9. Medicine and Ethics

Outside of very well integrated groups, it is hard to legislate about word meanings. In everyday discourse, words tend to take on a life of their own independent of any stipulations. This is important to remember when it comes to the words ‘ethics’, ‘morals’, and ‘morality’. Even though some philosophers define these nouns in such a way that they receive distinct meanings (‘ethics’, for instance, is often defined as ‘the philosophy of morals’), in many everyday contexts they are synonymous. Etymologically, ‘ethics’ comes from the ancient Greek words ‘ethikos’ and ‘ethos’. The latter meant something like ‘the place of living’ and the former ‘arising from habit’. That is, both have to do with custom or practice. ‘Morals’ and ‘morality’ are derived from Latin words such as ‘moralis’, ‘mos’, and ‘mor-’ meaning ‘custom’. Today, what is moral or ethical is often contrasted with what comes about only by habit or custom.

The words ‘ethics’ and ‘morality’ are today sometimes used in a purely descriptive sense and sometimes in a normative. To say ‘their ethics (ethical system, morals, or morality) contains the following main rules: …’ is mostly only meant to describe in a neutral way what rules a certain group adheres to, whereas to say ‘this is what the laws dictate, but this is what ethics requires’ is to use ‘ethics’ in a normative sense. Normally, to call someone ‘a morally responsible person’ is to say that he wants to act in accordance with the right norms. But to call someone ‘a moralist’ is to say that he is a bit too keen on judging the moral behavior of others.

Where a moral system is accepted, it regulates the life of human beings. It even regulates how animals are treated. Through history, and still today, we find different cultures and subcultures with different and conflicting rules of how one should to act. Philosophers have tried to settle these conflicts by means of reasoning, but so far they have not succeeded in obtaining complete consensus even among themselves. Therefore, in some of the subchapters below, we will present a common tripartite classification of philosophers’ ethical positions:

- deontology (deontological ethics or duty ethics)
- consequentialism (one main form being utilitarianism or utilitarian ethics)
- virtue ethics.

A moral-philosophical system gives an elaborate presentation and defense of what should be regarded as morally right and wrong actions as well as morally good and bad persons, properties and states of affairs. During this presentation, we will argue in favor of a certain kind of modern virtue ethics, but we hope nonetheless that we make such a fair presentation of duty ethics and utilitarian ethics, that the readers are able to start to think on their own about the merits and demerits of these positions. Our views affect, let it be noted, how we look upon medical ethics. We will not deal with ethical nihilism and complete ethical subjectivism, i.e., the positions that end up saying that there are no super- or inter-personal moral constraints whatsoever on our actions. If some readers find all the ethical systems we present absurd, then there are reasons to think that they are either ethical nihilists or subjectivists.

Medical ethics is often, together with disciplines such as business ethics, environmental ethics, and computer ethics, called ‘applied ethics’. This label gives the impression that medical ethics is only a matter of applying to the medical field a set of pre-existing abstract moral principles. And so it is for deontological and consequentialist ethicists, but is not for virtue ethicists. The reason is that virtue ethics does not basically rely on verbally explicit moral rules. In Chapter 5 we argued that epistemology has to extend beyond ‘knowing that’ and take even the existence of ‘knowing how’ into account. Here, we will argue that, similarly, ethics has to take its kind of know-how into account too.

In *The Reflective Practitioner*, Donald Schön describes the relationship between theory and practice (or ‘knowing that’ and ‘knowing how’) as follows:

> In the varied topography of professional practice, there is a high, hard ground where practitioners can make effective use of research-based-theory and technique, and there is a swampy lowland where situations are confusing 'messes' incapable of technical solutions.

We regard this as being as true of medical ethics as of medical practice. Even medical ethics is influenced by the special topography of swampy lowlands; too seldom are there any straight ethical highways in this area.

Medical ethics may be defined as a multidisciplinary research and education discipline that historically, empirically, and philosophically scrutinizes moral and ethical aspects of health care in general, clinical activities, and medical research. It evaluates merits, risks, and social concerns of activities in the field of medicine. Often, new medical discoveries and clinical developments imply new ethical considerations. Just think of the development of DNA-technology, the cultivation of stem cells and cloning, prenatal diagnosis, fetus reduction, and xenotransplantation. Clinical ethics also bring in the physician-patient relationship, the physician's relations to relatives of the patient, relations between doctors and other health care professionals, and relations between doctors and the society at large. Medical ethics is part of biomedical ethics, which also includes things such as environmental ethics, animal rights, and the ethics of food production.

Moral philosophers have not only been discussing ethical systems and applied ethics. During the twentieth century, especially the latter half, Anglo-American philosophers were preoccupied with what they call 'meta-ethics', i.e., problems that are related to ethics but nonetheless are 'beyond' (= 'meta') ethics. Typical meta-ethical issues are (i) analyses of moral language (e.g., 'what does it *mean* to claim that something is morally good?'), and (ii) the questions whether there can be objectivity and/or rationality in the realm of morals. Some philosophers claim that such meta-ethical investigations are completely neutral with respect to substantive ethical positions, but others contest this and say that meta-ethics has repercussions on substantive ethics; we align with the latter ones. Meta-ethicists who believe either that moral statements are true or false or that (even if not true or false) their validity can be rationally discussed, are

called ‘cognitivists’; non-cognitivism in meta-ethics easily implies ethical nihilism in substantive ethics.

