in fact, the two questions have rarely been kept quite distinct.

4. Even where we can get clear about what authority is owed obedience, further difficulties arise when we try to define the limits of such obedience. In modern societies everyone agrees that any authority that commands immoral acts ought to be disobeyed; but this is one of those tautological [see Glossary] propositions, so common in popular morality, that convey

no real information. ·It is empty, because it boils down to saying that immoral acts ought not to be performed.· The right question to ask is:

> What acts are there that *remain* immoral after being commanded by a rightful authority?

There's no clear principle on which to base an answer. It has sometimes been said that the law can't override definite duties; but the obligation of fidelity to contracts is perfectly definite, yet we don't think it's right to fulfill a contract if, since it was made, a law is passed that forbids the conduct that would be needed to keep the contract. And there's practical disagreement on this question among people who wouldn't knowingly adopt the utilitarian method of answering it by value-balancing the different outcomes. Some say that the duties of family relations must yield to the duty of obedience to the law, so that a son ought not to aid a parent, actively or passively, in escaping punishment for crime; while others would consider this rule too inhuman to be imposed, and yet others would draw the line between helping and not-hindering. When a rightly constituted government commands acts that are unjust and oppressive to others, common sense recoils from saying either that •all such commands ought to be obeyed or that •all ought to be disobeyed; but apart from utilitarian considerations I can't find any clear accepted principle for distinguishing the unjust commands of a legitimate government that ought to be obeyed from those that ought not to be obeyed. And then some legal theorists hold that we're not strictly bound to obey laws commanding something that isn't otherwise a duty, or forbidding something that isn't otherwise a sin; on the ground that in the case of duties prescribed only by positive [see Glossary] laws the alternatives of •obeying or •accepting the penalty are morally open to us. [Sidgwick here has a footnote quoting a passage from William Blackstone to that effect.] But others think this principle is too lax; and certainly if any particular law met with a widespread preference for penalty over obedience, the law would be thought to have failed. On the other hand, there seems to be no agreement about whether one is bound to submit to unjust penalties.

In the face of all this difference of opinion, it seems idle to maintain that there's a clear and precise first principle of order that the common reason and conscience of mankind sees intuitively to be true. No doubt there's a vague general habit of obedience to laws (even bad ones), which can fairly claim the universal consensus of civilised society; but when we try to state an explicit •principle corresponding to this •habit, the consensus seems to vanish and we are drawn into controversies that seem to have no solution except what the utilitarian method offers. (I haven't thought it worthwhile to enter into the special difficulties of international law.)

5. My next topic is good faith, i.e. fidelity to promises. This is a natural place to discuss it because some thinkers have based the duty of •law-observance on a prior duty of •fulfilling a contract. The social contract that we have been considering is at best merely a convenient fiction—a device that lets us neatly express the mutual jural [see Glossary] relations of the members of a civilised community. Such a fiction is out of place in an account of the ethical principles of common sense. But *historically* the duty of law-observance has often been closely linked with the duty of good faith. Much constitutional law in certain ages and countries has been established or confirmed by explicit compacts in which different sections of the community agree on certain rules for the future government. The duty of •obeying these rules thus presents itself as a duty of •fidelity to compact. This is even more the case when it's a matter of imposing not a law but a law-giver whose authority is strengthened by an oath of allegiance from his subjects or a representative portion of

them. But even in those cases it's a palpable fiction that the citizens generally are bound by an agreement that only a few of them have actually entered into.

Some moralists have classified the duty of •keeping promises with •veracity, or even identified the two. There is a certain analogy between the two: we fulfill the obligations of veracity and of good faith alike by bringing it about that words correspond with facts—in fidelity by making fact correspond with statement, and in veracity by making statement correspond with fact. But the analogy is obviously superficial and imperfect; we aren't bound to make our actions correspond with all of our assertions but only with our promises. If I merely *assert my intention* of abstaining from alcohol for a year, and then after a week take some, I am (at worst) ridiculed as inconsistent; but if I *promise* to abstain, I am blamed as untrustworthy. The essence of the duty of good faith, then, is conformity not to •my own statement but to •expectations that I have intentionally raised in others.

When a promise has been understood in a sense not intended by the promiser, is he bound to satisfy expectations that he did not voluntarily create? Common sense says Yes in some cases—ones where the expectation is one that most men would form under those circumstances. But this seems to be one of the indefinite duties of •justice, and not properly of •good faith, because strictly speaking no promise has been made. The normal effect of language is to convey the speaker's meaning to the hearer,. . . .and we always suppose this to have taken place when we speak of a promise. If for some reason this normal effect doesn't happen, we can say that no promise—or no perfect promise—has been made.

So the moral obligation of a promise is perfectly constituted when both parties understand it in the same sense. We use 'promise' to cover not only •words but •all signs and even •tacit understandings that aren't explicitly signified in any way, if they clearly form a part of the engagement. The promiser is bound to perform what both he and the promisee understood to be undertaken.

6. Then is this obligation intuitively seen to be independent and certain?

It is often said to be so, and perhaps it seems so to unreflective common sense. But reflection turns up a number of qualifications of the principle—some clear and precise, others somewhat indefinite.

First: most thoughtful people would admit that the obligation of a promise can be annulled by the promisee. If he is dead or otherwise incapable of granting release, we have an exceptional case that isn't easy to solve.[1]

Second: a promise to perform an immoral act is thought not to be binding, because the prior obligation not to perform the act is paramount—just as in law a contract to do what one isn't legally free to do is invalid—otherwise one could evade any moral obligation by promising not to fulfill it, which is clearly absurd. . . . The same principle applies to immoral •omissions or •non-actions, though here we have to distinguish different kinds or degrees of obligatoriness in duties, because clearly a promise can sometimes make it obligatory to abstain from doing what it would otherwise have been a duty to do. For example: if I have promised •an undeserving friend to give him all the money I can spare, it becomes my duty not to give money to •a meritorious hospital, though apart from the promise it might have been

[1] Vows to God constitute another exception. Many think that if these are binding, there must be some way to understand God as granting release from them. But discussion of this lies outside the scope of the present work.

my duty to prefer the hospital to the friend. But we have seen the difficulty of defining the limits of strict duty in many cases—e.g. how far ought the promise of aid to a friend override the duty of giving one's children a good education? So the extent to which the obligation of a promise overrides prior obligations becomes in practice somewhat obscure.

7. When we look into the conditions under which promises are made and the consequences of keeping them, we encounter further qualifications of the duty of fidelity to promises—ones that are harder to think about or get agreement about. **First**: there's much dispute over how far promises obtained by 'fraud or force' are binding. . . . Suppose that a promise is made in consequence of a fraudulent statement, but is made quite unconditionally. Probably most people, if they clearly understood that the promise wouldn't have been made if it weren't for the false statement, would regard the promise as not binding. But the false statement may be only one consideration among others, and it may be of any degree of weight; and we probably wouldn't feel justified in breaking a promise because a single fraudulent statement had been a *part* of the inducement to make it—especially if the falsehood wasn't explicitly asserted but only suggested, or if no falsehood came into it but only a concealment of relevant facts. Some kinds of concealment are treated as legitimate by our law: in most contracts of sale, for example, the law adopts the principle of *caveat emptor* [Latin: 'let the buyer beware'], and enforces the contract even if the seller didn't disclose defects in the article sold, unless he somehow produced the belief that it was free from such defects. Still, this doesn't settle the moral question how far a promise is binding if concealment of relevant facts was used to obtain it. And what if an erroneous impression wasn't deliberately produced but was either shared by the promisee or produced unintentionally?. . . . On all these points common sense seems doubtful; and somewhat similar difficulties arise when we try to define the obligation of promises partly obtained by illegal violence and intimidation.

8. **Secondly**: Suppose that a promise has been made freely and fairly but when the time comes to fulfill it circumstances have changed so that the effects of keeping it may be quite other than what was foreseen when the promise was made. Probably everyone would agree that the promisee ought to release the promiser. But it's hard to decide how far the promiser is bound if the promiser refuses to release him. Some would say that he is bound in all cases; others would hold that a considerable change of circumstances removes the obligation, and they might add that all engagements should be *understood* to be binding only if relevant circumstances remain substantially the same. But such an understanding would greatly reduce the theoretical *definiteness* of the duty.

For a different view of this problem, let us return to the case of promises made to those who are now dead or temporarily out of the reach of communications. In such a case there's no way to get release from the promise, yet keeping it may be really opposed to the wishes—or what would have been the wishes—of both parties. Some say that it is our duty to carry out the 'intention' of the promise, but that is ambiguous. It mat refer to

- **(i)** the meaning that the promisee attached to the words of the promise, as distinct from any other meaning that the common usage of words might allow, or
- **(ii)** something that includes the end-result of the promise's being kept which the promisee had in view in exacting it.

Now we don't commonly think that the promiser is concerned with **(ii)**; he certainly hasn't promised to aim at •the end that the promisee has in view, but only at •some particular means

to it; and if he thinks these means are not conducive to the end, that doesn't—in ordinary cases—release him from his promise. But in the case we are now supposing, where relevant circumstances have changed, and the promise can't be revised, probably most people would say that the promiser should carry out what he sincerely thinks would have been the promisee's intention. But that makes the obligation very vague, because it's hard to tell from a man's wishes in one set of circumstances what he would have wanted in circumstances that were complicatedly different from these; and in practice this view of the obligation of a promise generally leads to great divergence of opinion. So it's not surprising that some hold that even here •the obligation should be interpreted strictly, while others go to the other extreme and maintain that •it ceases altogether!

Thirdly: A promise can't cancel a prior obligation; and most people would agree, as an application of this rule, no promise can make it right to harm anyone. *Anyone*? What about the parties to the promise? It doesn't seem to be commonly held that

> •a man is as strictly bound not to injure himself as he is to avoid harming others;

or, therefore, that

> •a promise is not binding because it was foolish. and will bring pain or burden to the promiser out of proportion to the good done to the promisee.

But in an extreme case, where the sacrifice is very disproportionate to the gain, many conscientious people would think that the promise should be broken rather than kept. Then what about a case where fulfilling the promise may harm the promisee? When we say that it is wrong to harm anyone, we don't commonly mean that it's wrong to bring actual harm to someone even if he thinks it is a benefit; for it seems clearly a crime for me to give someone what I know to be poison, even though he is stubbornly convinced that it is safe. But now suppose that I have promised x to do something which, before I fulfill the promise, I discover is likely to harm him. The circumstances are precisely the same; only my view of them has changed. If x thinks otherwise and calls on me to keep my promise, is it right to obey him? Surely no-one would say this in an extreme case such as that of the poison. But if the rule doesn't hold for an extreme case, where can we draw the line? Common sense gives no clear answer.

9. I have declared that a promise is binding only if it is understood in the same way by both parties; and such an understanding is ordinarily achieved clearly enough, so far as explicit words or signs are concerned. But even here obscurity and misunderstanding sometimes occur; and when it comes to the *tacit* understandings that often enter into promises, a lack of definite agreement is likely enough. So it becomes practically important to decide the question raised on page 128: what duty does a promiser have of satisfying expectations that he didn't intend to create? I called this a duty not so much of •good faith as of •justice, which prescribes the fulfilment of normal expectations. The common way of determining what these are seems to be this: We form the conception of an average or normal man, and consider what expectations *he* would have in the circumstances, inferring this from what men generally expect in similar circumstances. So we appeal to the uses of language and tacit understandings that are customary among people in relations like those of the promiser and promisee ·in the case we are thinking about·. These interpretations and understandings are not obligatory on someone making a promise, but they set •a standard that we presume to be known to all men and accepted by them, except when •it is explicitly rejected. If one of the parties to a promise has deviated from this common standard without explicitly

saying so, we think he ought to suffer any loss resulting from the misunderstanding. This criterion is generally applicable; but it can't be applied if custom is ambiguous or shifting, and then the claims of the parties create a problem that is very hard—if not strictly impossible—to solve. [Sidgwick didn't say '. . . if "custom" is ambiguous or shifting'. He seems to mean 'if there are changes and unclarities in the facts about what is customarily understood and expected'.]

I have been assuming that the promiser can choose his own words, and that if the promisee finds them ambiguous he can get them modified or. . . .explained by the promiser. But in promises made to the community as a condition of obtaining some office or salary, a certain unalterable form of words has to be used. Here the difficulties of moral interpretation are much increased. You might say:

> 'The promise should be interpreted in the sense in which its terms are understood by the community';

and if their usage is uniform and unambiguous this rule of interpretation suffices. But words are often used in different ways—and with different degrees of strictness—by different members of the same society; so it often happens that a promise to the community isn't understood in any one sense; and the question arises 'Is the promiser bound to keep the promise in the sense in which it will be most commonly interpreted, or may he select any of its possible meanings?' And if the formula is fairly old, 'Ought it to be interpreted in the sense that its words now generally have, or in the sense they had when it was first formulated. . . .?' It's hard to get any clear answer to these questions from common sense; and it is made even harder by the fact that there are often strong inducements to make these formal engagements, which cause even reasonably honest people to take them in a strained and unnatural sense. When this happens often enough, a new general understanding grows up about the meaning of the engagements: they are understood in a sense differing indefinitely from their original one, or even—the worst degradation—as 'mere forms'. The question then arises: 'How far, for a conscientious person, can this process of relaxation or perversion modify the moral obligation of the promise?' When the process is complete, we are clearly right in adopting the new understanding as far as good faith is concerned, even if it obviously conflicts with the natural meanings of the words. . . . But the process is usually incomplete because some of the community still understand the engagement in its original strict sense; and then the obligation becomes hard to determine, and the judgments of conscientious people about it become divergent and perplexed.

To sum up the results of the discussion, it seems that there's a clear consensus only for the principle that a promise, explicit or tacit, is binding if

- the promiser has a clear belief about the sense in which the promise is understood by the promisee,
- the promisee is still in a position to grant release from the promise but is unwilling to do so,
- the promise wasn't obtained by force or fraud,
- it doesn't conflict with definite prior obligations,
- the promiser doesn't believe that keeping the promise will harm the promisee or inflict a disproportionate sacrifice on the promiser, and
- relevant circumstances haven't changed since the promise was made.

If any of these conditions fails, the consensus fades away and the common moral perceptions of thoughtful folk fall into obscurity and disagreement.

Chapter 7: Classification of duties. Veracity

1. Now that I have discussed benevolence, justice, and the observance of law and contract, it may seem I have taken in the whole sphere of social duty, and that any other maxims accepted by common sense must be applications of the principles I have been trying to define.

[Sidgwick starts explaining why this may seem to be so, but then he apparently changes direction and focuses on a different idea about how to classify duties, namely into *social* and *self-regarding*. He rejects that as a basic theoretical classification, for reasons including the fact that it cuts *thorough* some virtues that seem to be unitary:] Consider the acts morally prescribed under the head of *courage*. It seems clear that this virtue has been prominent in historic systems of morality because of the great social importance it must always have when communities of men are continually called on to fight for their existence and well-being; but still the quality of bravery is essentially the same, whether it is exhibited for selfish or social ends.

When we are trying to pin down the kinds of conduct commended or prescribed in any commonly recognised list of virtues, the maxims we come up with are clearly not absolute and independent: the quality named by our virtue-word is agreed to be praiseworthy only when it promotes individual or general welfare, and when it works against these ends it becomes blameworthy, even if it remains intrinsically just the same. I have already called attention to one or two examples of this, and I'll illustrate it at length in the following chapters. But though there's a great deal of this in our moral thought, I want now to focus on duties that seem *not* to fit this pattern, because they are specially characteristic of the method that we call intuitionism.

One of the most important of these is •veracity; and it's convenient to take this immediately after discussing •fidelity to promises, because these two duties, despite their fundamental differences, have a good deal in common: each prescribes a certain correspondence between words and facts, so that the questions that arise when we try to make the maxims precise are somewhat similar. For example:

- •My duty [see Glossary] of fidelity involves acting in conformity not with the *admissible* meaning of certain words but with the meaning that I know the promisee took them to have;
- •My duty of veracity is not to utter words that *would* give beliefs corresponding to mine to any hearers who understood me according to common usage, but to utter words that I believe *will* have this effect on the hearers that I actually have.

This is usually a simple matter, because the natural effect of language is to convey our beliefs to others, and we commonly know quite well whether or not we are doing this. With veracity as with fidelity—·i.e. with statements as with promises·—a certain difficulty arises from the use of set forms; and most of chapter 6's discussion of the similar difficulty applies here, with obvious modifications.

[Biographical background to what comes next: When in 1859 Sidgwick became a Fellow of Trinity College, Cambridge, he formally signified his acceptance of certain doctrines of the Church of England. Every new Fellow had to do this, though many of them (not Sidgwick) regarded it as an empty formality. A decade later he no longer believed all those doctrines and astonished everyone by resigning his Fellowship.] In the case of formulae imposed by law, such as declarations of religious belief, may we understand the terms in any sense

that they commonly have, or must we take them in the sense intended by the legislature that imposed them? Another difficulty is created by the fact that the strong inducement offered for their acceptance leads to gradual degradation or perversion of their meaning; they are continually stretched until there gradually comes to be a new general understanding of the meaning of certain phrases; and there's a continual dispute about whether we can truthfully use the phrases in this new meaning. A similar process continually alters the meaning of conventional expressions in polite society: when a man declares that he 'has great pleasure in accepting' a tiresome invitation, or is 'the obedient servant' of someone he regards as an inferior, he is using phrases that were probably once deceptive. If they don't now deceive, common sense condemns as over-scrupulous the refusal to use them where it is customary to do so. But common sense seems doubtful and perplexed when the process of degradation is incomplete and there are still persons who may be deceived; as in the use of the message that one is 'not at home' to an inconvenient visitor.

But apart from the use of conventional phrases, the rule 'to speak the truth' is not hard to apply; and the simplicity and definiteness of this maxim have led many moralists to regard it as an unchallengeable instance of an *ethical axiom*. But patient reflection [Sidgwick's phrase] will show that the common sense of mankind doesn't really agree.

2. For a start: Is veracity an absolute and independent duty or a special case of some more general principle? There's no clearly agreed answer to this. We find Kant saying that the duty to speak the truth is *owed to oneself*, because 'a lie is an abandonment or, as it were, annihilation of the dignity of man'. And a somewhat weakened version of this seems to be involved in the view that lying is prohibited by the code of honour—the view that lying for selfish ends, especially out of fear, is low and base. The code of honour *requires* lying in some circumstances, though here it clearly diverges from the morality of common sense. Still, common sense doesn't seem to decide clearly between these two:

- Truth-speaking is absolutely a duty, needing no further justification.
- Truth-speaking is merely each man's general right to have truth spoken to him by his fellows—a right that can be forfeited or suspended in certain circumstances.