Since, in philosophy, the term ‘meta’ means ‘coming after’ or ‘beyond’, also mere presentations of moral and moral-philosophical systems can be regarded as a kind of meta-ethics. To such a meta-ethics we now turn.

9.1 Deontology

The Greek ‘deon’ means duty. Classical deontological ethics categorically prescribes that certain well defined actions are obligatory or are prohibited. A prototypical example is the Ten Commandments or the Decalogue from the Old Testament, which have played a dominant moral role in Christianity and (to a lesser extent) in Judaism. Here are some of these commandments (we do not give them number, since there is no culture independent way to do this):

- Honor your father and your mother!
- You shall not kill (human beings)!
- You shall not steal!
- You shall not bear false witness against your neighbor!
- You shall not covet your neighbor’s wife!
- You shall not covet your neighbor’s house or anything that belongs to your neighbor!

As they stand, these imperatives tell us what to do independently of both context and consequences. In no way can a deontologist say what Groucho Marx is reported to have said: ‘Those are my principles; if you don’t like them I have got others!’ The duties stated are meant to be *absolute* (i.e., they do not allow any exceptions), *categorical* (i.e., they are not made dependent on any consequences of the action), and *universal* (i.e., they are not, as nowadays presented, norms only in relation to a certain culture). In this vein, the commandment ‘You shall not kill!’ says that I am not under any circumstances allowed to kill anyone, even in a case of euthanasia, not even myself, whatever consequences of suffering I may then have to endure; and this is so irrespective of which culture I happen to belong to.

Two more things should be noted. First, it is harder to know exactly how to act in order to conform to the obligations (e.g., 'Honor your father and your mother') than to conform to the prohibitions (e.g., 'You shall *not* kill, steal, etc.'). Second, the last two commandments in our list do not prohibit actions, they *prohibit desires*. Such norms will be left aside. Like most modern moral philosophers, we will restrict our discussions to rules concerned with how to act.

In relation to every imperative of the form 'Do A!' it is formally possible to ask: 'Why should I do A?'. If the answer is 'It is your duty to do B, and doing B implies doing A', then it becomes formally possible to ask: 'But why is it my duty to do B?' If the next answer brings in C, one can ask 'Why C?', and so on. In order to justify a substantive moral norm, one has in some way to end this justificatory regress somewhere. If there are duties at all, then there has to be at least one duty that is self-justificatory.

To traditional Christian believers, lines of moral justifications end, first, with the Ten Commandments, and then – absolutely – with the answer: 'It is your duty to follow the Ten Commandments because God has said it is your duty', period. Most philosophers, however, find this answer question-begging. They want an answer also to the question 'Why should it be my duty to do what God says it is my duty to do?' The philosopher Immanuel Kant (1724-1804) is famous for having tried to exchange 'God' for 'a will enlightened by reason', a *rational will*. Although being himself a firm believer, Kant wanted to make the basic ethical norms independent of religion. In this undertaking, he claimed to have found a central principle, The Categorical Imperative (soon to be presented), by means of which other presumed categorical norms could be tested. At the end of the justificatory line that Kant proposes, we find: 'It is your duty to follow The Categorical Imperative and the commandments it gives rise to because every rational will wants this to be its duty'. In order to arrive at what we will label 'Kant's Four Commandments', Kant argued somewhat as in the brief reconstruction below; the words are for a while put directly into Kant's mouth.

Step 1: All basic moral norms have to be linguistically formulated as *categorical imperatives*: 'Do A!' or 'Don't do A!'. Hypothetical

imperatives such as '*If* you want A, *then* you have to do B!' and '*If* you want A, *then* you cannot do B!' can never in themselves state a basic moral norm. Why? Because they make the prescription dependent on a pre-given goal (A), and if this goal is a mere subjective whim, the prescription has nothing to do with morals. On the other hand, if there is a categorical norm that requires the doing of A, then this norm bears the moral burden, not the hypothetical imperative 'if A then B'. The imperative 'If you want to be a good doctor, then you have to develop your clinical and ethical skills' is a hypothetical imperative; it does not state that you have to try to become a good doctor.

Step 2: A basic moral norm cannot pick out anything that is spatiotemporally specific. Sensuous desires can be directed at one or several particular spatiotemporal objects ('I want to play with him and no one else!'), but reason and rational wills can directly be concerned only with a-temporal entities such as concepts, judgments, logical relations, and principles. Therefore, no *fundamental* moral rule can have the form '*This* person should do A!', '*I* should do A', or 'Do A at *place x* and *time t*'. Basic norms must be universal and have the form 'Do A!' or the form 'Don't do A!' Therefore, they cannot possibly favor or disfavor a particular person as such.

Step 3: If there is a moral norm, then it must in principle be possible for persons to *will* to act on it. There can be morals only where there is freedom; purely instinctual reactions and other forms of pre-determined behavior is neither moral nor immoral. Since the basic norms have to be stated in the form of categorical (step 1) and universal (step 2) imperatives, the last presupposition of a special 'causality of freedom' has the following consequence: in relation to a norm, it must not only be possible that I as a particular person want to conform to it, it must also be possible that the norm in question is *willed by everyone*. Therefore, when you are trying to find out whether you can will to act in conformance with a certain 'action maxim', you cannot think only of yourself, you have to ask yourself whether you think that the maxim can in principle be collectively willed. If your answer is 'yes', you can will that it becomes a law for everyone. That is:

- Act only according to that maxim (= categorical general imperative) by which you can also will that it would become a universal law.

This is *The Categorical Imperative*. It is not a substantial imperative on a par with the Ten Commandments, but a test that substantial imperatives have to pass in order to be able to count as duties.

Step 4: Since the requirement stated is very formal, it should be called 'the *form* of the categorical imperative'. Nonetheless, we can derive from it a more substantial formulation, 'the *matter* of the categorical imperative':

- Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end.

This imperative of Kant is similar to the Christian Golden Rule: 'Do unto others as you would have them do unto you' (Jesus, according to Matthew 7:12), but it is clearly distinct from the negative imperative 'Do *not* unto others as you would *not* have them do unto you', which exists in other religions too.

Kant's imperative does not say that we are never allowed to use another human being as means; only that we are never allowed to use another person *only* as a means. On the other hand, this absolute respect for the autonomy of persons means that we are not even allowed to reduce our own autonomy, which, for instance, we can do by drinking alcohol or taking drugs.

When Kant derives the second formulation of his categorical imperative from the first, he seems to rely heavily on what he puts into his notion of a 'will'. According to Kant, only *persons*, i.e., mature responsible ('mündige') human beings, can have the kind of will that is mentioned in the first and formal formulation of the categorical imperative. Such a free will can create an end for itself merely by willing it, and Kant seems to falsely think that since persons are able to freely create ends for themselves, persons are *thereby* also ends in themselves. In Kant's opinion, animals are not persons even though they can have conscious perceptions,

desires, and feelings of pleasure and pain. His reason is that they cannot be persons since they cannot self-consciously will anything.

(Out of the two imperatives stated, Kant then derives a third formulation, which has quite a political ring to it. It says that it is always one's duty to act as if one were 'a legislating member' in 'the universal kingdom of ends', i.e., in the formal or informal society constituted by all persons. This is Kant's *form-and-matter* formulation of his categorical imperative.)

Step 5: Kant gives four examples of categorical general imperatives (maxims) that he thinks pass the test of the form of the categorical imperative, and which, therefore, should be regarded as absolute duties. They are: 'Do not commit suicide!', 'Develop your talents!', 'Do not make false promises!', and 'Help others!' Unlike the duties of the Decalogue, these commandments are not 'duties for God's sake', they are 'duties for duty's sake'. To a true Kantian, the question 'Why is it my duty to do what my rational will tells me to do?' is a nonsensical question. All that can possibly be said has already been said. Moral justification has reached its end point.

In the second formulation of the categorical imperative, Kant uses the expression (with our italics) 'treat humanity, whether in *your own person* or in the person of *any other*'. Some of one's duties are concerned only with oneself, and some of them are concerned only with other persons. If this distinction is crossed with another distinction that Kant makes, namely one between *perfect* and *imperfect* duties (soon to be commented on), then his four substantial norms can be fitted into the following fourfold matrix:

	Duties to oneself	*Duties to others*
Perfect duties	Do not commit suicide!	Do not make false promises!
Imperfect duties	Develop your talents!	Help others!

Table 1: *Kant's Four Commandments.*

The two *perfect duties* are prohibitions, and the two *imperfect duties* are obligations. The distinction may have something to do with the fact that it is often easier to know how to conform to a prohibition than how to act in order to fulfill an obligation, but it may also have to do with the fact that – no doubt – one can conform to the perfect duties even when one is asleep or is taking a rest. Imperfect duties in Kant's sense must not be conflated with what in moral philosophy is called 'supererogatory actions', i.e., actions that are regarded as good or as having good consequences, but which are *not* regarded as *duties*. Donating money to charity is mostly regarded as a supererogatory action; there is no corresponding duty. One may also talk of actions that are 'juridically supererogatory'; they are not required by the law but are nonetheless in the spirit of the law. It might be noted that most patients who praise the job of their doctor, would probably rather like to hear him respond by saying 'I worked a bit extra hard because you were suffering so much', than simply 'I did it out of duty'. This is not odd. Many patients want a somewhat personal relationship, but duties are impersonal. To work, out of empathy, more and harder than duty requires is good, but it is a supererogatory goodness.

It is hard to see and understand exactly how Kant reasons when he tests his four commandments against The Categorical Imperative and lays claim to show that they have to be regarded as absolute duties. Nonetheless we will try to give at least the flavor of the way he argues; we will in our exposition take 'Do not make false promises!' as our example.

In Chapter 4.2, we presented the structure of 'reductio ad absurdum' arguments. In such arguments, a certain view or hypothesis is shown to imply a logical contradiction or to contain some other absurd consequence; whereupon it is claimed that, therefore, the negation of the original view or hypothesis has to be regarded as true. Kant tries to do something similar with imperatives.