Each man is thought to have a natural right to personal security generally, but not if he's trying to harm others in life and property; so if we may even *kill* in defence of ourselves and others, it seems strange if we may not *lie* if lying will defend us better against a clear invasion of our rights; and common sense doesn't seem to prohibit this decisively. Another example: the orderly and systematic slaughter that we call 'war' is thought to be perfectly right in certain circumstances, painful and revolting though it is; similarly in the word-contests of the law-courts, the lawyer is commonly held to be justified in untruthfulness within strict rules and limits; for it's thought to be over-scrupulous for an advocate to refuse to say what he knows to be false, if he is instructed to say it. Again, common sense seems to concede that deception is sometimes right when it's designed to benefit the person who is deceived. e.g. speaking falsely to an invalid if this is the only way to conceal facts that might produce a dangerous shock; and I don't see that anyone shrinks from telling fictions to children, on matters on which it is thought better they didn't know the truth. But if we ever allow that benevolent deception is lawful, I don't see how we can decide when and how far it is admissible, except by. . . .weighing the gain of each particular deception against the risk to confidence involved in all violation of truth.

Then there's the much argued question of religious deception—' pious fraud'. Common sense *now* pronounces against the broad rule that it is all right to tell falsehoods in the interests of religion. But moral persons do accept a subtler form of that same principle. It's sometimes said that the most important religious truths can't be conveyed into the minds of ordinary men except by being enclosed, as it were, in a shell of fiction; so that by relating such fictions as if they were facts, we are really performing an act of substantial veracity.[1] Reflecting on this argument, we see that it's not so clear, after all, what veracity is. [The next few sentences depart rather freely from Sidgwick's words, preserving their content but making the passage clearer because briefer.] When someone asserts that P, his hearer may come to believe that P and also to infer a consequence Q from it; the speaker may have foreseen this and intended to get across Q as well as P. Two complexities arise, of which we have just seen one: the speaker may think it important to communicate Q , and not know how to do this except by way of inference from P, which is false. If Q's truth is much more important than P's falsity, some people will credit this speaker with veracity. But others have the exactly opposite view about the duty of veracity. Suppose that the speaker asserts P, which is true, wanting the hearer to infer from this that Q, which is false. Some people hold that here veracity has been maintained, our only absolute duty being to make our actual affirmations true. They think:

> Human converse *ideally* involves perfect sincerity and candour, and we ought to rejoice in exhibiting these virtues where we can. But in our actual world concealment is often required for the well-being of society, and can be legitimately effected by any means short of actual falsehood.

This involves the quite common view that in defence of a secret we may not indeed *lie*, i.e. *directly* produce beliefs contrary to fact; but we may 'throw the inquirer on a wrong scent', i.e. *indirectly* produce a false belief through a natural inference from a true thing that we say. Other people say that if deception is to be practised at all, it is mere formalism to object to one method of deception more than another.

So reflection seems to show that the rule of veracity, as commonly accepted, can't be elevated into a definite moral axiom, because there's no real agreement on how far we are bound to impart true beliefs to others; and while it's contrary to common sense to demand absolute candour in all circumstances, no self-evident secondary principle tells us clearly *when* it is not to be demanded.

3. But we mustn't overlook one method of exhibiting *a priori* the absolute duty of truth because if it is valid it would seem that the exceptions and qualifications that I have mentioned ·ought to be rejected, and· have been accepted by common sense only through carelessness and shallowness of thought.

It goes like this:

> 'If it were once generally understood that lies were justifiable under certain circumstances, it would immediately become quite useless to tell lies because no-one would believe them; and the moralist can't lay down a rule which, if generally accepted, would be suicidal.

There seem to be three answers to this. **(a)** It isn't necessarily an evil that men's confidence in each other's assertions should *in certain special circumstances* be impaired or destroyed. So far from being an evil, it may even be the result we're trying to produce: a good way to protect legitimate

[1] For example, certain religious people have held recently that it is right to affirm solemnly 'God created the world in six days and rested on the seventh' *as a way of saying* that 1:6 is the divinely ordered proportion between rest and labour.

secrets would be to let everyone know that if they asked questions about those matters they will be answered with lies. And we shouldn't be restrained from pronouncing it lawful to meet deceit with deceit by the fear of spoiling the security that rogues now get from the veracity of honest men! No doubt the end-result of general untruthfulness in these circumstances would be that such falsehoods would no longer be told; but unless this result is undesirable, the prospect of it isn't a reason why the falsehoods shouldn't be told as long as they are useful. **(b)** Because men's beliefs generally aren't formed purely on rational grounds, experience shows that untruthfulness can remain partially effective in circumstances where it is generally understood to be legitimate. We see this in the law-courts. Jurymen know that it's regarded as an advocate's duty to state as plausibly as he can whatever he has been instructed to say on behalf of his criminal client, and yet a skillful pleader can often produce an impression that he sincerely believes his client to be innocent; and there's still debate about how far this kind of hypocrisy is justifiable. **(c)** It can't be assumed as certain that it is never right to act on a maxim of which the universal application would be an undoubted evil. This assumption may seem to be involved in an admitted ethical axiom, namely that what is right for me must be right for everyone in similar conditions (see page 98). But. . . .suppose that an agent •knows that the maxim he wants to act on is *not* universally accepted, and •reasonably believes that his act won't significantly contribute to its becoming so. [The next sentence is exactly as Sidgwick wrote it.] In this case the axiom will practically only mean that it will be right for all persons to do as the agent does, if they are sincerely convinced that the act will not be widely imitated; and this conviction must vanish if it *is* widely imitated. These conditions are *possible*; so the axiom that I'm discussing can only serve to direct our attention to an important danger of unveracity, which constitutes a strong but not rigorously conclusive utilitarian argument for speaking the truth. (This axiom will be further discussed in IV/5.3.)

[The chapter ends with a half-page note criticising the treatment of veracity in Leslie Stephen's *Science of Ethics*, which was published between the second and third editions of the present work. The exchange hardly seems worth pursuing now, except for the last sentence of the note:] The general question raised by Mr Stephen as to the value of 'internal rules' expressed in the form 'Be this', in contrast to external rules expressed in the form 'Do this', will be dealt with in chapter 14.1.

Chapter 8: Other social duties and virtues

1. When we start asking how far the minor social duties and virtues that common sense recognises are anything more than special cases of the general or particular benevolence discussed in chapter 4, the loudest call for our attention comes from duties involving the existence and legitimacy of feelings that are the opposite of benevolent.

Malevolent affections [see Glossary] are as natural to man as benevolent ones are; not indeed in the same sense ·of 'natural'·—for man tends to have normally some kindly feeling for any fellow-man when nothing special is making him love or hate. . . ., but special causes of malevolent feeling continually occur, and exemplify a psychological law like the law explaining the growth of benevolent feelings. For just as we're apt to love those who cause pleasure for us whether deliberately or otherwise; so also, by strict analogy, we naturally dislike those who have done us harm, whether consciously from malevolence or selfishness or even unconsciously, as when someone is an obstacle to our getting something we greatly want. We naturally feel ill-will towards a rival who deprives us of something we both wanted; which is why it's easy get someone who strongly desires superiority to dislike anyone who is more successful or prosperous than he is; and ·this dislike·, this envy, however repulsive to our moral sense, seems as natural as any other malevolent emotion. Each of the elements into which we can analyse malevolent affection has an exact counterpart in the analysis of the benevolent affection. For example, malevolence includes

> a dislike of the presence of its object, a desire to give it pain, and a capacity to get pleasure from the pain thus given.[1]

·In that, replace 'dislike' by 'like' and (twice) 'pain' by 'pleasure', and the result characterises benevolence.·

How far is it right and proper to indulge in malevolent emotions? Common sense's answer to this isn't easy to formulate. Some would say that they ought to be repressed altogether or as far as possible. And no doubt we blame all envy (though sometimes to exclude it altogether requires a magnanimity that we praise); and we regard as virtues or natural excellences •the good-humour that prevents one from feeling even pain to a material extent let alone resentment from trifling annoyances inflicted by others, •the meekness that doesn't resent even graver injuries, •the mildness and gentleness that refrain from retaliation, and •the ability to forgive rapidly and easily. We're even apt to praise the mercy that spares punishment, even *deserved* punishment; because though we never exactly disapprove of the infliction of deserved punishment, and hold it to be a duty of government and sometimes of private persons to inflict it, we think that this duty admits of exceptions; we think that in exceptional cases there can be reasons—not strictly relevant to the question of justice—for remitting punishment, and we admire the sympathetic nature that eagerly employs these legitimate occasions for remission.

[1] Men also get pleasure from the pains and losses of others in ways too: •from the sense of power that explains much of the wanton cruelty of schoolboys, despots, etc., •from a sense of their own superiority or security in contrast with the failures and struggles of others, or •even from the excitement caused by the manifestation or representation of any strong feeling in others—a real tragedy is interesting in the same way as a fictitious one. These facts, though psychologically interesting, present no important ethical problems, because no-one doubts that pain ought not to be inflicted from such motives as these.

On the other hand common sense admits that instinctive resentment for wrong is legitimate and proper; and even a more sustained and deliberate malevolence is commonly approved as virtuous indignation. How, then, are we to reconcile these diverse approvals? There's some difficulty about this even as regards external duty, ·the duty to *act* in this or that way·. Here is why:

> Common sense firmly holds that •the punishment of adults ought generally to be inflicted by government, and that •a private individual who has been wronged ought not to 'take the law into his own hands'; yet in all societies there are harms to individuals that the law doesn't punish adequately, if at all, and for which effective requital is often possible without breaking the law; and there's no clear agreement about how we should deal with these.

The Christian code is widely thought to demand a complete and absolute forgiveness of such offences, and many Christians have tried to obey this rule by getting the offences out of their minds as much as they can, or at least not letting the memory of them affect their outward conduct. But few people would deny this:

> If a man wrongs me in some way, that's a reason to expect that he will later do bad things to myself or to others, so I am obliged as a rational being to take precautions against this;

and probably most would admit that such precautions may include the infliction of punishment, where impunity would tempt the offender to repeat the unpunished offence. Well, then, *how far* is forgiveness practically possible? That seems admittedly to depend on **(a)** how far the punishment is really required in the interests of society, and **(b)** how far, if so, it will be adequately inflicted if the person wronged refrains from inflicting it. But this way of settling the question is hard to distinguish from the utilitarian method.

And there's trouble over the legitimacy of malevolent feeling. Here again thoughtful persons disagree; many would say that although anger is legitimate, it ought to be directed always against wrong acts and not against the agent, because even where the anger may legitimately prompt us to punish him it ought never to overcome our kindly feeling towards him. If this state of mind is possible, it seems to be the simplest reconciliation of •the general maxim of benevolence with •the admitted duty of inflicting punishment. On the other hand, it is maintained with some reason •that

> to retain a genuine kindly feeling towards a man, while gratifying a strong impulse of aversion to his acts by inflicting pain on him,

requires an emotional complexity that is too far out of the reach of ordinary men to be prescribed as a duty; and that •we must allow a suspension of benevolence towards wrong-doers until they have been punished. Some distinguish instinctive resentment from deliberate resentment, saying that the former is legitimate when required for self-defence and the repression of mutual violence, but that deliberate resentment is not similarly needed because if we act deliberately we can act from a better motive than that. But others think that the interests of society require a deliberate and sustained desire to punish wrong-doers, because the mere desire to be just won't in practice be strong enough to repress offences; and that replacing •natural resentment by the desire for justice is as serious a mistake as replacing •a natural appetite in eating and drinking by prudence or replacing •filial affection by mere dutifulness.[1]

[1] Butler (Sermon 8, on *Resentment*) recognises that deliberate resentment 'has in fact a good influence on the affairs of the world', though 'it were much to be wished that men would act from a better principle'.

We might distinguish the impulse to inflict pain from the desire for the pleasure one will get from inflicting it; so that while we approve the impulse in certain circumstances we may regard the desire for pleasure as altogether inadmissible. But a man under the influence of a strong passion of resentment can hardly keep out of his mind any anticipation of the pleasure he'll feel when the passion is gratified; and if so, he also can't keep out any desire for this gratification. So if it's important for society's well-being that men should get great satisfaction from the punishment of a nefarious criminal, it may be going too far to declare the desire for this satisfaction to be absolutely wrong; though we may say that a man ought not to cherish this desire, and gloat over the anticipated pleasure.

Summing up: a superficial view of the matter leads us to condemn *all* malevolent feelings and the acts they lead to, as contrary to the general duty of benevolence; but the common sense of thoughtful persons recognises that this rule should be relaxed in the interests of society. Common sense is unsure about the limits or principles of this relaxation, but is inclined to let it [here = 'the limits'] be fixed by considerations of expediency.

2. The remaining virtues that are clearly and exclusively social can easily be seen not to have independent maxims; they are special cases of virtues that I have already discussed. So I needn't discuss them in detail—I'm not aiming to generate a complete glossary of ethical terms—but for illustration's sake I should perhaps discuss one or two of them. I select for examination *liberality* with its cognate notions, because of the prominence that it has had in earlier times and because of a certain complexity in people's feelings about it. Considered as a **virtue**, liberality seems to be merely *benevolence, as exhibited in the giving of money, beyond the limits of strict duty as commonly recognised*. And liberality can also be called a **duty**: we don't like it when someone performing one of the somewhat indefinite duties listed in chapter 4 tries to be *exact*; we think that a certain *excess* is needed if the duty is to be well done; ·and that is where liberality comes in·. In the case of the poor, this graceful excess is perhaps excluded by prudence: a poor man might make a great sacrifice in giving a small gift, but we would call this 'generous' but hardly 'liberal'—liberality seems to require a gift to be large. So it seems that only the rich can be liberal; and the admiration liberality commonly arouses seems to contain an element that is aesthetic rather than moral. We're all apt to admire *power*, and we recognise the latent power of wealth gracefully exhibited by careless profusion in giving happiness to others. Indeed the vulgar admire the same carelessness as manifested even in selfish luxury.

The sphere of liberality, then, lies mostly in the fulfilment of the indefinite duties of benevolence. But it also—and especially—shows up in a certain borderland between •justice and •benevolence, namely in the full satisfaction of all customary expectations, even when these are indefinite and uncertain. [Sidgwick gives two sorts of examples. **(1)** x could get y to work for him for less than x is paying; x could get z to pay him more than he does for working for z; in each case x's liberality shows in his choosing not to haggle. **(2)** A liberal man is involved in some matter of law or contract where there's unclarity about *amounts*, and x 'unhesitatingly adopts the interpretation that is least in his own favour'. Then:] We describe as 'mean' a man who does the opposite of all this—meanness being the vice that is opposite to the virtue of liberality. . . . The common disapproval of meanness, like the approval of liberality, includes an element that isn't strictly moral ·but rather aesthetic·. Just as a certain carelessness of money is admired as a sign of power and

superiority, the opposite habit is a symbol of inferiority. The mean man is apt to be despised as having the bad taste to show this symbol needlessly, preferring a little gain to the respect of his fellow-men.

But meanness has a wider sphere than liberality, and refers not merely to •the taking or refusing of money but to •taking advantages generally; in this wider sense the opposite virtue is generosity [see Glossary].

In the area of their overlap, generosity seems to differ from liberality in

- •partly transcending liberality,
- •partly referring more to feelings than to actions, and
- •implying a more complete triumph of unselfish over selfish impulses.

Generosity in this wider sense is strikingly exhibited in conflict and competition of all kinds, where it is sometimes called 'chivalry'. The essence of this beautiful virtue is the achievement of benevolence in circumstances that make it specially difficult and therefore specially admirable. Generosity or chivalry towards adversaries or competitors seems to consist in showing as much regard for their well-being as is compatible with the ends and conditions of conflict—e.g. trying to achieve ideal justice in the conflict, not merely by observing all the rules and tacit understandings under which it is conducted, but by resigning [here = 'giving up'] even accidental advantages. Such resignation isn't regarded as a strict duty; and there's no agreement about how far it is right and virtuous, for some people would praise conduct that others would regard as quixotic and extravagant.

Thus, the terms 'liberality' and 'generosity' in their strictly ethical uses, name the virtue of benevolence (perhaps including justice to some extent) as exhibited in special ways and in special conditions. Examination of the other minor social virtues would lead to similar general results, though it mightn't always be easy to agree on their definitions.

Chapter 9: Self-regarding virtues

1. The morality of common sense assume or postulates an ultimate harmony between **(i)** self-interest and **(ii)** virtue: it's generally thought that it's a *duty to yourself* to do your duty generally and to develop your virtues—this being always conducive to your true interests and well-being. And common sense (in modern Europe) recognises a strict duty to preserve your own life, even when that life will be one in which pain outweighs pleasure. Indeed, it is held to be right and praiseworthy to encounter certain death

- •in the performance of strict duty, or
- •for the preservation of someone else's life, or
- •for any very important gain to society,

but not merely

- •to avoid pain to oneself.

At the same time, within the limits fixed by this and other duties, common sense considers that it is a duty to seek our own happiness except when we can promote the welfare of

others by sacrificing it.[1] This 'due concern about our own interest or happiness' may be called the duty of 'prudence'. It's less obvious that •men don't adequately desire their own greatest good than that •their efforts to achieve this are not well directed; and for that reason when prudence is thought of as a virtue or excellence the focus is almost entirely on its intellectual side. Seen in that way, prudence can be said to be merely *wisdom made more definite by the acceptance of self-interest as its sole ultimate end*—the habit of •calculating carefully the best means to achieving our own interests, and •resisting all irrational impulses that could upset our calculations or prevent us from acting on them.

2. But there are current notions of particular virtues that might be called 'self-regarding', though it's not quite clear whether •they are merely special cases of prudence or •have independent maxims. The most prominent of these is *temperance*, one of the four cardinal virtues recognised in ancient times. In its ordinary use, temperance is the habit of controlling the principal appetites (i.e. desires with an immediate physical cause). Common sense recognises as useful and desirable the habit of moderating and controlling our desires generally, but it is less clear and emphatic about this.

Everyone agrees that our appetites need control; but we can't have a maxim of temperance unless we determine

•within what limits,
•on what principle, and
•for what purpose

they ought to be controlled. As for the appetites for food, drink, sleep, stimulants, etc., no-one doubts that the natural purpose of gratifying them is bodily health and vigour, and that they should be reined in when they tend to defeat this purpose (including mental health insofar as it depends on the general state of the body). And the indulgence of a bodily appetite is obviously •imprudent if it involves the loss of any greater pleasure, and •wrong if it interferes with the performance of duties; though it's not clear how far this latter indulgence would commonly be condemned as 'intemperance'.

Some people derive from •the obvious truth that bodily health is the chief natural purpose of the appetites •a more rigid rule of restraint that goes beyond prudence. They say that. . . .we should seek the pleasure of gratifying an appetite *only* when such gratification is positively conducive to health. Considering how markedly this condemns the usual habits of the moral rich, we might be disposed to say that it is clearly at variance with common sense; but it often meets with verbal assent.

There's an intermediate view according to which the gratification of appetite is to be sought—not indeed for its own sake, but—for more purposes than mere health and strength, e.g. to produce 'cheerfulness, and the cultivation of the social affections' (quoting Whewell). Many people seem to accept something like this, not always consciously: solitary indulgence in the pleasures of the table is often regarded with something like moral aversion; and that banquets that moral people give or enjoy are vaguely supposed to aim not at •the common indulgence of sensual appetites but

[1] Kant argues that this can't be regarded as a duty because everyone 'inevitably wills' means to promote his own happiness . But, as I contended in I/4.1 , a man doesn't 'inevitably will' to do what he believes will be most conducive to his own *greatest* happiness.
The view in the text is that of Butler ('The nature of virtue', appended to his *Analogy of Religion*), who admits that 'nature has not given us as strong a sense of disapproval of imprudence and folly as of falsehood, injustice, and cruelty', but points out that for various reasons such strongly sensed disapproval is less needed in the former case.

•the promotion of conviviality and enjoyable conversation. . . . Still it would be going too far to credit common sense with accepting the maxim that no sensual pleasures are ever to be sought except when they positively promote pleasures of a higher kind.

3. That was mainly about the appetites for food and drink. But the area where •morally prescribed regulation most clearly and definitely goes beyond mere prudence is that of the sexual appetite; •it is indicated by the special notion of *purity* or (the same thing except a bit more external and superficial) *chastity*.

You might at first sight think this:

> Common-sense **morality** restricts the sexual appetite merely by confining its indulgence within the limits of **legal** marriage; but because this natural impulse is so powerful and easily aroused it's especially necessary to prohibit any acts—internal or external—that might lead the person to cross these limits.

This is largely true; but on reflection we'll find that our common notion of **purity** involves a standard that is independent of **law**, for two reasons: **(a)** conformity to the law doesn't guarantee purity; and **(b)** sexual intercourse that doesn't conform to the law isn't always thought to be impure. The two notions—illegality and impurity—are sometimes run together, but that is a mere mistake.[1] But it's not clear what this standard of ours *is*. When we interrogate the moral consciousness of mankind, we seem to get two answers, one stricter than the other, analogous to the two versions of temperance discussed in section 2. They agree that the sexual appetite ought to be indulged only as a means to some higher end, and never merely for the sake of sensual gratification; but then they part company. Some people say that the propagation of the species— obviously the primary natural purpose—is also the only permissible purpose. Others see a different purpose as perfectly admissible and right, namely the development of mutual affection in a union designed to be permanent. The practical difference between the two views is considerable; so that this question needs to be asked and answered. But any attempt to lay down minute and detailed rules about this seems to be condemned by common sense as tending to defeat the purpose of purity, because such minuteness of moral legislation invites men to exercise their thoughts on this subject to an extent that is practically dangerous.[2]

The virtue of purity is not merely self-regarding, so it doesn't really belong in this chapter; but I put it here because of the convenience of discussing it along with temperance. Some would go further and say that purity should be treated as a distinctly *social* virtue; the propagation and rearing of children is one of the most important of social interests, and thee people hold that purity is simply a sentiment protective of these important functions, supporting the rules that we think are needed to secure their proper performance. But it seems clear that although common sense undoubtedly does •recognise that the sentiment of purity is conducive to the best possible provision for the continuance of the human race, it doesn't •regard that as the crux of this rule of duty and the sole criterion in deciding whether acts violate the rule.

[1] When illegality of union is taken to be prohibited directly and specifically—not merely for reasons of prudence and benevolence—it is being regarded as a violation of order rather than of purity.

[2] Part of what brought medieval casuistry [= 'practical ethics'] into disrepute was the failure to see that purity itself forbids too minute a system of rules for the observance of purity.

There seem to be no similar special questions regarding most other desires. We do recognise a general duty of self-control; but this is merely as a means to acting rationally (however we understand that); it only requires us not to yield to any impulse that would make us act against ends or rules that we have accepted. Among moral persons there's a tendency to the ascetic opinion that the gratification of merely sensual impulse is in itself objectionable; but this doesn't seem to be the view of common sense in particular cases—we don't condemn intense enjoyment of muscular exercise, or warmth, or bathing. The only other natural impulses that the common sense of our age and country thinks it right or admirable to repress—apart from what prudence and benevolence would dictate—are the promptings of pain and fear. An important instance of this is the absolute prohibition of suicide even when it's very probable that the rest of a man's life will be •miserable and •burdensome to others. But there are other cases also where praise is apparently bestowed on endurance of pain and danger beyond what is conducive to happiness. We'll see this in the next chapter.

Chapter 10: Courage, humility, etc.

1. Some prominent excellences of character are commonly admired and inculcated without, apparently, any clear reference to private or general happiness. They are usually conducive to one or other of those ends, but sometimes they seem to turn conduct against them.

Courage arouses general admiration, whether it is shown in self-defence or in aiding others or even when it doesn't bring any benefit that we can see. In Christian societies sincere *humility* often receives unqualified praise, in spite of what a man may lose by underrating his own abilities. It will be well, therefore, to examine how far in either case we can elicit a clear and independent maxim defining the conduct commended under each of these notions.

Courage: We take courage to be a disposition to face danger of any kind without shrinking. We sometimes also describe as 'courageous' those who bear pain unflinchingly, but it's more usual to call this 'fortitude'. What commonly recognised duties are there involving courage or fortitude? It seems clear that any answer to this will bring in other maxims and purposes. No-one would deny this:

> The only times when we have a duty to •face danger or •bear avoidable pain is when the need for this comes up in the course of ·some other· duty.[1]

And not always, even then: with duties such as those of general benevolence, it would be commonly allowed that

- •the agent's pain and danger are relevant in deciding how far his duty extends in the given case;
- •we're obliged to endure pain only when that will prevent obviously greater pain to someone else, or will achieve a more important amount of positive good;

[1] With pain that can't be avoided, fortitude will suppress outcries and lamentations; though if these relieve the sufferer without annoying others, it seems doubtful that this is a *duty*.

> •and we're obliged to run risks only when the chance of additional benefit for someone else outweighs the cost and chance of loss to ourselves if we fail.

The duty of benevolence, as commonly estimated, may not stretch as far as this. (See chapter 4.5.)

But when courage is viewed as an •excellence rather than a •duty, it seems to have a more independent position in our moral estimation. And this view of courage fits the common application of the notion better than does the 'duty' view of it; many acts of courage are not entirely under the control of the will, and therefore can't be strict duties. •Danger is often sudden, and has to be met without deliberation, so our way of meeting it can only be *semi*-voluntary. •And although given time for deliberation a naturally timid man may be able to control fear (as he can anger or appetite), preventing it from taking effect in dereliction of duty, this won't be enough for him to perform courageous acts that require great energy. Why not? Because the timid virtuous man's energy is likely to have been exhausted by the effort to control his fear; in battle he can stand still to be killed as well as the courageous man, but not charge with the same forceful rush or strike with the same vigour and precision.

Given that courage is not completely voluntary, we have to replace the question 'Are we strictly obliged to show courage?' by the question 'Is courage a desirable quality?'. There's no room for doubt that we commonly find it morally admirable without reference to any purpose served by it—when the dangers that call it forth could be avoided without any failure of duty. But a man who needlessly runs into danger beyond a certain degree we describe as 'foolhardy'. Where then is the limit to be fixed? On utilitarian principles we should balance as accurately as we can

> the amount of danger incurred in the given case

against

> the probable benefit of cultivating and developing by practice a habit that is so often needed for the performance of important duties.

This will obviously give different results for different states of society and different callings and professions; because most people need this instinctive courage less in civilised societies than in semi-barbarous ones, and civilians need it less than soldiers. Perhaps mankind's instinctive admiration for acts of daring doesn't altogether observe this limit; but when we do try thoughtfully to justify this admiration, we commonly do it in some such way as this; and common sense doesn't point to any limit depending on a different principle.

2. Humility: The virtue of courage is prominent in pagan ethics, and in the code of honour that is a sort of survival of the pagan view of morality; whereas humility especially belongs to the ideal that Christianity sets before mankind. But the common account of this virtue is somewhat paradoxical. It is generally said that humility prescribes a low opinion of our own merits; but if our merits are comparatively high it seems strange to direct us to have a low opinion of them. This may be said in reply:

> Even if our merits are high when compared with those of ordinary men, there are always some superior merits that we can compare ourselves with, right up to *ideal* excellence, of which we all fall far short. That's the kind of comparison that we ought to make, dwelling on our many faults and not on our merits.

But surely in human life's most important deliberations—in determining what work we will do and what social roles we will aspire to—a good decision often requires us to compare our qualifications with those of average men. And it seems just as irrational to underrate as to overrate ourselves. . . .

The word 'humility' isn't always used eulogistically; but when it is, what is being praised in the common judgments

using the notion of *humility* is a quality that •doesn't influence our opinions of ourselves—for here as in other opinions we ought to aim at nothing but truth—but •influences our emotion by tending to repress two different seductive emotions, one entirely self-regarding, the other relating to others and partly showing in social behaviour. One is the emotion of *self-admiration*, which arises naturally from the contemplation of our own merits. . . .and tends to cause such contemplation. This admiring self-satisfaction is generally condemned, but I don't think this comes from an intuition that claims to be basic; rather, it is commonly derived from the belief that such self-admiration, even if it is well-grounded, tends to hinder our progress towards higher virtue. The mere fact of our feeling this admiration is thought to be evidence that •we haven't sufficiently compared ourselves with our ideal or that •our ideal isn't high enough; and it is thought that our moral progress needs us to have a high ideal and to keep it continually in mind. But we obviously need to be careful how we apply this maxim. Everyone agrees that self-respect is an important aid to right conduct; and moralists point to the satisfactions of a good conscience as part of the natural reward that Providence has attached to virtue; and it's hard to separate •the glow of self-approval from performing a virtuous action from •the satisfied self-consciousness that humility seems to repress. Perhaps we can say that the feeling of self-approval *is* natural and a legitimate pleasure which is liable to impede moral progress if it is prolonged and cultivated; and that humility prescribes a repression of self-satisfaction that will tend on the whole to stop this prolongation. On this view the maxim of humility is clearly a derived one, and is subsidiary to the end of progress in virtue generally. As for pride and self-satisfaction *not* based on our own conduct and its results but on external and accidental advantages, these are condemned as involving a false and absurd view about the nature of real merit.

But most of us take pleasure in the respect and admiration not only of ourselves but still more of others. The desire for this is also regarded as to some extent legitimate, and even as a valuable aid to morality; but it's a dangerously seductive impulse which often acts in opposition to duty, so it is felt to stand in special need of self-control. Humility, however, consists not so much in •controlling this desire as in •repressing our natural inclination to get others to satisfy it. We're inclined •to demand 'tokens of respect' from others, some symbol of their recognition of how fine we are, and •to complain if our demands are not met. Humility tells us to repress such claims and demands. In many cases where others have a strict duty to express reverence towards someone, *he* is thought to have a duty not to demand it. But common sense holds that there's a limit at which this quality of behaviour passes over into a fault; for the omission of marks of respect[1] is sometimes an insult that we are prompted to repel by impulses that are commonly regarded as legitimate and even virtuous—sense of dignity, self-respect, proper pride, and so on. I don't think there's any agreed formula for determining where this limit lies.

[1] Omission of the customary •marks of respect for officials would be a breach of established order; but I'm not talking about •these, because the special political reason for requiring them goes far beyond the sphere of the virtue of humility.

Chapter 11: Review of the morality of common sense

1. I have now carried out the plan laid down in chapter 1 for examining the morality of common sense, doing it in as much detail as seemed desirable. I haven't discussed all the terms in our common moral vocabulary, but I think I have covered all that are important •in themselves or •relative to my present inquiry. The ones I have omitted don't bring in independent maxims: the conduct designated by them is either •prescribed merely as a means to performing duties already discussed or •really identical with the whole or part of some of these. . . .

Let us look back at what I have been doing and the results I have come up with.

At the outset I admitted the existence of apparently independent moral intuitions, this being the thesis that many moralists have concentrated their efforts on trying to prove. It seemed undeniable that men judge some acts to be right and others to be wrong *in themselves*, without. . . .taking their consequences into account at all, except for consequences that are included in the common notion of the act. But we saw that when these judgments are passed on particular actions, they seem (at least for thoughtful people) to bring in some general rule of duty; and that in the frequent cases of doubt or conflict of judgments about the rightness of any action, appeal is commonly made to such rules or maxims, as the ultimately valid principles of moral knowledge. To put the morality of common sense into a scientific form, therefore, I needed to get as exact a statement as possible of these generally recognised principles. I didn't evade this task by pleading the unscientific character of common morality. The moral opinions of ordinary folk are indeed loose, shifting, and mutually contradictory, but it doesn't follow that we can't sift out from this fluid mass of opinion a deposit of clear and precise principles that would be accepted by everyone. *Can* we do this? The question, I thought, shouldn't be answered *a priori*, without a fair trial; and part of the task of chapters 3–10 has been to prepare materials for this trial. I have tried impartially to discover, simply by thinking about common moral discourse, what the principles or maxims are by which actions are judged to be right and reasonable in different parts of life. Please note that I haven't introduced *anywhere* views of mine that I don't think would be accepted generally; my sole aim has been to make explicit the implied premises of our common moral reasoning. I shall now subject the results of this survey to a final examination, in order to decide whether these general formulae have the marks by which self-evident truths are distinguished from mere opinions.

2. How can an apparently self-evident proposition qualify as having the highest possible degree of certainty? Four conditions are jointly sufficient for this—conditions that must be more or less satisfied by the premises of •our reasoning in any inquiry if •it is to lead us forcefully to trustworthy conclusions.

(1) The terms of the proposition must be clear and precise. The rival originators of modern methodology, Descartes and Bacon, both stress this, and Bacon's warning against the 'badly defined notions' is especially needed in ethical discussion. My chief business in the preceding chapters has been to free the common terms of ethics, as far as possible, from objection on this score.

(2) The proposition must be found to be self-evident *by careful reflection*. Most people are liable to think they have

an **intuition** when really it is only

- •an **impression or impulse** which doesn't look like a dictate of reason when it is carefully observed, or
- •an **opinion** to which the familiarity that comes from frequent hearing gives an appearance of self-evidentness—an appearance that attentive reflection disperses.

In such cases Descartes's method of testing the ultimate premises of our reasonings by asking ourselves if we clearly and distinctly apprehend them to be true—though he was wrong to think it gives a complete protection against error—can be really useful. A strict demand for self-evidentness in our premises can protect us from the influence of our own irrational impulses on our judgments; it distinguishes as inadequate the mere external support of authority and tradition, and blocks the more subtle and hidden effect of these in shaping our minds into an easy and unquestioning admission of common but unwarranted assumptions.

This test is especially needed in ethics. On one side: any strong sentiment, however purely subjective, is apt to seem like an intuition: when we *want* something we're apt to call it *desirable*, and when conduct gives us keen pleasure we're strongly tempted to give it our moral approval—and detecting these illusions requires careful thought. On the other side: of the rules of conduct that we customarily obey, many are shown by reflection to be really derived from some external authority; so that even if they are unquestionably obligatory, they can't be discovered intuitively. This is of course the case with the positive law of the community: we certainly ought to obey it, generally speaking, but of course we can't learn what it is by any process of abstract reflection, but only by consulting reports and statutes. These sources of knowledge, however, are so definite and conspicuous that we don't risk confusing what we learn from them with the results of abstract thought. It's not like that with the •traditional and •customary rules of behaviour that exist in every society, supplementing the law's work in regulating conduct; it's much harder to distinguish *these* from rules that a moral man is called on to define for himself by applying intuitively known principles. . . .

Consider two systems of rules that I compared with morality in I/3.2—the law of •honour, and the law of •fashion or etiquette. I remarked that the common terms 'honourable' and 'dishonourable' are ambiguous. No doubt they are sometimes used, like ethical terms, as implying an absolute standard. But when we speak of the 'code of honour' we seem to be referring to rules that are ultimately fixed by the general opinion of well-bred persons; when this opinion condemns a man we admit that he is in a sense 'dishonoured', even if we think his conduct acceptable or even intrinsically admirable (see I/3.2). Similarly with the rules of fashion or etiquette: looked at from the point of view of reason, some seem useful and commendable, some indifferent and arbitrary, some perhaps absurd and burdensome; but we recognise that the final authority on matters of etiquette is the custom of polite society, which doesn't feel obliged to reduce its rules to rational principles. But each individual in any society commonly finds in himself a seemingly complete knowledge of the rules of honour and etiquette, and an impulse to obey them without requiring any reason for doing so. He seems to see at a glance what is •honourable and polite just as clearly as he sees what is •right; and he would have to think hard to realise that in •the former cases custom and opinion are the final authority from which there is no appeal. And even in rules regarded as clearly *moral*, we can usually find an element that seems to us as clearly conventional as the code of honour, when we think about the morality of other men, even in our own age and country. So we can reasonably

suspect a similar element in our own moral code, and must admit the importance of testing rigorously *any* rule that we have a habitual impulse to obey, to see whether it really comes from a clear intuition of rightness.

(3) The propositions accepted as self-evident must be consistent with one another. Obviously, any collision between two intuitions proves that at least one of them contains some error. Ethical writers often treat this point very lightly. They seem to regard a conflict of ultimate rules as a difficulty that may be ignored or set aside to be solved later, without any slur on the scientific status of the conflicting formulae. Whereas such a collision is absolute proof that at least one of the formulae needs qualification; and suggests a doubt whether the correctly qualified proposition will look as self-evident as the simpler but inadequate one did; and whether we haven't mistaken a derivative and subordinate proposition for an ultimate and independent axiom.

[Sidgwick's point here is as follows. I am inclined to accept as ultimate the propositions

(a) I ought never to do A, and
(b) I ought never to do B;

and then I realise that these two 'collide' in cases where it is impossible for me to obey both at once. To deal with this I 'qualify' = 'modify' one of them, dropping **(b)** in favour of

(c) I ought never to do B in circumstances where P is the case,

with a value of P that prevents the collision. The odds are that **(c)** won't strike me as self-evident in the way that the simpler and seemingly basic **(b)** did.]

(4) My confidence in the validity of something I have asserted is likely to be weakened if someone else denies it. And in fact 'universal' or 'general' consent has often been thought to be, all by itself, sufficient evidence of the truth of the most important beliefs; and in practice it's the only evidence on which most of mankind can rely. A proposition accepted as true on this ground alone isn't self-evident or rigorously demonstrable; but our usual confident acceptance of the generalisations of the empirical sciences rests—even in the minds of experts—largely on the belief that other experts have seen the evidence for these generalisations and pretty much agree that it is adequate. And it's easy to see that the certainty of our beliefs won't survive if there *are* significant disagreements. If *any* of my judgments are in direct conflict with a judgment by someone else, there must be error somewhere; and if I don't have any reason to suspect error in the other person's mind rather than in my own, the upshot of my thoughtfully putting his judgment and mine side by side is that I'm forced into a temporary state of neutrality. The total result in my mind is not exactly suspension of judgment, but an alternation and conflict between positive affirmation by one act of thought and the neutrality that results from another; it's very different from scientific certitude.