Let us assume (contrary to what we want to prove) that no one has to take the imperative 'Do not make false promises!' as stating a duty. This means that no one would feel compelled to fulfill their promises; and, probably, also that many people would allow themselves to make false promises as soon as they would benefit from it. It may then happen, that people so seldom fulfill their promises that no one dares to take a promise seriously any more. In a community where this is the case, the citizens can

no longer make any real promises; they cannot even fool anyone by making a false promise, since the whole institution of promising has become extinct. (Another similar case is a society where each and every coin might be a fake, and every person knows this. In such a society, there would be no real money any more, and, because of this, nobody could fool anyone with fake money.) According to Kant, it is absurd for a rational person to will that his community should contain the possibility of having no institution of promising. Therefore, a rational will must will that the maxim ‘Do not make false promises!’ becomes a universal law.

In the modern world, bureaucrats and other appointed rule watchers can sometimes spontaneously vent a Kantian way of thinking. Often when they become confronted by a person who wants them to make an exception to the rules they are meant to apply, they ask rhetorically: ‘Now, look, what do you think would happen if I and my colleagues should always do like this?’ When asking so, they think or hope that the person in front of them shall immediately understand that this would mean – with very negative consequences – the death of the whole institution. ‘What do you think would happen with the parking system’, the parking man asks, ‘if I and my colleagues did not bother about the fees?’ The dummy way to understand Kant’s universalistic moral reasoning is to take him to ask rhetorically in relation to every person and every culture:

- What would your society, and thereby your own life, look like if everyone gave false promises all the time?
- What would your society, and thereby your own life, look like if everyone committed suicide?
- What would your society, and thereby your own life, look like if everyone neglected his talents all the time?
- What would your society, and thereby your own life, look like if no one ever helped any other person at all?

Modern physicians have to fulfill many roles, often also that of the bureaucrat and the rule watcher. Therefore, they may often say to patients: ‘What do you think would happen if I and my colleagues should always do the way you want me to do?’ A patient may, with some good reasons, request his doctor to give him sick leave even though he is not suffering

from any well-defined disease; well knowing that it is possible for this specific doctor to deceive the regional social insurance office. But if all physicians provide false certificates whenever a patient benefits from such a certificate, medical certificates would probably after some time become completely futile.

A contemporary American philosopher, Alan Gewirth (1912-2004), has tried to ground norms in a way that makes him similar to Kant. He claims that human agency is necessary evaluational, in the sense that all actions involve the agent's view of his purposes as being good; at least in the sense that he wants to see them realized. Now different agents have, of course, different purposes, but common to all agents, according to Gewirth, is that they must regard the freedom and well-being necessary to all successful action as necessary goods. Hence, since it would be logically inconsistent for an agent to accept that other agents may interfere with goods that he regards as necessary, he must claim rights to freedom and well-being. And since it is his being an agent that makes it necessary for him to make this claim, he must also accept the universalized claim that all agents equally have rights to freedom and well-being. Accordingly, a universalist theory of human rights is derived from the individual agent's necessary evaluative judgement that he must have freedom and well-being.

The main problem for classical deontology is its inflexibility. In relation to almost every categorical norm, there seems to be possible cases where one ought to make one's action an exception to the norm. For instance, most people believing in the Ten Commandments have with respect to 'You shall not kill!' made an exception for wars and war-like situations; and in relation to the norm 'Don't lie!' everyday language contains a distinction between 'lies' and 'white lies'. Especially in medicine, it is quite clear that physicians sometimes can help patients by means of (white) lies. A related problem is that it is easy to conceive of situations where deontological norms (say, 'Help others!' and 'Don't lie!') conflict; in such a situation one has to make an exception to at least one of the norms.

A famous modern attempt to leave this predicament behind but nonetheless remain a deontologist (i.e., not bring in the consequences of one's actions when evaluating them) has been made by W. D. Ross (1877-1971). His view is best understood by means of an analogy with the particle mechanics of classical physics. If, by means of this theoretical

framework, one wants to understand in principle or predict in detail the motion of a specific particle (call it ‘Alfa’), one obtains situations such as the following four. (i) Alfa is affected only by a gravitational force from one other particle (Beta). To follow this force (Newton’s law of gravitation) will then, metaphorically speaking, be the duty of the particle. However, (ii) if Alfa is similarly affected by the particle Gamma, then it is its duty to adapt to the gravitational forces from both Beta and Gamma. Furthermore, (iii) if Alfa, Beta, and Gamma are electrically charged, then it is the duty of Alfa to adapt to the corresponding electric forces (Coulomb’s law) too. In other situations, (iv) there can be many more particles and some further kinds of forces. All such partial forces, even those of different kinds, do automatically combine and give rise to a determinate motion of Alfa. In classical mechanics, this adding of *partial forces* is represented by *the law for the superposition of forces*. By means of this law, the physicists can add together all the partial forces into a *resultant force*, by means of which, in turn, the movement of Alfa can be directly calculated.

Let now Alfa be a person. According to Ross, he can then in a certain situation both be subject to many kinds of duties (partial forces), which Ross calls ‘prima facie duties’ (i.e., duties on ‘first appearance’), and have duties towards many persons (particles). When Alfa has taken the whole situation into account, there emerges an *actual duty* (‘resultant force’), which Alfa has to perform quite independently of the consequences both for him and for others. Without claiming completeness, Ross lists six classes of prima facie duties: (1) duties depending on previous acts of one’s own such as ‘Keep your promises!’ and ‘Repair your wrongdoings!’; (2) duties having the general form ‘Be grateful to those who help you!’; (3) duties of justice; (4) duties of beneficience; (5) duties of self-improvement; and (6) duties of not injuring others. A specific prima facie duty becomes an actual duty when it is not overridden by any other prima facie duty.