It seems clear that the maxims of the morality of common sense—if my account of it chapters 3–10 is mainly correct—don't generally satisfy the four conditions I have just laid down. When they are left as somewhat vague generalities, as we meet them in ordinary discourse, we're inclined to give them unquestioning assent; and that assent is *approximately universal* in the sense that any dissent is eccentric and paradoxical. But when we try to give these maxims the definiteness that science requires, we find that we can't do this without losing the universality of acceptance. In some cases we have to choose between alternatives that are equally or nearly equally plausible, and common sense doesn't decide between them. In other cases the moral notion seems to resist all efforts to extract a definite rule from it; in yet others it brings together elements that we can't reduce to a common standard except by applying the utilitarian method or something like it. Even where we do

seem able to get common sense to give a fairly clear reply to the questions we raise in our pursuit of definiteness, the resulting principle is so complicated that its self-evidentness becomes dubious or vanishes altogether. Thus, in each case what at first seemed like an intuition turns out to be either •the mere expression of a vague impulse, needing regulation and limitation that has to be drawn from some other source, or •a current opinion whose reasonableness has still to be shown by a reference to some other principle.

So that I can adequately present this result, please travel with me again through the series of principles drawn from common sense in chapters 3–10, so that we can examine them from a different point of view. Our main aim so far has been to discover what the deliverances of common sense actually are; now we have to ask what claim they have to the status of intuitive truths.

Throughout this examination I'll be making a double appeal—to *your* moral consciousness, and to common sense as expressed generally by the people whose moral judgment you're willing to rely on. In each case I shall ask:

(1) Can you state a clear, precise, self-evident first principle that you're prepared to use in judging conduct of this kind?

(2) If you can, is this principle the one that is commonly applied by those whom you take to represent common sense?

[At this point Sidgwick says in a footnote:] I have been accused of leaving the determinations of common sense very loose and indefinite. So indeed I have. If I were trying to bring out a more •positive result from this examination, I ought certainly to have discussed further how we are to identify the 'experts' on whose 'consensus' we are to rely. But my scientific conclusions are so •negative that I thought it hardly necessary to go into this. I have been careful not to exaggerate the doubtfulness and inconsistency of common sense; if it turns out to be more doubtful and inconsistent than I have made it out to be, my argument will only be strengthened.

3. Let us start with the duty of acting wisely, discussed in **chapter 3**. We may seem here to have an undoubtedly self-evident axiom: acting wisely seemed to mean

- •taking the right means to the best ends, i.e.
- •taking the means that reason indicates to the ends that reason prescribes.

And it's evident that it must be right to act reasonably. Equally undeniable is the. . . .negative aspect of this principle, namely that it's wrong to act in opposition to rational judgment. From this, together with the empirical fact that we have impulses conflicting with reason, we get—as another self-evident principle—the maxim of temperance or self-control in its widest interpretation, namely that reason should never give way to appetite or passion.[1] And these principles ·of wisdom and temperance· have sometimes been solemnly offered as answering the basic question of ethics and supplying a comprehensive basis for a doctrine of how to behave.

But this statement of principles turns out to be a brief circuit leading us back to the point from which we started. (This happens annoyingly often in the course of ethical reflection!) To make sure that the point is understood: the maxims just given have two senses—in one sense they are self-evident but are also insignificant; in the other sense they point us more or less clearly towards an important duty, but

[1] In chapter 9 I treated temperance as a special application of prudence, i.e. self-love moralised. That's because that seemed to be the view of common sense, which I was trying to follow as closely as possible, both in •stating the principles of common sense and in •the order of their exposition.

in so doing they lose their self-evidentness. **First sense:** If the rules of wisdom and self-control mean **(1)** that we ought always to do what we see to be reasonable, and **(2)** that we are not to yield to any impulse urging us in an opposite direction, they simply affirm that it is our duty—**(1)** generally, and **(2)** under special temptations—to do what we judge to be our duty;[1] and say nothing about the method and principles by which duty is to be determined.

Second sense: But these rules are sometimes understood to prescribe the development of a habit of acting rationally, i.e. of basing each act on specific principles and ends, rather than letting it to be determined by instinctive impulses. ·This has real content, but· considering it as a universal and absolute rule of duty I can't see it as self-evidently true. . . . It presents us with the question 'Is reason's command *always* a good? Is it the case that the perfection of the conscious self must always be favoured by reason's being predominant over mere impulse, however great that predominance is?' It surely isn't self-evident that the right answer is Yes, i.e. that reason's predominance can't be carried too far. Perhaps there are limits to how much control reason should have; perhaps reason itself sets them, in the knowledge that rational ends are sometimes better achieved by those who don't directly aim at them as rational. Certainly common sense is inclined to hold that in many matters instinct is a better spring of action than reason: it is commonly said that

- •a healthy appetite is a better guide to diet than a doctor's prescription; and that
- •marriage is better undertaken as a consequence of falling in love than in the carrying out of a calm and deliberate plan;

and I noted in chapter 4 that services springing from spontaneous affection have a certain excellence that similar acts done from pure sense of duty don't have. Experience seems to show also that many acts requiring promptness and vigour are likely to be more energetic and effective, and that many acts requiring tact and delicacy are likely to be more graceful and pleasant to others, if they're performed not in conscious obedience to the dictates of reason but from other motives. For my present purposes, I don't need to know *how much* truth there is in this; it's enough I don't know intuitively that there's no truth in it. I don't know that there may not be—to use Plato's analogy—*over-government* in the individual soul no less than in the state. So the residuum of clear intuition that we have so far obtained is the empty proposition that it's our duty to do what we judge to be our duty!

4. Let us pass now to what in **chapter 4** I called the duties of the affections, i.e. rules that prescribe some degree of •love or of •the services that naturally spring from love in the relationships where it is expected and desired. Let's start with the question: 'How much of those services are we obliged to give if we don't feel the love?' In many cases this is answered differently by different persons, and no setting of the limit seems self-evident. Similarly with the question: 'Is love itself a duty?' On one hand: love is at most only partially under the control of the will, and when it *is* produced by voluntary effort there's thought to be something unsatisfactory and unattractive in it. On the other hand: in certain relationships it seems to be commonly regarded as a duty. On those points the doctrine of common sense is a rough compromise between conflicting lines of thought rather than something derivable from a clear and universally accepted principle. And if we confine ourselves to the relationships where common sense is sure there's

[1] Such a judgment may be objectively wrong; but while it's the judgment that I have it would be wrong for me not to act in accordance with it.

a broad moral obligation at least to give services such as love naturally prompts, there's still something—·actually, two things·—unsatisfactory about the rules of external duty that are commonly recognised in these relationships: •they aren't definite and precise, and •the details of the duties they prescribe don't seem to be based on independent intuitions. Consider the duty of parents to children. We have no doubt about this duty as a part of the present order of society, something that distributes among the adults the proper growth and training of the next generation. But when we think about this arrangement we can't see intuitively that it's the best possible. It is plausible to maintain that children would be better trained, physically and mentally, if they were brought up under the supervision of physicians and philosophers in large institutions maintained out of the general taxes. We can't decide *a priori* which of these alternatives is preferable; we have to bring in generalisations that psychologists and sociologists have obtained by empirical study of human nature in actual societies. Well, then, let's consider the duty of parents by itself and not as connected with this social order: it certainly isn't self-evident that we owe more to our own children than to others whose happiness we could equally affect. . . . Some people hold that my special duty to my own children arises from the fact that it's I who brought them into existence. It seems to follow from this that I have a right to lessen their happiness, provided I don't turn it into a negative quantity; because if it weren't for me they wouldn't have existed at all, their status as my children gives them no claim on me for anything more than an existence that is over-all above zero in respect of happiness. We might even infer a parental right to extinguish one's children painlessly at any point in their existence, provided that their life up to that point has been on the whole worth having; for how can persons who would have had *no* life but for me fairly complain that they aren't allowed more than a certain quantity?[1] I'm not saying that these doctrines are even implicitly held by common sense; I aim only to show that here as elsewhere the pursuit of an irrefutable intuition may draw us into a nest of paradoxes.

So it seems that we can't, after all, say that the special duty of parents to children, considered by itself, is clearly self-evident; and it was easy to show in chapter 4 common sense's limits for it are indeterminate.

We needn't linger on the rule prescribing the duty of children to parents. Common sense thinks that this *may* be merely a particular case of gratitude, and we have no clear intuition of what is due to parents who don't deserve gratitude. The moral relation of husband and wife seems to depend chiefly on contract and definite understanding. It is usually thought that morality as well as law prescribes conditions for all connubial contracts; and in our own age and country it is held that they should be monogamous and permanent. But clearly that neither of these ·moral· opinions would be maintained to be a primary intuition. Can any of the legal regulations of the union of the sexes be derived from some intuitive principle of purity? I will address that shortly [page 173]; but as for conjugal duties that aren't prescribed by law, probably no-one these days would maintain that they can be known *a priori*—there's not enough general agreement about what they are for that to be plausible.[2]

[1] A view similar to this has often been maintained regarding what God is in justice bound to do for human beings, given his quasi-parental relation to them.

[2] It's relevant here to remember the remarkable variety of suggestions for the better regulation of marriage that reflective minds seem to be led to once they are disentangled from the web of tradition and custom. . . .

In these domestic relations the duties of affection are commonly regarded as imperative and important; if we can't find any independent and self-evident principles for determining *them*, I needn't spend time showing that we can't find such principles either for the less intimate ties—of kindred, neighbourhood, etc.—that link us to other human beings. . . .

There are certain obligations towards human beings generally that are, speaking broadly, unquestionable [Sidgwick's phrase]. For example, the duty (whether of justice or benevolence) to •refrain from causing pain to others against their will, except as deserved punishment; and to •make reparation for any pain we have caused. But when we try to define the limits of these duties, asking *how far* we can legitimately go in causing pain to other men (or other sentient beings) so as to get happiness for ourselves or third persons—or even to confer a greater good on the sufferer himself if the pain is inflicted against his will—we can't obtain any clear and generally accepted principle for settling this, unless the utilitarian formula is openly admitted. And we have seen that there's a fundamental doubt about how far reparation is due for harm that was involuntarily caused.

Similarly, everyone agrees that we have a general duty to help our fellow-men—especially those in special need—and that when we can greatly benefit someone through a small loss to ourselves we're obliged to do that; but when we ask 'How much of our own happiness are we obliged to give up so as to promote the happiness of others?', common sense doesn't clearly accept the utilitarian principle but it doesn't definitely affirm any other.

And even the common principle of gratitude, which everyone immediately feels to be strict, seems to be essentially indeterminate because of the unsolved question:

> 'Ought the requital of a benefit be proportionate to what the benefit is worth to the recipient or to what it cost the benefactor?'

5. When we consider the element of justice that presented itself in **chapter 5** as *gratitude universalised*, the same difficulty recurs in a more complicated form. For now the question is:

> 'Ought the requital of good desert be proportioned to •the benefit rendered or to •the effort made to render it?'

[We'll now meet the second passage—the other was on page 135—where Sidgwick uses 'retribution' to cover rewards as well as punishments.] And if we look hard at the common moral notion of retributive justice, it seems to imply the metaphysical doctrine of free will. [Sidgwick now reminds us that retributive justice *looks back*, paying no attention to the possible consequences of this or that policy for rewards and punishments. He continues:] If every excellence in a man's actions or productions seems to come ultimately from causes other than himself, his claim to requital appears to vanish. On the other hand it is obviously paradoxical in estimating desert to *omit* moral and intellectual excellences due to heredity and education. (Why intellectual excellences? Because they come into moral evaluation: good intention without foresight is commonly regarded as a very imperfect merit.) Even if we cut through this speculative difficulty by leaving the ultimate reward of real desert to divine justice, we still can't find any clear principles for constructing a scale of merit. And much the same can be said *mutatis mutandis* [see Glossary] of the scale of demerit that criminal justice seems to require.

Even if these difficulties were overcome, we would still be only *starting* on the puzzles that beset the attempt to find self-evident principles on which to base a system of justice. Chapter 5's study of the contents of the notion of justice didn't turn up a single precise principle; all it

provided was a *swarm* of principles that are liable to come into conflict with each other. Some of them, looked at in isolation, do have the air of being self-evident truths, but they don't reliably carry with them any intuitively discoverable definition of the boundaries and relations between them. For example, in constructing an ideally perfect distribution of the means of happiness we have to take into account the notion that I call *fitness* (often confounded with *desert* but really distinct from it). What there is to be distributed in society includes not merely •the means of obtaining pleasurable passive feelings but also •functions and instruments that are important sources of happiness but which should be given to those who ·are *fit* to have them, i.e. to those· who can fulfill the functions and use the instruments. And even as regards the material means of comfort and luxury—in short: wealth—the same amount doesn't produce the same amount of happiness in every case; and it seems reasonable that the means of refined and varied pleasure should be given to those who have the corresponding capacities for enjoyment.[1] But the fittest may not be the most deserving, so that this principle can conflict with the principle of requiting desert.

And each principle, as we saw earlier [starting on page 131], is liable to collide with the widely-accepted doctrine that

> the proper ultimate end of law is to secure for everyone the greatest possible freedom of action, and each individual is obliged not to interfere with others, and that is his *only* obligation apart from any that he has taken on through a free contract.

And when we look into this 'freedom' principle itself we find that it can't provide a practical basis for social construction unless it is limited and qualified in ways that make it less like •an independent principle than like •a 'middle axiom' of utilitarianism; and that a lot of stretching is needed to make it cover the most important rights that positive law guarantees. How for example are we to justify *ownership*? On the grounds that it makes men free? or on the grounds that it provides the only adequate motive for labour? And we can't derive from this supposed basic principle the questions that arise concerning the limits of the right of property—e.g. whether it includes the right of bequest. Nor again is the enforcement of contracts a way of making people free: strictly speaking, a man is more free when no one of his volitions is allowed to cause an external control of any other. And if we disregard this as a paradoxical subtlety, we are met on the opposite side by the puzzle that if abstract freedom is consistent with *any* engagement of future services, it must be consistent with ones that are perpetual and unlimited, and so even with actual *slavery*. And here's something that makes this question especially important: many writers have tried to reconcile our duty to obey positive laws with the abstract right of freedom, by supposing that each individual has a 'tacit compact' or understanding with the rest of his community. But this 'compact' or 'understanding turns out to be too obviously fictitious to be put forward as a basis for moral duty—witness the endlessly various ways in which its friends have modified it. Many of them hold that the only abstractly justifiable social order is one where no laws are imposed without the *explicit* consent of those who are to obey them. But we couldn't construct society on this basis; and such representative governments as *have* been established appear to embody this idea ·of explicit agreement· by means of sweeping limitations and transparent fictions. And it became obvious also that the. . . .the most perfect conformity between a government's actions and the wishes

[1] Many people seem to hold that wealth is rightly distributed when cultivated persons are wealthy and the uncultivated have just enough to live on, because the former are better able to get happiness from wealth than the latter.

of the majority of its subjects need by no means result in the greatest civil freedom in the society so governed.

But even if we could construct a satisfying ideal social order, including an ideal form of government, we would still have to reconcile •the duty of bringing this about with •the duty of conforming to society's actual order. We are strongly convinced that positive laws ought, generally speaking, to be obeyed; and our notion of justice seems to include a general duty to satisfy the expectations created by custom and precedent. But if the actual order of society deviates much from what we think ought to exist, the duty of conforming to it becomes obscure and doubtful. And common sense doesn't regard it as an axiom that laws ought to be obeyed. Indeed, everyone agrees that they ought to be disobeyed when they command what is wrong; though we can't extract from common sense any clear general view as to what remains wrong after it has been commanded by the sovereign. Also, the positive laws that ought to be obeyed because they are the law must be commands issued by a (morally) rightful authority; and these won't *always* coincide with legally enforced commands, because the law-courts may be temporarily subservient to a usurper. And again: a sovereign who has habitually been obeyed may be one whom it has become right to rebel against (it's generally admitted that this is sometimes right). So we need principles for settling *when* usurpation becomes legitimate and *when* rebellion is justifiable; and they aren't forthcoming from common sense; though we can say that common sense leans more towards the utilitarian method on this topic of sovereignty than it does on matters of private morality.

We're even further from being able to state the general duty of satisfying 'natural expectations'—i.e. ones that an average man would form in the given circumstances—in the form of a clear and precise moral axiom. No doubt a just man will generally satisfy customary claims; but it can't be maintained that the mere existence of a custom gives each person a clear obligation to conform to it if he hasn't promised to do so; especially because bad customs can only be abolished by individuals venturing to disregard them.

6. We have still to examine (whether as a branch of justice or under a separate heading) the duty ·discussed in **chapter 6**, namely that· of fulfilling explicit promises and distinct understandings. The special confidence in this principle that moralists have generally felt is strikingly illustrated by the attempts to extend its scope that I have just mentioned; and it does surpass in simplicity, certainty, and definiteness the moral rules I have discussed so far. So this seems to be our best chance of finding one of the ethical axioms we are searching for. Now, we saw that the notion of a *promise* needs to be made precise with several details that aren't commonly thought about; but that doesn't rule out the possibility that the notion can be used in forming a maxim that will, once it has been stated and understood, be accepted by everyone as self-evident. The uneducated majority couldn't define a *circle* as *a figure bounded by a line all of whose points are equidistant from the centre*; but when they meet this definition, they'll accept it as perfectly expressing the notion of roundness that they've always had in their minds. And I think that this sort of *potential universality of acceptance* can be claimed for the propositions that

- •to be binding a promise must be understood by promiser and promisee in the same sense at the time of promising;
- •a promise is relative to the promisee and can be annulled by him; and
- •a promise can't override determinate prior obligations.

But it's not like that with the other qualifications that we had to discuss. Should we include *them* in the notion

of *promise*?—in answering this common sense splits. When (for example) we ask: How binding is a promise if

- •it was made in consequence of false statements, though it wasn't understood to be conditional on their truth?
- •important circumstances were concealed?
- •we were somehow led to believe that the consequences of keeping the promise would be different from what they turn out to be?
- •the promise was given under compulsion?
- •circumstances have relevantly altered since the promise was given?
- •we find that the results of keeping the promise will be different from what we foresaw when we promised?
- •we now see that keeping the promise will involve a sacrifice out of proportion to the benefit received by the promisee?
- •we now see that the keeping the promise will be harmful to him, though he doesn't think so?

different conscientious people would answer these and other questions (both generally and in particular cases) in different ways. We might get a decided majority for some of these qualifications and against others, but there wouldn't be a clear consensus about any of them. Furthermore, the mere discussion of these points shows that the confidence with which the 'unsophisticated conscience' asserts unreservedly that *promises ought to be kept* is due to carelessness, and that when the above questions are fairly considered this confidence changes into hesitation and perplexity. Some of the discussion of these questions suggests that our principle about promising is a special case of more comprehensive utilitarian principle.

And our distrust of what common sense ordinarily says about promise-keeping is reinforced when we think about *where* it belongs in a classified system of moral obligations. We saw that fidelity to promises is ranked with veracity, as though the mere fact of my having said that I would do something gave me a duty to do it. But on reflection we see that the obligation arises from the reliance that someone else has placed on my assertion; that the breach of duty is constituted by the disappointment of expectations that I have voluntarily raised. And when we see this, we become less sure of the absoluteness of the duty; it now seems to depend on how much harm is done by disappointing expectations; and if keeping the promise would involve an amount of harm that thoroughly outweighs the harm brought by disappointment, we shrink from saying that the promise ought to be kept.

The case of veracity (discussed in **chapter 7**) can be dismissed more briefly, because it was even easier to show that the common statement of the unqualified duty of truth-speaking is made thoughtlessly, and can't be accepted by a reflective mind as an absolute first principle. Firstly, we found no clear agreement about the basic nature of the obligation; or about its exact scope—i.e. are we obliged to do what we can to make the facts fit

- •our actual affirmation as understood by the hearer?
- •whatever inferences we the hearer is likely to draw from this?
- •both?

To achieve perfect candour and sincerity, we must aim at both; and of course we do admire the exhibition of these virtues; but few people will maintain that they *ought* to be exhibited at all times. Secondly, common sense seems to admit, though vaguely and reluctantly, that the veracity principle, however defined, doesn't hold universally; at any rate it isn't thought to be clearly wrong to tell untruths to children, madmen, invalids, enemies, robbers, or even

persons who ask questions they have no right to ask (if a mere refusal to answer would practically reveal an important secret). And when we consider the generally admitted limitations, it seems even clearer than it was with promising that they are commonly determined by implicit or explicit utilitarian reasonings.

7. Given that common sense's versions of the prescriptions of justice, promise-keeping and truth-telling can't be converted into first principles of scientific ethics, it's hardly necessary to inquire whether such axioms can be extracted from •such minor maxims of social behaviour as the maxim of liberality or the rules restraining the malevolent affections. Indeed, it became clear in **chapter 8** that common sense's only escape from inconsistency or hopeless vagueness concerning the proper regulation of resentment is to adopt the 'interests of society' as the ultimate standard. Similarly with such virtues as courage and humility, which in **chapter 10** we couldn't classify as either social or self-regarding. We can't definitely distinguish courage from foolhardiness except in terms of •probability that the daring act will promote the well-being of the agent or of others, or of •some definite rule of duty prescribed under some other notion.

It is true that among what are commonly called 'duties to self' we find the duty of self-preservation prescribed with apparent absoluteness, at least so far as the sacrifice of one's life is not imperatively required for the preservation of the lives of others, or for the attainment of some result conceived to be very important to society. ·I discussed this in **chapter 9**·. But I think that when common sense is confronted with the question:

> Is a man obliged to stay alive in he can foresee that the life remaining to him will be miserable ·for him· and burdensome to others—e.g. if he has a fatal disease that make it impossible for him to do work of any kind during the weeks or months of agony that remain to him?

although it (common sense) would deny the legitimacy of suicide even under these conditions, it would admit that it needed to find reasons for that denial. This implies that the universal wrongness of suicide is not self-evident. And I think the reasons that would be found—other than ones derived from revealed religion—would turn out to be broadly utilitarian. It would be urged that if any exceptions to the prohibition of suicide were allowed, that would encourage the suicidal impulse in other cases where suicide would really be a weak and cowardly failure of social duty; it would also probably be urged that the toleration of suicide would make secret murders easier. Thus, in this case as in the others the independent axiom we are searching for seems to disappear when we look closely.

Reflection seems also to show that the duties of temperance, self-control, and other related virtues are clear and definite only when conceived as subordinate either to prudence (the usual case) or to benevolence or some definite rule of social duty, or at least to some end—such as 'furtherance of moral progress'—the conception of which involves the notion of some duty that is supposed to be already determinate. The authority of common sense can't be claimed for *any* restriction even of the bodily appetites for food and drink unless it is thus subordinated.

For the sexual appetite, however, a special regulation seems to be prescribed on the basis of some independent principle under the notion of purity or chastity. When I examined this notion in chapter 9 we saw that common sense, far from being explicit about this, is actually averse to explicitness about it. Because my primary aim back there was to give a faithful exposition of the morality of common sense, I allowed my inquiry to be brought to a halt by this (as

it seemed) clearly recognisable sentiment. But now that my primary purpose is to test whether the commonly accepted moral principles are intuitively evident, I have to over-ride this aversion. My question is this:

> Is it possible to acquire rational conviction regarding the acts allowed or forbidden under the notion of purity (or chastity) and its opposite?

And I can't answer that without subjecting *purity* to the same close scrutiny I have tried to give to the other leading notions of ethics. Here I'll only need to give the briefest account of such a scrutiny. I am aware that in giving even this I'm sure to cause a certain offence to minds trained in good moral habits; but I claim the same permission that is commonly granted to the anatomy-teacher who also has to direct the student's attention to objects that a healthy mind naturally prefers not to contemplate.

8. Well, then, what is the conduct that *purity* forbids? (The principle is easier to discussed in its negative aspect.) Because the normal and obvious end of sexual intercourse is the propagation of the species, some people have thought that

- All sexual activity that isn't a means to procreation should be prohibited.

But this restricts conjugal intercourse much too severely for common sense. Shall we then say this?

- Purity forbids sexual activity that isn't a means to procreation except between legally married couples.

On reflection this turns out to be unsatisfactory also. For one thing: we might condemn the conduct of a couple who deliberately omit to fulfill legal conditions and make a contract that the law declines to enforce, but we shouldn't call their union 'impure'. And, secondly, we feel that positive law ought to maintain *purity* but sometimes doesn't (being like *justice* in both those respects). Then what kind of sexual relations are we to call essentially impure, whether or not sanctioned by law and custom? There seem to be no clear principles with any claim to self-evidentness that could give a generally acceptable answer to this question. It would be hard to state such a principle even for settling what degree of blood-relatedness between husband and wife makes their union incestuous, though the general aversion to incest is a specially intense moral sentiment; and we would find it even harder—indefinitely much harder—to find a rationale for the prohibited degrees of blood-relatedness. [In what follows: 'polygyny' = one man with two or more wives; 'polyandry' = one woman with two or more husbands.] Another problem: probably few people would regard a legal polygynous connection as impure, even if they disapprove of the law and state of society that allows it; but if legal polygyny is not impure, is polyandry impure? (I mean legal and customary polyandry, which is fairly common among the lower races of man.) If it isn't, then on what rational principle can the notion be applied to institutions and conduct? [How does that question relate to what precedes it? The puzzle about this is there in the original; it isn't an artifact of this version.] And another: Where divorce by mutual consent, with subsequent marriage, is legalised, we don't call this an offence against purity; but once we allow that freely changing partners is allowed in principle, it seems paradoxical to distinguish purity from impurity merely by how slowly the change is made,[1] and to condemn as impure even 'free love' when it is earnestly advocated as a means not to mere sensual license but to a completer harmony of sentiment between men and women.

[1] I'm not challenging anyone to say exactly *how* slowly; I'm asking whether we can really think that the decision depends at all on considerations of this kind.

We might then fall back on *mutual affection* (as distinguished from mere appetite) as constituting the essence of pure sexual relations. But this, while too •lax from one point of view, seems from another point of view to be too •severe for common sense. We don't condemn loveless marriages as *impure*, although we disapprove of them as productive of unhappiness. Such marriages are indeed sometimes criticised as 'legalised prostitution', but we feel that phrase to be extravagant and paradoxical; and it's not even clear that we even *disapprove* of loveless marriages under all circumstances—consider the case of royal alliances.

How are we to judge such institutions as those of Plato's Commonwealth, where women and children were to be held in common, while sexual indulgence would be regulated with strict reference to social ends? Our habitual standards seem to get no grip on such novel circumstances.

In fact, reflection on the current sexual morality reveals *two* bases for it: **(a)** first and chiefly, the maintenance of a social order that is believed to be best for the prosperous continuance of the human race; **(b)** secondly, the protection of individuals' habits of feeling that are believed to be generally most important to their perfection or their happiness. It's commonly believed that both these ends can be achieved by the same regulations, and in an ideal state of society perhaps they would be; but in actual life there's often a partial separation and incompatibility between them. And in any case if the repression of sexual license is prescribed merely as a means to these ends, the claim that it's needed for them can't be self-evident—it would have to be supported empirically. **(a)** We can't reasonably be sure, without evidence from sociological observations, that a certain amount of sexual license will be incompatible with maintaining the population in sufficient numbers and good condition. **(b)** And although it's certainly clear that someone whose sexual relations are of a merely sensual kind misses the highest and best development of his emotional nature, we can't know *a priori* that this lower kind of relation interferes with the development of higher ones (and indeed experience doesn't show that it always does). And the **(b)** line of thought has a further difficulty. [What follows breaks away from Sidgwick's rather clotted wording, but its content is precisely his.] There is

(i) a life of celibacy,
(ii) indulgence in sexual freedom purely for sensual pleasure, and
(iii) the development of relations higher than those in **(ii)**.

We may despise someone who goes in for **(i)** (unless he does so for some noble end) but we don't *blame* him. If we're going to blame someone who goes in for **(ii)** because that isn't good for **(iii)**, why wouldn't we also blame **(i)** for the same reason?

9. I could say much more about the perplexities we encounter in trying to define the rule of purity or chastity. But I don't want to extend the discussion beyond what is needed to complete my argument. The conclusion announced in the paragraph 'It seems clear. . . ' on page 165 has now been sufficiently justified. We examined the moral notions that present themselves with a *prima facie* claim to provide independent and self-evident rules of morality; and in each case we found that from the rules of conduct that common sense really supports we can't derive any proposition that even *appears* to qualify as a scientific axiom. [Sidgwick adds, regarding common sense's attempts to co-ordinate its principles, that he has already discussed these fully enough. He has shown that instead of •co-ordination there is often •collision; this requires borders; and common sense's drawing of those tends to be vague and inconsistent.]

. . . .Nothing that I have said even *tends* to show that we don't have distinct moral impulses—claiming authority over

all others, and prescribing or forbidding kinds of conduct—concerning which there's fairly general agreement among educated people of the same time and country. My thesis is only that the objects of these impulses can't be scientifically determined by any reflective analysis of common sense. The notions of benevolence, justice, good faith, veracity, purity etc. don't lose their significance for us just because we can't define them with precision. The main part of the conduct prescribed under each notion is clear enough; and the general rule prescribing it doesn't necessarily lose its force because in each case •there's a margin of conduct involved in obscurity and perplexity, or •the rule turns out not to be absolute and independent. In short, the morality of common sense may still be perfectly adequate to give practical guidance to common people in common circumstances; but the attempt to elevate it into a system of intuitional ethics highlights its inevitable imperfections without helping us to remove them.[1]

Chapter 12: Motives or springs of action as subjects of moral judgment

1. I was careful in chapter 1 to point out that our common moral judgments concern motives as well as intentions. Indeed, reflecting on motives and judging them to be good or bad is a prominent element in our notion of *conscientiousness*. So if I am to complete my examination of the intuitional method I need to •consider this comparison [Sidgwick's word] of motives and •discover how far it can be made systematic and pursued to conclusions of scientific value. This is a convenient place for treating of this part of the subject: an important school of English moralists has maintained that desires and affections, rather than actions, are the proper subjects of the ethical judgment; and this is a natural fall-back position when systematic reflection on the morality of common sense has shown us how hard it is to get a precise and satisfactory account of rightness and wrongness in actions.

The term 'motive' is commonly used in two ways. The 'motive' of an action may be

(a) those of its foreseen consequences that the agent desired in willing; or

(b) the desire or conscious impulse that led to the action.

The two meanings do in a way correspond, because whenever **(b)** the impulses are different there must always be some **(a)** difference in their respective objects. But for our present purpose **(b)** is more convenient to work with: what we have to deal with in our lives is our own impulsive nature, controlling, resisting, indulging the different impulses; so what we primarily have to estimate is the ethical value of *these*; and we often find that two psychologically very different impulses are directed towards essentially the same end, though regarded from different points of view. For example, a man could be driven by •appetite or by •rational

[1] The more positive treatment of common-sense morality in its relation to utilitarianism, to which I shall proceed in IV/3, is an indispensable supplement to the negative criticism that I have just completed.

self-love to seek a particular sensual gratification. . . . In this chapter, then, I shall use 'motive' to refer to *the desires for particular results that we think we can achieve through voluntary acts, by which desires we are stimulated to will those acts*. [There's a footnote here describing and criticising Green's way of distinguishing 'motives' from 'desires'.]

The issue in any internal conflict is not usually thought to be between •good and •bad motives but between •better and •less good. . . .motives. If there's any kind of motive that we commonly judge to be *intrinsically* bad in any circumstances, it is malevolent affection [see Glossary], i.e. the desire—however aroused—to inflict pain or harm on some other sentient being. And reflection shows (as we saw in chapter 8) that common sense doesn't condemn even this kind of impulse as absolutely bad, because we commonly recognise the existence of 'legitimate resentment' and 'righteous indignation'. Moralists try to distinguish

> •anger 'against the act' from •anger 'against the agent' and
>
> •the impulse to inflict pain from •the desire for the pleasure one will get from inflicting pain;

but isn't it beyond the capacity of ordinary human nature to maintain these distinctions in practice? At any rate common sense doesn't condemns as absolutely bad any motives other than deliberate malevolence. The other motives that are commonly criticised seem to be *seductive* (Bentham's word) rather than *bad*. That is, they prompt men to forbidden conduct with conspicuous force and frequency; but careful thought shows us that there *are* certain narrow limits within which their operation is legitimate.

Our common judgments about which kinds of motives are better than which seem to lay claim to some intuitive knowledge about this; and our present question is: How far does this intuitive knowledge satisfy the conditions laid down in chapter 11? I argued in chapter 1.2 that this comparison of motives is not the **normal form** of our common moral judgments, and I see no reason to think that it is their **original form**. I think that in the normal development of man's moral consciousness—both in the individual and in the race—moral judgments are first passed on outward actions, and that motives don't come to be definitely considered till later. . . . But that doesn't rule out the thesis that the comparison of motives is the **final and most perfect form** of moral judgments. It might qualify for that status if it were true that

> the comparison of motives, when pursued by different thinkers independently, leads to results that are clear and mutually consistent; and it is free from the puzzles and difficulties that beset other developments of the intuitional method.

·But that is not how things stand·. When we try carefully to arrange motives in order of excellence, we encounter versions of •many (if not all) of the difficulties we met when discussing of the commonly received principles of conduct, and •other difficulties that we didn't meet back then. And when intuitive moralists try to overcome these difficulties they turn out to differ from one another •more, and •more deeply, here—in the rank-ordering of motives—than we found them to do over the rightness of actions.

2. ·The **inclusion problem**·: Are we to include in our list of motives the moral sentiments, i.e. impulses towards particular kinds of virtuous conduct as such, e.g. candour, veracity, fortitude. It seems wrong to exclude them, because we observe them as distinct and independent impulses in most well-trained minds—sometimes as intense impulses, as when we describe a man as 'enthusiastically brave' or 'intensely veracious' or 'having a passion for justice'. But including them confronts us with a dilemma, starting from

the question *Are the objects of these impulses represented by the very notions that we have been examining?*

If they are, then after we have decided that impulse x is better than impulse y, all the perplexities set forth in chapters 3–10 will recur, before we can act on our decision. What's the use of recognising the superiority of the impulse to do *justice* if we don't know what it is just to do?

If they aren't, because the objects they prompt us to realise are conceived more simply, without the complexities that our complete reflection on common sense forced us to recognise, then certainly won't find agreement about the relations amongst these impulses. For example, ought we to follow the impulse to speak the truth? There will inevitably be disputes about this when veracity seems opposed to the general good or to the interests of some person—i.e. when it conflicts with 'universal' or 'particular' benevolence. Hutcheson explicitly ranks these benevolent impulses higher than candour, veracity, and fortitude, reserving the highest moral approval for 'the most extensive benevolence' or 'calm, stable, universal goodwill to all'.[1] But this view, which in practice works out as equivalent to utilitarianism, would certainly be disputed by most intuitional moralists. And some of these moralists (such as Kant) hold that no action is good unless it is motivated by a pure regard for duty, a pure choice of right as right; whereas Hutcheson, who stands at the opposite pole of intuitional ethics, also identifies the love of virtue as a separate impulse, but he treats it as having the same rank and the same effects as universal benevolence.

Moralists also diverge widely in their views about the ethical value of self-love. Butler seems to regard it as one of two superior and naturally authoritative impulses, the other being conscience; indeed, in a passage quoted on page 55, he even concedes that it would be reasonable for conscience to yield to it if the two could possibly conflict. Other moralists (and Butler in the appendix on virtue to *The Analogy of Religion*) appear to place self-love among virtuous impulses under the name of 'prudence' but to rank it rather low among them, and would have it yield to nobler virtues in cases of conflict. Others exclude it from virtue altogether; e.g. Kant in his one of his works says that the end of self-love, one's own happiness, can't be an end for Moral Reason, and that the force of the reasonable will, which is what virtue consists in, is always exhibited *in resistance to* natural egoistic impulses.[2]

Martineau, whose system is built on the basis that I am now examining, tries to avoid some of the difficulties I have just pointed out by refusing to admit that there are any virtuous impulses except the 'preference for the superior of the competing springs of action in each case' of a conflict of motives. In his *Types of Ethical Theory* he writes:

> 'I can't admit either the *loves of virtues*—of candour, veracity, fortitude—or the virtues themselves, as so many additional impulses over and above those from the conflict of which they are formed. I don't confess my fault *in order to he candid*. . . unless I am a prig. I never think of candour as something that is or will be predicable of me at all.'

I'm not sure whether he •really means to deny that anyone ever acts from a conscious desire to realise an ideal of candour or fortitude, or •merely means to express disapproval of acting with such a motivation. In the former sense his

[1] Francis Hutcheson, *System of Moral Philosophy* I/4.10

[2] That is in Kant's *Doctrine of Virtue* ·which is part of his *Metaphysic of Morals*·. The ethical view briefly expounded in the *Critique of Pure Reason* appears to be much more like Butler's.

statement seems to me a psychological paradox, in conflict with ordinary experience; in the latter sense it seems to be an ethical paradox—a striking example of the diversity of judgments about the ranking of motives.