Ross’ move may at first look very reasonable, but it contains a great flaw. No one has so far been able to construe for this ethical system what the principle of superposition of forces is in classical mechanics. That is, in situations where more than one prima facie duty is relevant, Ross supplies no rule for how to weigh the relevant duties in order to find out what our actual duties look like; here he resorts to intuition. As discussions after Ross has shown, there are even good reasons to think that no

‘superposition principle for prima facie duties’ can ever be construed. We will present this case in Chapter 9.3.

*****

Even if Ross’ proposal to exchange absolute duties for prima facie duties should, contrary to fact, be completely without problems, it leaves a big moral problem untouched. It takes account only of how one man should find out how to act morally in a given particular situation, but modern states and institutions need verbally explicit general norms in order to function; states need written laws. How to look upon such necessary rules from a moral point of view? If singular actions can be morally right, wrong, or indifferent, the same must be true of institutional norms; and it seems incredible to think that all of them are morally indifferent. We now meet from a societal perspective the kind of justificatory regress that we have already pointed out in relation to the Ten Commandments and Kant’s categorical imperative. It now looks as follows.

Institutional laws might be regarded as justified as long as they stay within the confines of the state in which they are applied. Most modern societies then make a distinction between ordinary laws and some more fundamental and basic laws, often called ‘constitutional laws’. The latter not only regulates how ordinary laws should be made (and, often, also how the institutional laws themselves are allowed to be changed), they are also meant in some way to justify the procedures by means of which the ordinary laws come into being. But what justifies the constitutional laws themselves? Shouldn’t the power that enforces our laws be a morally legitimate power? Even if some of the constitutional laws can be regarded as being merely hypothetical imperatives, i.e., rules that function as means to a pre-given goal such as to find the will of the citizens, this cannot be true for all of them; at least not if everything is made explicit. Then there should (in the exemplified case) be a categorical norm that says that the citizens shall rule. When there is a bill of rights connected to a constitution, then the justificatory question appears at once: how are these rights justified? For instance: how is the Universal Declaration of Human Rights of the UN justified?

This problem is not necessarily a problem only in relation to the society at large. Several medical rules have a moral aspect to them, and this needs moral justification. Neither the rule to respect the autonomy of patients nor the rule that requires informed consent in clinical research are rules that are meant only to be efficient means for the work of doctors and researchers.

A way to solve this rule-justificatory problem in a non-religious and in a partly non-Kantian way has been proposed by the two German philosophers, Karl-Otto Apel (b. 1922) and Jürgen Habermas (b. 1929). It is called '*discourse ethics*', and they regard it as a modern defense of deontology. Instead of trying to ground the basic categorical norms in what a rational will has to will, they try to ground them in what a group of people who want to communicate with each other have to regard as an ideal form of behavior. Like most late-twentieth century philosophers, they think that language is necessarily a social phenomenon, and this creates in one respect a sharp contrast between their position and that of Kant. The free will and the reason that Kant speaks of can belong to a mature person alone, but the pragmatic language principles and the communicative reason that Apel and Habermas focus on are necessarily inter-personal.

We will now first state Habermas' version of the discourse ethic counterpart to Kant's first (formal) formulation of The Categorical Imperative, and then try to explain why discourse ethicists think that some pragmatic language principles are at one and the same time both necessary presuppositions for linguistic communication (discourses) and categorical imperatives. Like Kant's categorical imperative, Habermas' so-called '*principle of universalization*' is a formal principle against which all more substantial norms are meant to be measured. Whereas Kant says that 'a norm is valid if and only if: it can be rationally willed that it would become a universal law', Habermas says:

- a norm *is valid* if and only if: the consequences and side effects of its general observance for the satisfaction of each person's particular interests are acceptable (without domination) to all.

How is this deontological universal principle taken out of the wizard's hat? Let us look at it in the light of the norms (a) 'When talking, do this in order to be understood!', (b) 'Don't lie!', and (c) 'Give other people the

rights that you yourself claim!' Most individual language acts take something for granted. If someone asks you 'Have John stopped smoking yet?', he takes it for granted that John have been a smoker; otherwise his question makes no sense. Therefore, you might reason as follows: (1) there is this question about John; (2) *how is it possible* to put forward?; (3) answer: the speaker presupposes that John has been a smoker. In a similar way, Apel and Habermas make inferences from more general facts to presuppositions for the same facts. Their so-called 'transcendental inferences' can be schematized as follows:

(1) fact: there is communication by means of language
(2) question: how is such communication possible?
(3) answer: by means of (among other things) pragmatic language principles
(4) norm derivations: (a) if people never give the same words the same meaning, language stops functioning, therefore the pragmatic language principle 'You ought to give words the same meaning as your co-speakers!' is at the same time a valid norm; (b) if everybody lies all the time, language stops functioning, therefore the pragmatic language principle 'Don't lie!' is at the same time a valid norm.