3. ·The **rank-ordering problem**·: Even if we set aside the moral sentiments and self-love, we still won't be able to construct a scale of motives, arranged in order of merit, for which there'll be anything like a clear consent even of cultivated and thoughtful persons. We do seem to be generally agreed on one or two points—e.g. that bodily appetites are inferior to benevolent affections and intellectual desires, and perhaps that impulses to attend to our own individual well-being rank below impulses that we class as other-regarding or disinterested. But a few vague statements of this kind are about as far as we can go. For example, when we compare •personal affections with •the love of knowledge or of beauty or •the passion for any kind of ideal, much doubt and divergence of opinion show up. There isn't even agreement about the relative ranks of the benevolent affections taken by themselves: some people prefer the more intense though narrower affections, while others would prefer the calmer and wider feelings. And there's disagreement about love. As I said in on page 115, love is a complex emotion which commonly includes not just •a desire for the good or happiness of the beloved but also •a desire for union or intimacy of some kind; and people disagree about which of these contributes more to the value, the rank, of a given instance of love.

The love of fame is an important and widely operative motive that would be ranked differently by different persons. Some would rank this 'spur that the clear spirit doth raise' higher than any but the moral sentiments, while others think it degrading to depend on popular favour for one's happiness. [The 'spur' phrase is from Milton's *Lycidas*.]

Also, the more we look into the actual promptings that precede any volition, the more we find *complexity of motive* to be the rule rather than the exception, at least in educated persons; and this composition of impulses creates a basic perplexity regarding the principles on which our decision is to be made, even if we are clear about the relative worth of the elementary impulses. That's because the compound usually contains higher and lower elements, and we can't get rid of the lower ones; as I said on page 95, though we can often suppress and expel a motive by firmly resisting it, it doesn't seem possible to exclude it if we perform the action that it prompts us to do. Suppose we are impelled

- •in one direction by a combination of high and low motives, and
- •in another direction by an impulse that ranks between those two in the scale.

How shall we decide which course to follow? It's not an uncommon situation. Examples:

- •An injured man is pushed towards punishing the offender by a regard for justice and a desire for revenge, and towards sparing him by an impulse of pity.
- •A Jew of liberal views is moved to eat pork by a desire to vindicate true religious liberty combined with a liking for pork, and restrained from eating it by a desire not to shock the feelings of his friends.

How are we to deal with such cases? You won't suggest that we should estimate the relative proportions of the different motives and decide accordingly! Qualitative analysis of our motives is to some extent possible to us, but the quantitative analysis that this would require is not in our power.

Apart from this difficulty arising from complexity of motives, we can't assign a definite and constant ethical value to each kind of motive without reference to •the circumstances under which it has arisen, •the extent of

indulgence that it demands, and •the consequences to which this indulgence would lead in any particular case. I can illustrate this by reference to Martineau's table of springs of action arranged in descending order of merit.

(1) Primary sentiment of reverence.
(2) Primary affection of compassion
(3) Primary affections, parental and social; with (approximately) generosity and gratitude.
(4) Primary sentiments of wonder and admiration.
(5) Secondary sentiments: love of culture.
(6) Causal energy: love of power, love of liberty.
(7) Primary passions: antipathy, fear, resentment.
(8) Secondary affections (sentimental indulgence of sympathetic feelings).
(9) Love of gain (reflective derivative from appetite).
(10) Primary animal propension: spontaneous activity (unselective).
(11) Primary organic propensions; appetites.
(12) Secondary organic propensions; love of ease and sensual pleasure.
(13) Secondary passions: censoriousness, vindictiveness, suspiciousness.

This scale seems to me open to much criticism, from the points of view of psychology and of ethics;[1] but, granting that it corresponds broadly to the judgments that men commonly pass as to the different elevatedness of different motives, it seems to me utterly paradoxical to say that each class of motives is *always* to be preferred to the class below it, without regard to circumstances and consequences. So far as it's true that 'the conscience says to everyone, "Don't eat till you are hungry and stop when you are hungry no more"', it is not because a 'regulative right is clearly vested in primary instinctive needs, relatively to their secondaries', but because experience has shown that it is usually dangerous to one's health to gratify the palate when one isn't hungry, and it's in view of this danger that the conscience operates. If we condemn a ship's captain who, 'caught in a fog off a lee shore, neglects through idleness and love of ease to slacken speed and take cautious soundings and open his steamwhistle', it's not because we intuitively discern **(7)** fear to be a higher motive than **(12)** love of ease, but because we judge that the consequences he is disregarding are much more important than the gratification obtained; in a case where fear was not in this way backed up by prudence, our judgment would certainly be different. Common sense holds rather that most natural impulses have their proper spheres within which they should normally operate, so that a question of the form 'Should motive x yield to motive y?' can't be answered decisively in the general way in which Martineau answers it; the right answer depends on the particular circumstances of the particular case. We see that each of these is possible:

- •a motive that we commonly rank as higher wrongly intrudes into the proper sphere of one that we rank as lower;
- •a lower motive invades the sphere of a higher one.

It's only because the former is much less likely that it naturally falls into the background in ethical discussions and exhortations that have a practical aim. And another complication: as the character of a moral agent improves, the motives we rank as 'higher' tend to be developed, so that their normal sphere of operation is enlarged at the expense of the lower. So in moral regulation and culture relating to motives there are two aims: **(i)** to keep the 'lower' motive

[1] For example, why is the class of 'passions' so strangely restricted? Why is conjugal affection omitted? Is wonder really a definite motive? Is it right to rank 'censoriousness' with 'vindictiveness' as one of the 'lowest passions'? And so on.

within its proper sphere so long as we can't substitute for it the equally effective operation of a higher motive; and **(ii)** to substitute 'higher' motives for 'lower' ones •gradually and •as far as we can do it without danger, up to some limit. We can't specify the limit, but certainly think of it—for the most part—as falling short of completely supplanting the lower motive.

To illustrate this I'll return to the passion of resentment. Reflective common sense holds that •this malevolent impulse, as long as it is limited to resentment against wrong and operates in aid of justice, has a legitimate sphere of action in the social life of human beings; that •suppressing it would do great harm unless we could intensify the ordinary man's concern for justice or for social well-being so that the total strength of motives prompting to the punishment of crime wouldn't be lessened. No doubt it is 'to be wished', as Butler says [see footnote on page 155] that men would repress wrong from these higher motives rather than from passionate resentment; but we can't hope to change human beings in this way except by a slow and gradual process of elevation of character; so granting that there's a conflict between **(2)** compassion and **(7)** resentment, it is *not* the case that as a general rule compassion ought to prevail. We ought rather—with Butler—to regard resentment as a valuable 'balance to the weakness of pity', which would be liable to prevent the carrying out of justice if resentment were excluded.

Or consider the impulse that comes lowest (among those not condemned altogether) in Martineau's scale, namely **(12)** the 'love of ease and sensual pleasure'. No doubt this impulse. . . .continually leads men to shirk their strict duty or not perform it thoroughly, or in some less definite way to fall below their own ideal of conduct; which is why preachers and practical moralists habitually argue for its repression. Yet common sense recognises cases where even this impulse ought to prevail over impulses ranked above it in Martineau's scale; we often find men prompted—for example by 'love of gain'—to shorten *unduly* their hours of recreation; and such a case we would judge it best that victory should remain on the side of the **(12)** 'love of ease and pleasure' and that the encroachment of **(9)** 'love of gain' should be repelled.

But I don't think that in either of these cases the conflict of motives would *stay* as I have described it. The struggle might *begin* as a duel between resentment and compassion, or between love of ease and love of gain, but it wouldn't be *fought out* in those terms. As the conflict went on, the higher motives would inevitably be called in:

- •regard for justice and social well-being on the side of resentment,
- •regard for health and long-term efficiency for work on the side of love of ease;

and these intervening higher motives would decide the struggle, so far as it was decided rightly and as we should approve. That is certainly what would happen in my own case if the conflict were at all serious and its decision deliberate; and this is my final reason for holding that a scale like Martineau's, arranging motives according to their moral rank, can't have more than a very subordinate ethical importance. I admit that it may give a rough indication of the kinds of desires that it's ordinarily best to encourage and indulge, as compared with other kinds that are likely to collide with them; and we might use this for quick settlements of the trifling conflicts of motive that the varying and complex play of needs, habits, interests, and their accompanying emotions continually arouses in our daily life. But when a serious question of conduct arises I can't conceive myself deciding it morally by comparing motives below the highest. The question has to be submitted to the decision of the court

of whatever motive we regard as supremely regulative; so that the finally decisive comparison wouldn't be •between the lower motives primarily conflicting, but •between the effects of the different lines of conduct to which these lower motives prompt, considered in the light of whatever we regard as the ultimate ends of reasonable action. And I think this will be the course naturally taken not only by utilitarians but by everyone who follows Butler in regarding our passions and propensions as forming naturally a 'system or constitution' in which the •ends of lower impulses are subordinate as •means to the •ends of certain governing motives or are taken to be parts of these larger ends.

Chapter 13: Philosophical intuitionism

1. Is it then impossible to reach, by a deeper and sharper examination of our common moral thought, real ethical axioms—intuitive propositions that are really clear and certain?

This question leads us to the third phase of the intuitive method that I called 'philosophical intuitionism' in I/8.4 [see page 47]. We think of a philosopher as trying to do more than merely define and formulate the common moral opinions of mankind. His function is to tell men not •what they do think but •what they ought to think; he is expected to go beyond •common sense in his premises, and is allowed some divergence from •it in his conclusions. There are limits—though not well defined ones—to how far he is allowed to deviate. The truth of his premises will be tested by the acceptability of his conclusions; and if he is found to be in flagrant conflict with common opinion on any important point, his method is likely to be declared invalid. Still, though he is expected to establish and put together the main part of the commonly accepted moral rules, he isn't obliged to take them as the basis for his own system. His task is to state in full strength and clarity the primary intuitions of reason which can, handled scientifically, systematised and correct the common moral thought of mankind. At any rate, that's what we would expect orthodox thinkers to be trying to do throughout the history of moral philosophy.

To some extent that's what has happened. But moral philosophy—i.e. philosophy as applied to morality—has also been engaged in other tasks that are even harder than that of penetrating to the basic principles of duty.

- •In modern times especially, it has accepted the task of proving that someone's having a duty doesn't conflict with ·self-·interest, i.e. his happiness or welfare.
- •It has also tried to determine how right or good generally relates to the actual world, a task that could hardly be well done without an adequate explanation of the existence of evil.
- •It has been further distracted by questions—ones that I think belong to psychology rather than to ethics—about the 'innateness' of our notions of duty, and the origin of the faculty that provides them.

While concentrating on these difficult subjects, each of which has been mixed up with the discussion of basic moral

intuitions, philosophers have too easily •settled for ethical formulae that implicitly accept the morality of common sense *en bloc*, ignoring its defects; and •merely expressed some view about how this morality relates to the individual mind or to the actual universe. Perhaps they have also been hampered by the reasonable fear of losing the support of 'general assent' if they adopt for themselves and their readers too rigid a standard of scientific precision. In spite of all this, however, philosophers *have* provided us with a number of comprehensive moral propositions that they have advanced as certain and self-evident—ones that at first sight seem fit to be the first principles of scientific morality.

2. A warning that I have already given needs to be especially stressed here. Beware of a certain class of sham-axioms that are apt •to offer themselves to a mind seeking for a philosophical synthesis of practical rules, and •to delude the unwary through their appearance of being clearly self-evident. They appear certain and self-evident because they are basically tautological [see Glossary]: examine them and you'll find that they only say things of the form:

> It is right to do what is—in a certain department of life, under certain circumstances and conditions—right to be done.

The history of moral philosophy shows that. . . .even powerful intellects are liable to accept ·and value· such tautologies—sometimes expanded into circular reasonings, sometimes hidden in the depths of an obscure notion, often lying so near the surface that it's hard to understand how they could ever have been thought important.

Look for example at the time-honoured 'cardinal virtues'. If we're told that the dictates of wisdom and temperance can be summed up in these clear and certain principles:

> •It is right to act rationally (wisdom),
>
> •It is right that the lower parts of our nature should be governed by the higher (temperance),

we may at first feel that we're getting valuable information. But when we find that 'acting rationally' means the same as 'doing what we see to be right' (see page 166), and that the 'higher part' of our nature is explained as being *reason*. . . ., it becomes obvious that these 'principles' are tautologies. Another example: the principle of justice that *we ought to give every man his own*, this looks plausible until we find that we can't define 'his own' except as = 'that which it is right he should have'.

The 'principles' I have quoted can be found in modern writers, but it's worth noting that throughout the ethical speculation of Greece[1] we are offered tautological universal propositions about virtue or good conduct—or propositions that can be defended from the charge of tautology only by being taken as •definitions of the problem to be solved rather than as •attempts at its solution. For example, Plato and Aristotle appear to offer, as constructive moralists, the scientific knowledge of the good and bad in human life that Socrates said no-one has. They seem to agree that such good as can be realised in the lives of men and communities is chiefly *virtue* or (as Aristotle more precisely puts it) the *exercise of virtue*. . . . But how can we discover what kind of conduct counts as virtuous? The only answer we seem to get from Plato is that a given virtue consists in

[1] I'm aware of the special interest and value of ancient Greek ethical thought. Through much of the present work the influence of Plato and Aristotle on my discussions has been greater than that of any modern writer. But my topic here is just the ancient systems' offerings of general principles for determining what ought to be done.

- •knowledge of what is good in certain circumstances and relations, and
- •a harmony among the elements of man's appetitive nature, such that their resultant impulse always conforms with this knowledge.

But this knowledge (or at least its principles and method) are just what we wanted him to give us! We won't be satisfied by a mere account of the different situations in which we need it. And Aristotle doesn't bring us much closer to such knowledge when he tells us that the good in conduct is to be found somewhere between different kinds of bad. This at most tells us the *whereabouts* of virtue; it doesn't give us a method for finding it.

On the Stoic system as constructed by Zeno and Chrysippus[1] it may be unfair to pass a final judgment on the basis of the accounts given of it by adversaries like Plutarch, and semi-informed expositors such as Cicero, Diogenes Laertius, and Stobaeus. But, as far as we can tell the exposition of its general principles is a complicated network of circular reasonings by which the inquirer is continually deluded with an apparent approach to practical conclusions and continually led back to the point from which he set out.

Stoicism's most characteristic doctrine says that the ultimate end of action is 'life according to nature'. The motion that sustains this life spring was driven, according to the Stoics,

- •in the vegetable creation by a mere unfelt impulse;
- •in animals by an impulse accompanied by sensation,
- •in man by the direction of reason, which in him is naturally supreme over all blind impulses.

What directions does reason give? 'To live according to nature' is one answer, which gives us the circular exposition of ethical doctrine in its simplest form. Sometimes the answer is 'Live according to virtue', which leads us into a circle like—though actually neater and tighter than—the one we found in Plato and Aristotle. That is because the Stoics also defined virtue as knowledge of good and bad in different circumstances and relations; and whereas Plato and Aristotle held that the notion of •virtue is the chief component in the notion of •good for human beings, the Stoics thought that those two notions absolutely coincide. So this is the upshot: virtue is knowledge of what is good and ought to be chosen (or bad and ought to be rejected); and there's nothing good (or bad) except virtue (or vice). But that makes virtue a science that studies nothing but itself, emptying the notion of all practical content. To avoid this result and reconcile their system with common sense, the Stoics explained that some other things in human life are •preferable, in a way, though not strictly •good—including the primary objects of men's normal impulses. Well then, when our impulses are conflicting or ambiguous, how are we to choose among them?. . . . The Stoics' only answers were 'We should choose what it reasonable' or 'We should act in accordance with nature'; and those answers obviously bring us back to different points in the original circle.[2]

[1] I am not aiming here at the *later* Stoicism of the Romans—Seneca and Marcus Aurelius—in which the individual man's relation to humanity generally is more prominent than it is in the earlier form of the system.

[2] The Stoics sometimes tried to get ethical guidance from what is 'natural' in a different sense, steering by the complex fit between means and ends in the organic world. But since they held that the whole course of the universe is perfect and completely predetermined, they couldn't get from any observation of the facts any clear and consistent principle for selecting among alternatives of conduct. Their most characteristic practical precepts show a curious conflict between two tendencies: •to accept what is customary as 'natural', and •to reject what seems arbitrary as unreasonable.

Butler in his use of the Stoic formula seems to avoid this circular reasoning, but only by ignoring or suppressing the intrinsic reasonableness of right conduct. Butler assumes that •it is reasonable to live according to nature, and argues that •conscience—the faculty that imposes moral rules—is naturally supreme in man; from which he infers that •it is reasonable to obey conscience. But what's the status of the rules that conscience lays down? Do we know them merely **as the dictates of arbitrary authority**? If so, Butler would be giving absolute authority to the possibly unenlightened conscience of the individual; and he is much too cautious to do this. In fact, in more than one passage he explicitly adopts Clarke's doctrine that the true rules of morality are **essentially reasonable**. But if conscience is *reason applied to conduct* then Butler's argument seems to bend itself into the old circle—'It is reasonable to live according to nature, and natural to live according to reason'. . . .

3. Can we steer between •doctrines that bring us back to common opinion with all its imperfections and •doctrines that lead us round in a circle, getting through to self-evident moral principles of real significance? Common sense instinctively assumes that there are such principles, and many moralists have consciously believed in them and tried to state them; it would be disheartening to have to conclude that they are altogether illusory! And yet, the more we learn about man and his environment and take in the vast variety of human natures and circumstances at different times and places, the less inclined we are to believe that any definite code of absolute rules is applicable to all human beings without exception. We'll find that the truth lies between these two conclusions. There are some absolute principles of conduct whose truth is obvious once they are explicitly stated, but they're too abstract and too sweeping to show us immediately how we ought to act in any particular case. Particular duties have to be determined by some other method.

One such principle was given on page 98, where I pointed out (though not in these words) that

> If a kind of conduct that is right (or wrong) for me is not right (or wrong) for someone else, there must be some difference between us beyond the fact that we are different persons.

A corresponding and equally true proposition can be stated concerning what ought to be done *to*—not *by*—different individuals. [Sidgwick now runs the discussion through an unhelpful detour via the so-called Golden Rule: 'Do to others as you would have them do to you'. He gets back on course with this:] The self-evident principle must take some such negative form as this:

> 'It can't be right for A to treat B in a manner in which it would be wrong for B to treat A, unless some difference between the natures or circumstances of the two provides a reasonable ground for difference of treatment.'

This clearly doesn't give complete guidance. Its effect is merely to put the onus of proof on the man who treats someone else in a way that he would complain of if it were applied to himself. But common sense has amply recognised the practical importance of the maxim; and it appears to me to be self-evidently true.

The same basic principle is at work in the ordinary administration of law, or (as we say) of 'justice'. On page 140 I drew attention to 'impartiality in the application of general rules', as an important element in the common notion of justice—and the only one that can be intuitively known with perfect clearness and certainty. Again, this isn't sufficient for the complete determination of just conduct, because it doesn't help us to decide what kinds of rules should be thus impartially applied. But everyone agrees that conscious

partiality should be excluded from government and from human conduct generally—this being implied in the common notion of 'fairness' or 'equity'.

·I now come to a different principle, a key element in which is the notion of a *whole*·. The proposition that

> One ought to aim at one's own good

is sometimes offered as the maxim of rational self-love, i.e. prudence; but put like that it doesn't clearly avoid tautology, because we can define 'good' as 'what one ought to aim at'. But now try

> One ought to aim at one's own good *on the whole*.

This points to a principle which, when explicitly stated, is not tautological. In the footnote on page 57 I referred to it as the principle 'of impartial concern for all parts of our conscious life'; and we could express it concisely thus:

> Hereafter *as such* is not to count for less or more than Now.

This doesn't of course mean that the good of the present may not reasonably be preferred to that of the future because of its greater certainty; or that a week ten years hence can't be more important to us than a week now because our means or capacities of happiness will increase. The principle says only that the mere difference of priority and posteriority in time is not a reasonable ground for having more regard to the consciousness of one moment that to that of another. This usually comes up in practice in the thesis that a smaller present good is not to be preferred to a greater future good (allowing for difference of certainty), most commonly in the more specific thesis that it is reasonable to forgo present pleasure in order to get greater pleasure later on; but the principle needn't be restricted to a hedonistic application. It concerns *future versus present*, not *future pleasure versus present pleasure*.

The notion of the 'good on the whole' of a person is constructed by comparing and integrating the different 'goods' that succeed one another in the series of the person's conscious states; and we also have the notion of universal good—·the 'good on the whole' of humanity·—by comparing and integrating the goods of all individual humans. And here again, by considering how the parts relate to the whole and to each other, I obtain the self-evident principle that

> the good of any one person is no more important from the point of view (if I may put it like this) of the universe than the good of any other;

unless there are special grounds for believing that more good is likely to occur in the one case than in the other. And it is evident to me that

> as a rational being I am obliged to aim at good generally. . . .rather than at any particular part of it.

These two rational intuitions rigorously imply the maxim of benevolence in an abstract form, namely that

> each person is morally obliged to regard •the good of anyone else as much as •his own good, except when he judges it to be less, when impartially viewed, or less certainly knowable or attainable by him.

I have remarked that the duty of benevolence recognised by common sense falls short of this. But this could be fairly explained on the basis that each man, even if his eye is on universal good, ought in practice to aim at promoting the good of a limited number of people, generally in proportion to how closely he is related to them. I think that a 'plain man' in a modern civilised society, if he conscientiously faced the question:

> Would it ever be morally right for you to seek your own happiness if that would certainly involve a sacrifice of the greater happiness of some other person, without any counterbalancing gain to anyone else?'

would unhesitatingly answer 'No'.

I have tried to show how the principles of justice, prudence, and rational benevolence as commonly recognised contain self-evident elements that can be immediately known by abstract intuition, each involving how individuals and their particular ends relate to the wholes of which they are parts and to other parts of these wholes. The more or less clear grasp of these abstract truths is, I think, the permanent basis for the common belief that morality's fundamental precepts are essentially reasonable. These principles are often aligned with other precepts to which custom and general consent have given an *illusory* air of self-evidentness, but you have only to *think about* the two kinds of maxims to become sharply aware of the difference between them. 'I ought to speak the truth', 'I ought to keep my promises'—however true these are I know by reflection that they •aren't self-evident to me and •require rational justification. 'I ought not to prefer a present lesser good to a future greater good', 'I ought not to prefer my own lesser good to the greater good of someone else'[1]—these do present themselves as self-evident; as much as the mathematical axiom that if equals are added to equals the wholes are equal.

I can now explain why I have refrained from entering at length into the psychogenetic [see Glossary] question as to the origin of apparent moral intuitions. My reason involved the fundamental and clearly important distinction between

(1) the moral maxims that reflection shows not to possess ultimate validity, and

(2) the moral maxims that are or involve genuine ethical axioms.

No psychogenetic theory has ever been offered claiming to discredit **(2)** the propositions that I regard as really axiomatic on the grounds that their causes were such as had a tendency to make them false. And as regards **(1)** the other class of maxims, there's no *need* for a psychogenetic proof that they are untrustworthy when taken as absolutely and unqualifiedly true, because direct reflection shows me that they have no claim to be taken in that way. On the other hand, psychogenetic theory ·has a certain positive role here·: when it represents moral rules as being—roughly speaking—means to the ends of individual and social well-being, it tends to support the conclusions I have reached by a different method; because it leads us to regard other moral rules as subordinate to the principles of prudence and benevolence.

4. I wouldn't be as confident as I am of the conclusions I have reached in the preceding section if they didn't seem to be in substantial agreement—despite superficial differences—with the doctrines of the moralists who have worked hardest to find genuine intuitions of the practical reason among commonly received moral rules. I pointed out in the Note at the end of I/8 [page 47] that the earlier intuitional moralists show a more philosophical turn of thought on the whole than the later ones reacting against Hume. Among the earlier writers no-one shows more earnestness in trying to penetrate to really self-evident principles than Clarke.[2] For our behaviour towards our fellow-men Clarke lays down two fundamental 'rules of righteousness'. He states his Rule

[1] I'm supposing these propositions to be asserted after fully taking into account the difference of *certainty* between present and future good, and between one's own good and someone else's.

[2] Clarke's anxiety to exhibit the parallelism between ethical and mathematical truth. . . .•renders his general terminology inappropriate (saying that moral intuition is directed at 'relations and proportions' or 'fitnesses and unfitnesses of things'), and sometimes •leads him into absurdities (saying that 'a man who willfully acts contrary to justice wills things to be what they are not and can't be'). But these defects aren't relevant to my present purpose.

of Equity thus:

> 'My judgment that it is reasonable (unreasonable) that someone else should do x for me is also a judgment that it would be reasonable (unreasonable) for me in a similar situation to do x for him.'

This is of course the 'Golden Rule', precisely stated. [Sidgwick now gives a long quotation from Clarke about the obligation to 'universal love or benevolence', and follows that with a charge of 'tautology' from which he rescues it by giving it the non-tautological content:]

> 'The good or welfare of any one individual must as such be an object of rational aim to any other reasonable individual no less than his own similar good or welfare.'

(Note that the proposition that *universal benevolence is the right means to the achieving universal good* is not quite self-evident, because the end may not always be best attained by directly aiming at it. Thus rational benevolence, like rational self-love, may be self-limiting; may direct its own partial suppression in favour of other impulses.)

Among later moralists, Kant is especially noted for his rigour in extracting the purely rational element of the moral code; and his ethical view seems to me to coincide—at least to a considerable extent—with the view presented in the preceding section. I have already noted that his basic principle of duty is the 'formal' rule of 'acting on a maxim that one can will to be universal law'; and this, when appropriately restricted,[1] is an immediate practical corollary of the principle that I first noted in the section 3. And the only really ultimate end that he offer for virtue to aim at is the object of rational benevolence as commonly conceived—the happiness of other men.[2] He regards it as evident *a priori* that each man as a rational agent is obliged to aim at other men's happiness; in his view, indeed, I have a duty to seek my own happiness only because I consider it as a part of the happiness of mankind in general. I disagree with this last claim, because I agree with Butler that 'one's own happiness is a manifest obligation' independently of one's relation to other men; but the positive part of Kant's conclusion seems mainly to agree with the view of the duty of rational benevolence that I have given. (I don't accept his arguments for his conclusion. See the note at the end of this chapter.)

5. You'll have realised by now that the self-evident principles laid down in section 3 don't specially belong to 'intuitionism' in the restricted sense that I gave to this term at the start of this Book. The axiom of prudence, as I have given it, is a self-evident principle, implied in rational egoism as commonly accepted'.[3] And the axiom of justice or equity as above stated—that similar cases ought to be treated similarly—belongs as much to utilitarianism as to any system commonly called intuitional; while the axiom of rational benevolence is required as a rational basis for the utilitarian system.

So my search for really clear and certain ethical intuitions brings me at last to the basic principle of utilitarianism.

[1] I don't think that Kant *does* appropriately restrict it. See chapter 7.3 and IV/5.1.

[2] Kant gives the agent's own perfection as another absolute end; but his account of perfection implies that this isn't ultimately basic because it presupposes that there are other ends of reason. . . . He writes: 'The perfection that belongs to men generally. . . can only be the cultivation of one's power and one's will to satisfy the requirements of duty in general.'

[3] How does rational egoism relate to rational benevolence?' I regard that as the profoundest problem of ethics. My final answer to it will be given in the concluding chapter of this work.

I admit that recent utilitarians haven't tried to show the truth of their first principle through any procedure like the one I have given. Still, the 'proof' of the 'principle of utility' given by Mill, the most persuasive and probably most influential English utilitarian, obviously needs to be completed by a procedure like mine. (My quotations will be from chapters 1 and 2 of *Utilitarianism*.)

Mill begins by explaining that though 'questions of ultimate ends are not amenable to *proof* in the ordinary and popular meaning of the term', there's a broader meaning of 'proof' in which they are capable of proof: the subject is 'within the reach of the rational faculty. . . Considerations can be presented capable of determining the intellect to accept the utilitarian formula'. He makes clear that by 'acceptance of the utilitarian formula' he means the acceptance. . . .of 'the greatest amount of happiness altogether' as the ultimate 'end of human action' and 'standard of morality'; and he holds that the supreme 'directive rule of human conduct' is to promote that end. When he comes to give the 'proof'—in the broader sense before explained—of this rule or formula, he offers the following argument:

> 'The sole evidence it is possible to produce that anything is desirable is that people do actually desire it. . . No reason can be given why the general happiness is desirable except the fact that each person desires his own happiness. But this *is* a fact; so we have not only all the proof there could be for such a proposition, but all the proof that could possibly be demanded, that happiness is a good; that each person's happiness is a good to that person, and therefore that general happiness is a good to the aggregate of all persons.[1]

He then goes on to argue that pleasure, and pleasure alone, is what all men actually do desire.

. . . .When Mill says that 'the general happiness is desirable', he must be understood to mean—and his whole treatise shows that he *does* mean—that it is what each individual ought to desire, or at least. . . .to aim at achieving through his actions.[2] But this proposition is not established by Mill's reasoning, even if we grant that something's being actually •desired implies that it is in this sense •desirable. For an aggregate of actual desires, each directed towards a different part of the general happiness, does not constitute an actual desire for the general happiness existing in any individual; and Mill certainly wouldn't contend that a desire that doesn't exist in any individual can exist in an aggregate of individuals! So the proposition that the general happiness is desirable can't be established by this argument, which requires there to be an actual desire for the general happiness; there's a gap in Mill's argument, and it (I think) be filled only by a proposition like the one I have tried to exhibit as the intuition of rational benevolence.

Utilitarianism thus appears as the final form into which intuitionism tends to pass, when the demand for really self-evident first principles is pressed hard. But its passage isn't logically complete unless we interpret 'universal good' as 'universal happiness'. And this interpretation can't be justified by arguing, as Mill does, from •the psychological premise that happiness is the sole object of men's actual desires to •the ethical conclusion that it alone is desirable or good; because in I tried in I/4 to show that happiness or pleasure is not the only thing that each man actually desires for himself. The identification of •ultimate good with

[1] This is from *Utilitarianism* chapter 4; page 24 in the version on the website at www.earlymoderntexts.com.

[2] I have been accused of overlooking a confusion in Mill's mind between 'desirable' = 'what can be desired' and 'desirable' = 'ought to be desired'. The present paragraph shows that I am aware of this confusion, but I don't think that for my purposes I need to discuss it.

•happiness is properly to be reached, I think, by a more indirect line of reasoning that I'll try to explain in the next chapter.

Note: The present great influence of Kant's teaching makes it worthwhile to state briefly the arguments by which he tries to establish the duty of promoting the happiness of others, and my objections to them. In some places he tries to exhibit this duty as following from •his fundamental formula 'Act from a maxim that you can will to be universal law' in conjunction with •the fact (as he assumes it to be) that every man in severe difficulties must desire to get help from others. The maxim 'Everyone should be left to take care of himself without help or interference' is one that we could *conceive* existing as a universal law, Kant says, but we couldn't *will* to be such. 'A will that went that way this would be inconsistent with itself, for many cases may arise where the individual thus willing needs the benevolence and sympathy of others.' And in another work he explains at more length that the self-love that necessarily exists in everyone involves the desire to be loved by others and to be helped by them in case of need. So we necessarily constitute ourselves as an end for others, and claim that they shall contribute to our happiness; and so, according to Kant's fundamental principle, we must recognise our duty to make *their* happiness *our* end.

This is not a valid argument. *Every man in need wishes for the aid of others*—that's an empirical proposition that Kant cannot know *a priori*. We can certainly conceive a man in whom the spirit of independence and the distaste for incurring obligations is so strong that he would prefer •enduring any privations to •receiving aid from others. Furthermore, even if it's true that everyone in distress must at that moment necessarily wish for help from others, it's still possible for a strong man, after balancing the chances of life, to think that he and his like have more to gain, over-all, by the general adoption of the egoistic maxim, benevolence being likely to bring them more trouble than profit.

Elsewhere Kant reaches the same conclusion by an apparently different line of argument. . . . All particular ends at which men aim get their status as ends by the existence of impulses directed towards particular objects. Now we can't tell *a priori* that any of these special impulses forms part of the constitution of all men; so we can't state it as an absolute dictate of reason that we should aim at any such special object. Set all those particular empirical ends aside, then, and all that remains is the principle that 'all rational beings as such are ends to each' or, as Kant sometimes puts it, that 'humanity exists as an end in itself'.

Now, says Kant, if I confine myself to mere non-interference with others, I don't positively make humanity my end; my aims remain selfish, though restricted by this condition of non-interference with others. So my action is not truly virtuous, because virtue shows in—and *consists* in—the effort to bring about the end of reason in opposition to mere selfish impulses. Therefore 'the ends of the other person, who is himself an end, must of necessity be *my* ends if the representation of humanity as an end in itself is to have its full weight with me' and my action is to be truly rational and virtuous.

I can't accept this argument either. The conception of 'humanity as an end in itself' is perplexing; because by an *end* we commonly mean something to be brought about—·a state of affairs·—whereas 'humanity' is, as Kant says, 'a self-subsistent end' [i.e. it is a thing, not a state of affairs]. Also, there seems to be a logical flaw in the deduction of the principle of benevolence by means of this conception. The humanity that Kant maintains to be an end in itself is Man (or the aggregate of men) *in so far as rational* [Sidgwick's phrase].

But other men's subjective ends, which benevolence directs us to take as our own ends, would seem according to Kant's own view, to depend on and correspond to their non-rational impulses—their empirical desires and aversions. It's hard to see why, if man as a rational being is an absolute end to other rational beings, they must therefore adopt his subjective aims as determined by his non-rational impulses.

Chapter 14: Ultimate good

1. At the outset of this treatise I noted that there are two accounts of what ethical inquiry is *about*:

- rules of conduct—'the right';
- ends—'the good'.

I remarked that in the moral consciousness of modern Europe the two notions are *prima facie* distinct: it is commonly thought that the obligation to obey moral rules is absolute, but it isn't commonly held that the whole good of man lies in such obedience. The view that it *does* is vaguely and respectfully—but unmistakably!—repudiated as a Stoic paradox. The ultimate good or well-being of man is regarded rather as an upshot, the connection of which with his right conduct is indeed commonly held to be certain, but is often conceived as supernatural and therefore beyond the range of independent ethical speculation. [To make sure that is understood: the crucial distinction is between 'Man's good *lies in* moral obedience' and 'Man's good *results from* moral obedience'.] But if my conclusions in the preceding chapters are right, it would seem that ·this has things back-to-front, and that· the practical determination of right conduct *depends on* the determination of ultimate good. I'm talking about conclusions regarding the principles of **prudence** and **benevolence**, namely that

- most of the commonly received maxims of duty—even ones that at first sight look absolute and independent—turn out to be subordinate to them;
- they, along with the formal principle of justice or equity, are the only principles that are intuitively both clear and certain; and
- in their self-evident forms ·they are about *ends*, because· one is a precept to seek one's own good on the whole, and the other is a precept to seek others' good as well as one's own.

Thus we are brought back to the old question with which ethical theorising in Europe began—'What is the ultimate good for man?'—though not in the egoistic form in which the old question was raised. But when we examine the controversies that this question originally led to, we see that the investigation that has brought us back to it has excluded the once-popular answer that *general good consists solely in general virtue*. 'Virtue' means 'conformity to prescriptions that are the main part of common-sense morality', and we have seen that the exact determination of these prescriptions depends on what the general good is; so the once-popular answer leads us around a logical circle.

We can't evade this argument by adopting what on page 108 I called 'aesthetic intuitionism', regarding virtues as excellences of conduct clearly discernible by trained insight although their nature doesn't allow them to be stated in definite formulae. Our notions of special virtues don't become more independent by becoming more indefinite; they still contain, though perhaps less explicitly, the same reference to 'good' or 'well-being' as an ultimate standard. You can see this when you consider any virtue in relation to the vice or non-virtue that it tends to become when •pushed to an extreme or •exhibited under inappropriate conditions. Common sense may seem to regard the qualities on the upper line as intrinsically desirable; but when we relate them respectively to the items on the lower line

liberality frugality courage placability

profusion meanness foolhardiness weakness

we find that in each case common sense draws the line not by immediate intuition but by reference either to some definite maxim of duty, or to the general notion of 'good' or well-being. [Sidgwick has argued that *definite* duties involve references to ends; aesthetic intuitionism was trying to avoid that by making all duties indefinite.] And the same thing happens when we ask at what point candour, generosity and humility cease to be virtues by becoming excessive. Other commonly admired qualities—energy, zeal, self-control, thoughtfulness—are obviously regarded as virtues only when they are directed to good ends. The only so-called virtues that can be thought to be essentially and always virtues, and incapable of excess, are qualities such as wisdom, universal benevolence, and (in a sense) justice; and the notions of these obviously presuppose a determinate notion of *good*.

•Wisdom is insight into good and the means to good;

•benevolence is exhibited in actions called 'doing good';

•justice (when regarded as always a virtue) lies in distributing good (or evil) impartially according to proper rules.

If then we are asked 'What *is* this good that it is excellent to •know, to •bestow on others, to •distribute impartially, it would be obviously absurd to reply that it is just •this knowledge, •these beneficent purposes, •this impartial distribution!

And I can't see that this difficulty is coped with by regarding virtue as a quality of 'character' rather than of 'conduct', and expressing the moral law in the form 'Be this' rather than 'Do this'.[1] Of course from a practical point of view it is important to urge men to •aim at an ideal of character and •consider the effects of actions on character. But the thesis that character and its elements—faculties, habits, or dispositions—are the constituents of ultimate good it doesn't follow from that, and is indeed ruled out by the very concept of a faculty or disposition. A faculty or disposition is only a *tendency* to act or feel thus and so; and such a tendency isn't valuable in itself but for the acts and feelings in which it takes effect, or for the further consequences of these; and these consequences can't be *ultimately* good as long as they are conceived as merely states of faculties, dispositions, etc. . . .