These norms are said to be to be *implicitly present* in our language. What the discourse ethicists mean by this claim might be understood by means of an analogy (of ours). Assume that you are a zoologist that has discovered a completely new species and that after careful observations you have come to the conclusion that all the individuals you have observed have to be regarded as sick. One might then say that you have found an *ideal* (to be a healthy instance of the species) implicitly present in these really existing (sick) individuals. Discourse ethicists have come to the conclusion that human languages have an ideal built into their very functioning. In the ideal language community there are no lies. When people lie or consciously speak past each other, they do not kill language, but they make it deviate from its in-built ideal.

Here is another discourse ethical 'transcendental inference', and one whose conclusion comes closer to Habermas' principle of universalization:

1. fact: there are argumentative discourses
2. question: how are such discourses possible?
3. answer: by means of (among other things) pragmatic language principles that are specific for argumentative situations
4. norm derivation: (c) if *by arguments* you try to show that your opponent does not have the same rights in the discussion as you have, then you have to accept that there may be arguments that take away your own right to free argumentation; therefore, the pragmatic principle 'You should in discussions recognize the other participants as having rights equal to your own!' is at the same time a valid norm.

The last norm says that the ideal for discussions is that they are free from domination; they should only contain, to take the German phrase, 'Herrschaftsfreie Kommunikation'. If this norm is accepted, then it seems (for reasons of philosophical anthropology) natural to claim that all norms that can be communicatively validated have to take account of, as Habermas says: 'the satisfaction of each person's particular interests'. This means, to use the earlier analogy, that we cannot make our sick language healthy, and have an ideal communicative community, without a normative change in the whole community.

Habermas' principle of universalization is impossible to apply in real life situations (since it talks about *all* persons, consequences, and side effects), and he is aware of this fact. But he claims that this principle can function as a regulative idea that make us try to move in the right moral direction, and that we ought to approximate the principle when we put forward constitutional laws, bills of rights, or other kinds of basic moral norms. In order to come a bit closer to real life he has put forward another principle, which he calls 'the principle of discourse ethics':

- a norm can *lay claim to validity* if and only if it meets (or can meet) with the approval of all affected in their capacity as free and equal participants in a practical discourse.

In practice, we will never be able to finally validate a norm. Therefore, we have to rest content with a principle that tells us when we can at least lay claim to validity. According to the principle of discourse ethics, no number of thought experiments can be more validating than a real communicative exchange around the norm in question. For instance, without speaking with other people in your community you cannot find out whether promise breaking should be regarded as always forbidden or not. Note, though, that even if every discussion of such a norm takes place under certain given circumstances, it should aim (according to discourse ethics) at finding a moral norm whose validity is not culture bound, but is universal.

Discourse ethics is like Kant's duty ethics an ethical system that lays claim to be (i) deontological, (ii) in its centre formal, (iii) universalistic in the sense of not being culture bound, and (iv) cognitivist, since it thinks that the validity of moral statements can be rationally discussed. It differs from Kant in having exchanged a person-located reason for an inter-personal communicative reason.

At last, we want to point out a difference between Apel and Habermas. Apel is very much like Kant a philosopher who thinks he has found something that is beyond all doubt, whereas Habermas is a fallibilist. The problem for Apel is that he can *at most* lay claims to having shown that we are faced with an alternative: either to shut up or to accept the existence of norms. But then his norms are not true categorical imperatives, only hypothetical imperatives ('*if* you want to use language, *then* …'). Habermas is an outspoken fallibilist, who thinks that his principles may contain mistakes, but that, for the moment, they cohere quite well with most of our present moral particular judgments. In this opinion, he comes close to John Rawls' more elaborate views on what fallibilism in moral matters amount to. For our presentation of this view, and Rawls' famous notion of 'reflective equilibrium', the reader has to wait until Chapter 9.3; let it here only be noted that the very term 'moral fallibilism' is used neither by Habermas nor by Rawls.

9.2 Consequentialism

'Consequentialism' is an umbrella term for ethical systems that claim that it is only or mainly the *consequences* of a particular action (or kind of

action) that determines whether it should be regarded as being morally right, wrong, or neutral. Each specific consequentialist system has to state what kind of consequences should be regarded as being good, bad, or indifferent. In so far as good and bad consequences can be weighed against each other, the natural consequentialist counterpart to Kant's categorical imperative (the second formulation) can be stated thus:

- Among the actions possible for you, choose the one that maximizes the total amount of good consequences and minimizes the total amount of bad consequences that you think your action will (probably) produce.

In retrospect, one may come to the conclusion that what one thought was the objectively right action in fact was not so. One had estimated the consequences wrongly. Nonetheless, the basic imperative must have this form. It is impossible to act retrospectively, and one has to act on what one believes. The term 'probably' is inserted in order to make it explicitly clear that there is never in real life any complete and certain knowledge about the consequences spoken of. Rather, the epistemological problems that pop up when consequentialist imperatives shall be applied are huge. Because of these problems, there are in some contexts reasons to discuss whether or not one should stick to one of two more unspecific principles, one (the first below) states the thesis of 'positive consequentialism' and the other the thesis of 'negative consequentialism':

- Act so that you produce as much good consequences as possible.
- Act so that you produce as little bad consequences as possible.

The second principle is often given the formulation 'we demand the elimination of suffering rather than the promotion of happiness' (Karl Popper). It is claimed that at least politicians should stick to it. Why? Because, it is argued, if they make false predictions about the outcome of their policies, the consequences of acting on the reducing-suffering principle cannot be as disastrous as they can be when acting on the creating-happiness principle.

Most important and most well known among consequentialist systems are the utilitarian ones. We will briefly present four of them; they differ in how they define what is good and bad. They are (with their most famous protagonist within parenthesis):

1. simple hedonistic utilitarianism (Jeremy Bentham, 1748-1832)
2. qualitative hedonistic utilitarianism (John Stuart Mill, 1806-1873)
3. ideal utilitarianism (George Edward Moore, 1873-1958)
4. preference utilitarianism (Richard M. Hare, 1919-2002; Peter Singer, b. 1946).

According to simple hedonistic utilitarianism, pain and only pain is bad, and pleasure and only pleasure is good (or has utility). The term ‘utility’ is a bit remarkably made synonymous with ‘pleasure’; ‘hedone’ is the Greek word for pleasure. Normally, we distinguish between many different kinds of pleasures (e.g., in sex, in work, in music, etc.) and pains (aches, illnesses, anxieties, etc.), but according to Bentham this is at bottom only a difference in what causes the pleasures and pains, respectively. These can, though, differ with respect to intensity and duration. When the latter factors are taken into account, utilities can, he thinks, in principle be added (pains being negative magnitudes), and hedonic sums for amounts of pleasures and pains be calculated; the utility calculus is born. He even thinks that factors such as the certainty and the proximity of the expected pleasures and pains can be incorporated in the calculus. If we leave the latter out of account, his central thought on these matters can be mathematically represented as below.

Let us associate all pleasures with a certain positive number, p, and all pains with the negative number, -p. Furthermore, let us assume that each different degree of intensity of pleasure and pain can in a reasonable way be associated with a certain number, i_l. Let us then divide the life of a certain person during the temporal interval T into m number of smaller time intervals, called t_1 to t_m; these intervals shall be so small that they do not contain any changes of pleasures, pains, or intensities. All these things taken for granted, the actual total pleasure or the hedonic sum – $h_a(T)$ – for the life of the person *a* in the interval T can be mathematically represented

by formula 1 below. Formula 2 represents the aggregated pleasure for all sentient beings in the interval T:

$$(1)\quad h_a(T) = \sum_{n=1}^{n=m} t_n \cdot (i_l \cdot p) \qquad (2)\quad H(T) = \sum h_a(T)$$

(Formula 1 should be read: the total pleasure of person *a* in time interval T equals the sum of the pleasures in each time interval, 1 to m; the pleasure in each such interval being given by multiplying the length of the time interval with the intensity of the pleasure and the value of the pleasure. Formula 2 should be read: the total pleasure of all sentient beings in T equals the sum of all their individual pleasures.)

Bentham thinks it is our duty to try to maximize H(T) for a T that stretches from the moment of action as long into the future as it is possible to know about the consequences of the action. His categorical imperative is *the utility principle*:

- Among the actions possible for you, choose the one that maximizes the utility that your action produces.

Bentham makes it quite clear that he thinks that even animals can experience pleasure and pain, and that, therefore, even their experiences should be taken into account in the 'total amount of pleasure and pain' spoken of. Utilitarianism brought from the start animals into the moral realm. Nonetheless there is an ambiguity in the utility principle: should we relate our maximizing efforts only to the pleasures and pains of those sentient being that already exist (and probably will exist without extra effort on our part), or should we try to maximize pleasure absolutely, i.e., should we even try to create new sentient beings *only* for the purpose of getting as much pleasure in the world as is possible? In the latter case we might be forced to maximize the number of members of those species whose members normally feel more pleasure than pain during their lives, and to minimize the number (might even mean extermination) in species whose members normally feel more pain than pleasure (even if they are human beings). In what follows we will mainly write with the first

alternative (‘the prior existence view’) in mind, even though the second alternative (‘the total amount view’) is literally closer to the utility principle.

Despite utilitarianism’s explicit concern with animals, Bentham’s view is often regarded as being possible to condense into the so-called ‘greatest happiness principle’:

- One should always act so as to produce the greatest happiness for the greatest number of people.

As should be obvious from the formulas 1 and 2, the task of comparing for a certain action all possible alternative hedonic sums, H(T), or even a very small number of them, is an infinite task. For instance, how are we supposed to reason when applying the utility principle in clinical practice; especially in emergency situations which do not leave much room for reflection? Only extremely rough and intuitive utility estimates can be made – and questioned. Nonetheless, the utility principle immediately stirred the minds of social reformers. And it is easy to understand why: the principle makes no difference between different kinds of people, e.g., between noble men and ordinary men. Pleasures and pains are counted quite independently of what kind of person they reside in. The toothache of a beggar is regarded as being quite as bad as a similar toothache of a king. A famous dictum by Bentham (that leaves animals aside) is: ‘each to count for one, and none for more than one’.