2. I have been speaking only of particular virtues as exhibited in conduct judged to be objectively right; and you may think that this is too *external* a view of the virtue that claims to constitute ultimate good. The difficulty I have been presenting vanishes, you may say, if we penetrate beyond •the particular virtues to the root and essence of virtue in general, namely •the will's being set to do whatever is judged

[1] See Leslie Stephen, *The Science of Ethics*, chapter 4.16 .

to be right and to aim at bringing about whatever is judged to be best. This subjective rightness or goodness of the will doesn't depend on knowledge of what is objectively right or good, so it also doesn't depend on the presupposition of good as already known and determined—the presupposition that we have seen to be implied in the common conceptions of virtue as exhibited in outward acts. Well, I admit that the thesis that *subjective* rightness or goodness of will is the ultimate good doesn't involve the logical difficulty that I have been urging. But it is radically contrary to common sense, because the very notion of *subjective* rightness or goodness of will implies an *objective* standard which it directs us to seek but doesn't claim to supply. Consider:

> The right-seeking mind asks for directions, and we tell it •that all it has to aim at is this right-seeking itself, •that this is the sole ultimate good, and •that the only *effect* of right volition that can be good in itself is the subjective rightness of one's own or others' future volitions.

This is a palpable and violent paradox. It's true that a reasonable person can't recognise any more authoritative rule than the one telling him to do what *he judges* to be right; because when he is wondering what to do next he can't distinguish •doing what is objectively right from •conforming to his own subjective conception of rightness. But we're continually forced to make that distinction as regards the actions of others. . . ., and we continually judge conduct to be objectively wrong because it tends to cause pain and unhappiness to others, apart from any effect on the subjective rightness of their volitions. That is what we are doing when we recognise the mischief and danger of fanaticism—meaning by 'a fanatic' a man who insists on his own conception of rightness when it is plainly mistaken.

The same result can be reached without going out to anything as extreme as fanaticism. As I pointed out in chapter 11.3 (see also chapter 12.3), although the 'dictates of reason' are always to be obeyed, it doesn't follow that 'the dictation of reason'—the predominance of consciously moral motives over non-moral ones—is to be promoted without limits. Common sense holds that some things are likely to be better done if they are done from other motives than conscious obedience to practical reason, i.e. to conscience. So there's a real *question* here:

> How far should we go in aiming to make reason's dictate—the predominance of moral choice and moral effort—predominant in our lives?

By allowing this question to be asked, we concede that conscious rightness of volition is not the sole ultimate good. Summing up, then: neither •subjective rightness or goodness of volition nor •virtuous character (except as expressed in virtuous conduct) can constitute ultimate good; and we can't identify ultimate good with •virtuous conduct either, because our conceptions of various kinds of virtuous conduct presuppose the prior determination of the notion of good—the good that virtuous conduct is suppose to produce or promote or rightly distribute.

And this result regarding virtue applies even more obviously to the other talents, gifts, and graces that make up the common notion of human perfection. However immediately the excellent quality of such gifts and skills is recognised and admired, if you think a little you'll realise that they are valuable only because of the good life in which they are actualised or which will be promoted by their exercise.

3. Shall we then say this?–

> Ultimate good is desirable conscious life of which virtuous action is one element but not the whole thing?

This seems to be in harmony with common sense; and the

fact that particular virtues and talents and gifts are valued largely as means to •resultant goods doesn't rule out their exercise as an element of •ultimate good; just as the fact that a proper combination of physical action, nutrition, and rest is a •means to the maintenance of our animal life doesn't rule out their being indispensable •elements in such a life. But it's hard to conceive any kind of activity as both means and end, from the same point of view and in respect of the same quality; and in both those cases it's easy to distinguish the aspect in which the activities are to be regarded as means from that in which they are to be regarded as desirable in themselves. Start with the physical processes. It's in their purely physical aspect, as complex processes of corporeal change, that they are means to the maintenance of life; but considering them purely in this way, as movements of particles of matter, it seems impossible to regard them as in themselves either good or bad. . . . If any quality of human life is what is ultimately desirable, it must concern human life on its mental side—*consciousness*, for short.

But mental life as known to us includes pain as well as pleasure, and so far as it is painful it isn't desirable. So I can't accept a view of the well-being of human beings and other living things that is •suggested by current zoological conceptions and •maintained by influential writers. It's the view that when we describe as 'good' the manner of existence of any living organism, we are attributing to it a tendency either •to self-preservation or •to the preservation of its community or its race; so that what 'well-being' adds to mere 'being' is just promise of future being. If you think clearly about this you'll see that it is wrong. If all life were as undesirable as some portions of it have been in my experience and (I believe) in that of most men, I would judge any tendency to preserve it to be totally bad. Actually, we generally hold that human life, even as now lived, has on average a balance of happiness; so we regard what preserves life as generally good, and what destroys life as bad; and an important part of morality's function is indeed to maintain habits and sentiments that are needed for the continued existence, in full numbers, of a society of human beings under their actual conditions of life. But that's not because the mere existence of human organisms, even if prolonged to eternity, appears to me in any way desirable; it is assumed to be so only because it is supposed to be accompanied by consciousness that is on the whole desirable. It is therefore this *desirable consciousness* that we must regard as the ultimate good.

Similarly, when we judge virtuous activity to be a part of ultimate good, I think it's because we judge the accompanying consciousness to be desirable for the virtuous person; though this consideration doesn't adequately represent the importance of virtue to human well-being, because we have to consider its value as a means as well as its value as an end. Consider this: If virtuous life were combined with extreme pain, would it still be on the whole good for the virtuous agent? The answer 'Yes' was strongly supported in Greek philosophical discussion; but it's a paradox from which a modern thinker would recoil: he wouldn't say that the portion of life spent by a martyr in tortures was in itself desirable, though it might be his duty to suffer the pain for the good of others or even for his own eventual happiness.

4. If then ultimate good has to be conceived as *desirable consciousness* (including the consciousness of virtue as a *part*), should we identify this notion with *happiness* or *pleasure*, and say as the utilitarians do that general good is general happiness? You might well think that this conclusion as inevitable—especially if it seems to you that this:

> Other things that are called 'good' are only means to the end of making conscious life better or more desirable

is just one way of saying this:

> Other things that are called 'good' are only means to the end of happiness.

But ·let's not rush into that, because· some important distinctions remain to be considered. According to the view taken in II/2, when we affirm that ultimate good is happiness or pleasure, we imply •that nothing is desirable except desirable feelings; •that the desirability of each feeling can be known directly only by the person who has the feeling, and only at the time of having it; and •that therefore this judgment of the person must be taken as final on the question how far each element of feeling has the quality of ultimate good.[1] I don't think anyone would estimate in any other way the desirability of feeling considered merely as feeling; but you might say this:

> Our conscious experience includes not only feelings but also cognitions and volitions; the desirability of these must be taken into account, and it can't be estimated in the way you have described.

[In this context, a 'cognition' is a short-term belief or thinking-that-P.] But when we reflect on a cognition as a transient fact of an individual's mental life—distinguishing from •the feeling that normally accompanies it and from •any facts about whether it is true or valid—we can see that it's an element of consciousness which is neutral in respect of desirability; and the same is true of volitions when we set aside •the feelings associated with them, •their relation to any objective norm or ideal, and •all their consequences. I don't deny that in ordinary thought certain states of consciousness—such as cognition of truth, contemplation of beauty, volition to be free or virtuous—are sometimes judged to be preferable on grounds other than their pleasantness; but I stand by the explanation of this that I suggested in II/2.2 , namely that what we really prefer in such cases is not the present consciousness but either •its effects on future consciousness or else •something in the objective relations of the conscious being that isn't included in his present consciousness.

An example of that second alternative: a man who prefers the mental state of •apprehending truth to the state of •reliance on generally accepted fictions, even if the former state is more painful than the latter, and independently of any effect that he expects either state to have on his subsequent consciousness. In this case, as I see it, what he really prefers is not •the consciousness of knowing truth, considered merely as consciousness, because the pleasure in this is more than outweighed by the concomitant pain, but •the relation between his mind and something else that I call 'objective' because it exists independently of his cognition of it. Suppose you learn that something you have taken for truth is not really such; you'll certainly feel that your preference had been mistaken; but if your choice had really been between two elements of transient consciousness, its reasonableness couldn't be affected by any later discovery.

Another example: a man prefers freedom and poverty to a life of luxurious servitude, not because the pleasant consciousness of being free outweighs the expected comforts and securities of the other life, but because he has a predominant aversion to slavery. . . . If he comes later to believe that his conception of freedom was illusory—that we are all

[1] Final, that is, so far as the quality of the present feeling is concerned. When an estimate of the desirability of a feeling involves comparison with feelings that are only represented in idea [e.g. in memory], it can be erroneous through imperfections in the representation.

slaves of circumstances, destiny, etc.—that will lead him to regard his preference as mistaken.

When someone prefers conformity to virtue (or contemplation of beauty) to a state of consciousness that he recognises as pleasanter, this preference seems to come from his belief that his conception of virtue (or of beauty) corresponds to an ideal that is to some extent objective and valid for all minds. Apart from any consideration of future consequences, we would all agree that a man who sacrificed happiness to an erroneous conception of virtue or beauty had made a mistaken choice.

Still, it may be said that this is merely a question of definition—

> 'We can take "conscious life" in a wide sense, so as to include the objective relations of the conscious being implied in our notions of virtue, truth, beauty, freedom; and then we can regard cognition of truth, contemplation of beauty, free or virtuous action, as preferable to pleasure or happiness, while admitting that happiness must be included as a part of ultimate good.'

If that is right, the principle of rational benevolence—described in chapter 13 as an indubitable intuition of the practical reason—would direct us to the pursuit not merely of universal happiness but also of these 'ideal goods', as ends that are ultimately desirable for mankind generally.

5. But no careful and thoughtful person should find this plausible. To show this, I ask you to use the same twofold procedure that I asked you on page 166 to use in considering the absolute and independent validity of common moral precepts. I appeal **(i)** to your intuitive judgment, after due consideration, of the question when fairly placed before you; and **(ii)** to a comprehensive comparison of the ordinary judgments of mankind. **(i)** It seems clear to me after reflection that these objective relations of the conscious subject, when distinguished from the consciousness accompanying and resulting from them, are not ultimately and intrinsically desirable; any more than material objects are when considered apart from any relation to conscious existence. Granted, we have actual experience of preferences like the ones I have described, of which the ultimate object is something that isn't merely consciousness; but it seems to me that when we 'sit down in a cool hour' (Butler's phrase), the only way we can justify to ourselves the importance that we attach to any of these objects is by considering its conduciveness, in one way or another, to the happiness of sentient beings.

(ii) The argument involving the common sense of mankind can't be made completely forceful because some cultivated persons *do* habitually judge that knowledge, art, etc. are *ends* independently of the pleasure derived from them. But all these elements of 'ideal good' *do* •produce pleasure in various ways; and the amount of commendation they get from common sense is roughly proportional to *how* •productive they are. This seems obviously true of beauty; and it will hardly be denied in respect of any kind of social ideal; it is paradoxical to maintain that any degree of freedom, or any form of social order, would still be commonly regarded as desirable even if we were certain that it had no tendency to promote the general happiness. The case of knowledge is rather more complex; but certainly common sense is most impressed with the value of knowledge when its 'fruitfulness' has been demonstrated. [Sidgwick goes on to say that even when a scientific discovery doesn't look useful, •common sense allows that it may become useful in due course. But there are limits to •its patience.] Common sense is somewhat disposed to complain of the misdirection of valuable effort; so that the amount of honour commonly paid to science seems to be graduated, perhaps unconsciously, on a pretty

exact utilitarian scale. When there's a serious dispute about the legitimacy of any branch of scientific inquiry—as in the recent debate about vivisection—the controversy on both sides is generally conducted on an openly utilitarian basis.

The case of virtue needs to be discussed separately. A main aim of men's ordinary moral discourse is to encourage virtuous impulses and dispositions in each other; so that the very *question* whether this encouragement can go too far has a paradoxical air. But we have, though rarely, experienced cases in which the concentration of effort on the cultivation of virtue has been harmful to general happiness through •being intensified to the point of moral fanaticism and •neglecting other conditions of happiness. And I think we'll generally admit that in such cases the criterion for deciding how far the cultivation of virtue should be carried should be conduciveness to general happiness.

Still, common sense *does* resist accepting that happiness (when explained to mean a sum of pleasures) is the sole ultimate end and standard of right conduct. But this can be accounted for by four considerations.

A. The term 'pleasure' is not commonly used so as to include clearly *all* kinds of consciousness that we desire to retain or reproduce; in ordinary usage it suggests too prominently the coarser and commoner kinds of such feelings; and even those who are trying to use the word scientifically find it hard to •free their minds from the associations of ordinary usage and to mean by 'pleasure' nothing but 'desirable consciousness or feeling of whatever kind'. Also, we are constantly reminded of what Bentham called 'impure' pleasures—ones that will inevitably involve greater pain or the loss of more important pleasures—and we naturally shrink from including these, even hypothetically, in our conception of ultimate good; especially since we often have moral or aesthetic instincts warning us against them.

B. I showed in I/4 (see also II/3) that many important pleasures can be felt only if we have desires for things other than pleasure. For example, we miss the valuable pleasures that accompany the exercise of the benevolent affections if we don't have genuinely disinterested impulses to make others happy. So the very acceptance of pleasure as the *ultimate* end of conduct involves the practical rule that it shouldn't always be the *conscious* end; and that may help to explain the reluctance of common sense to regard pleasure as the sole thing ultimately desirable.

C. . . . .When happiness is spoken of as man's sole ultimate good, the idea most commonly suggested is that each individual is to seek his own happiness and neglect of that of others; and this offends both our sympathetic and our rational regard for others' happiness. (Rational regard? Yes; as I argued in chapter 13, reason shows me that if my happiness is desirable and a good then the equal happiness of anyone else must be equally desirable.) What common sense is averse to is the goal of •egoistic rather than of •universalistic hedonism. And certainly one's individual happiness is in many ways a poor target for one's supreme aim, apart from any clash it may have with rational or sympathetic benevolence. That's because it doesn't have the features that Aristotle says we 'divine' to belong to ultimate good: so far as it can be empirically foreseen, individual happiness is narrow and limited, short-lived, and shifting and insecure while it lasts. But universal happiness—desirable consciousness for the countless sentient beings present and to come—satisfies our imagination by its vastness, and sustains our resolution by its comparative security.

Here's an objection that someone might offer:

> If we require the individual to sacrifice his own happiness to the greater happiness of others on the ground that it's reasonable to do so, we're assigning to him a

> different ultimate end from the one we present as the ultimate good of the universe of sentient beings. We tell him that the universe's ultimate good is happiness, while his is conformity to reason.

I admit the substantial truth of this statement. . . ., but I don't see it as counting against the position I am maintaining. Why not? Because the individual is essentially and fundamentally different from the larger whole—the universe of sentient beings of which he is a part. The difference is that *he* has a known relation to similar parts of the same whole, while the whole itself has no such relation. So I don't see any inconsistency in holding that while

- if the aggregate of sentient beings could act collectively, it would be reasonable for it to have its own happiness as its only ultimate end, and
- if an individual were the only sentient being in the universe, it would be reasonable for him to have his own happiness as his only ultimate end,

it may be actually reasonable for an individual to sacrifice his own good or happiness for the greater happiness of others. [The crucial contrast there is between what *would* be the case if. . . etc. and what is *actually* the case in the world as it is.]

I admit that in ancient Greek philosophy men sometimes judged an act to be 'good' *for the agent* while recognising that its consequences would be painful to him—e.g. a heroic exchange of a life full of happiness for a painful death at the call of duty. I attribute this partly •to a confusion of thought between what it is reasonable for an individual to desire when he considers his own existence alone, and what he must recognise as reasonably to be desired when he takes the point of view of a larger whole; and partly to a faith deeply rooted in the moral consciousness of mankind, that there can't be really and ultimately any conflict between the two kinds of reasonableness.[1] [Sidgwick continues with some difficult remarks about how that Greek attitude figures in modern moral philosophy.]

D. From the universal point of view as much as from that of the individual, it seems true that happiness is likely to be better attained if the extent to which we set ourselves consciously to aim at it is carefully restricted. One important reason for this is that action is likely to be more effective if our effort is temporarily concentrated on more limited ends; but there is also another. The fullest development of happy life for each individual seems to require him to have external objects of interest other than the happiness of other conscious beings. So we can conclude that the pursuit of the ideal objects I have mentioned—virtue, truth, freedom, beauty, etc.—*for their own sakes* is indirectly and secondarily rational, though not primarily and absolutely so—not only because of the happiness that will result from their attainment but also because of the happiness that comes from the disinterested pursuit of them. And yet if we look for a final criterion of the comparative value of the different objects of men's enthusiastic pursuit,. . . . we'll conceive it to depend on how far each of them conduces to happiness.

If this view is rejected, can we construct any other coherent account of ultimate good? If we're not to systematise human activities by taking universal happiness as their common end, on what other principles are we to systematise them? They need to be principles that enable us not only •to make value-comparisons *among* the non-hedonistic ends

[1] We can illustrate this double explanation by a reference to Plato's *Gorgias*, where the ethical argument has a singularly mixed effect on the mind. It strikes us as partly a fairly skillful set of tricks, playing on a confusion of thought latent in the common notion of good; and partly a noble and stirring expression of a profound moral faith.

we have been considering but also •to provide a common standard for comparing these values with that of happiness—unless we're prepared to reject happiness as absolutely valueless! We have a practical need to answer not only questions of the type

> Should we should pursue truth rather than beauty? or freedom or some ideal constitution of society rather than either of those? or desert all of these in favour of a life of worship and religious contemplation?

but also questions of the type:

> How far we should follow any of these lines of endeavour when we foresee it resulting in pain for human or other sentient beings, or even the loss of pleasures that they might otherwise have enjoyed?[1]

I haven't found and can't construct any systematic answer to this question that appears to deserve serious consideration; so I am finally led to the conclusion (which at the close of the last chapter seemed to be premature) that when the intuitional method is rigorously applied it leads to the doctrine of *pure universalistic hedonism*. . . .that it is convenient to label as 'utilitarianism'.

[1] The controversy on vivisection illustrates well the need that I am pointing out. I haven't heard anyone in this controversy paradoxically deny that the pain of sentient beings is, just because it is pain, to be avoided.

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-unconscious 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 felicific [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 kindliness, 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, imposing[1] 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 laws 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 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 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 felicific 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 it 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 infelicific even though they may seem to be infelicific 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 infelicific. 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 felicific; 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 them[1]—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 unfelicific, 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 sometimes *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 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 felicific 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/assets/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 infelicific. 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 unaspiring; 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 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 who values conduct in proportion to its felicific 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.

[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.]

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 a 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 writers[1] seem to think that we can be led to see that the good of each coincides with the good of all by thinking had 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 allowing[2] 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 felicific (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 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 adequately[1] 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.