John Stuart Mill was heavily influenced by Bentham and his utilitarian father, James Mill, who was a close friend of Bentham. But Mill the junior found the original utilitarianism too simple. Pleasures, he claimed, can differ not only in their causes and their intensity, but in *kind*, too. Broadly speaking, in Mill’s opinion, there are two qualitatively distinct realms of pleasures. One lower realm of physical or almost-physical kinds of pleasures, some of which we share with animals, and one higher realm. The latter contains more intellectual kinds of pleasures, pleasures that are typical of educated human beings. And Mill was not alone. Many of Bentham’s contemporaries regarded his utilitarianism as a doctrine worthy only of swine or of small children. Shouldn’t we distinguish between the sensibilities of a Socrates, an infant, or a pig? The pleasure of playing

push-pin, it was said, is always a lower kind of pleasure than the pleasure in reading good literature. Put in modern jargon, Mill's view is that the pleasures of sex, drugs, and rock-and-roll should count lower in a utility calculus than the pleasures of reading, writing, and listening to Mozart. The higher pleasures were said to give rise to happiness, the lower ones merely to contentment.

To illustrate further the difference between Bentham and Mill we might look at a thought experiment provided by Roger Crisp. He asks you to think that you are about to be born, and are given the choice between being born as an oyster or as a great composer such as Haydn. Your life as an oyster would, in all probability, be a safe, soft, and long life; you would be floating in a moderate sensitive pleasure without any pains for some two hundred years. Your life as a composer, on the other hand, would probably be very intense, complex, creative, and exciting, but also much shorter and contain phases of severe suffering. What would you choose? Bentham would perhaps choose to be born as an oyster, but Mill would surely prefer the shorter life of a Haydn. Even though the oyster-versus-Haydn example is extreme, it might nonetheless illustrate a problem in clinical medicine: is it better for terminally ill patients to stay alive in a long persistent vegetative state, or is it better for them to live a more ordinary life for only a brief period of time?

What we have termed *simple* hedonistic utilitarianism is sometimes called *quantitative* utilitarianism. We have avoided the latter label, since it gives the false impression that Mill's qualitative utilitarianism, which posits the existence of *qualitatively different kinds* of pleasures, cannot be given a mathematical representation. But it can. All we have to assume is that in a reasonable way we can associate with each and every *kind* of pleasure and pain a certain positive (pleasure) or negative (pain) number, p_k. Mill's higher pleasures should then be attributed large positive numbers and the lower pleasures small positive numbers. On the assumption of such a pleasure-and-pain (utility) scale, and on the assumptions earlier stated, the two utility formulas presented are transformed into the following ones (formula 1' for the total amount of pleasure of one person in the interval T, and formula 2' for the aggregated pleasure of all sentient beings):

$$(1') \quad h_a(T) = \sum_{n=1}^{n=m} t_n \cdot (i_l \cdot p_k) \qquad (2') \quad H(T) = \sum h_a(T)$$

(Formula 1’ should be read: the total pleasure of person *a* in time interval T equals the sum of the pleasures in each time interval, 1 to m; the pleasure in each such interval being given by multiplying the length of the time interval with the intensity of the pleasure and the value of the specific kind of pleasure in question. Formula 2’ should be read: the total pleasure of all sentient beings in T equals the sum of all their individual pleasures.)

Qualitative hedonistic utilitarianism differs from simple utilitarianism not only in its acceptance of many different kinds of pleasures, since this difference has at least two repercussions. First, the existence of different kinds of pleasures makes the original utility calculus much more complicated. Second, the difference affects what persons can be set to make utility calculations. Every person who constructs a utility scale for qualitative utilitarianism must be familiar with every kind of pleasure and pain that he ranks, which means that only persons who are familiar with both the higher and the lower kind of pleasures can make the calculations. Mill took it for granted that those who know the higher pleasures normally also know all the lower pleasures, and that he himself could rank all kinds of pleasures. This might be doubted.

Both Bentham and Mill were, like Kant, part of the Enlightenment movement, and all of them tried, each in his own way, to connect moral thinking with political thinking in order to promote the creation of good societies.

Mill’s line of thinking about utilitarianism was taken one step further by Moore, who argued that some ideal things are good without necessarily being connected to any pleasure. According to Moore, having knowledge and experiencing beauty are good independently of the pleasure they may cause. One consequence of this view is that since ‘ideal utilities’ are (normally considered to be) quantitatively incommensurable with ‘pleasure utilities’, to accept ideal utilitarianism is to drop the whole of classical monistic utilitarianism in favor of a pluralistic utilitarianism.

To claim that one is interested in (has a preference for) having knowledge even when it does not gives rise to pleasure, is to introduce a

distinction between *pleasure* and *preference satisfaction*. Such a distinction was foreign to the classical utilitarianism of Bentham and Mill. They thought that everything that human beings directly desire can adequately be called pleasure, and that every true desire satisfaction gives rise to some pleasure. But, with Moore, we may well question this view. Therefore, many utilitarian philosophers have tried to create forms of utilitarianism where utility is defined in terms of preference satisfaction instead of pleasure. Of course, these philosophers take it for granted that we often have pleasure from preference satisfactions and pain from preference dissatisfactions; the point is that a utility calculus *also* has to take account of preference satisfactions and dissatisfactions where this is not the case. For instance, if a physician has a preference for being regarded as a good physician but *unknowingly to himself* is regarded as bad, shouldn't this preference dissatisfaction of his count, even though there is no experienced displeasure?

The shift from pleasure utilitarianism to preference utilitarianism does not affect the general formulation of the utility principle; it only redefines utility. Unhappily, however, neither does the shift solve the two outstanding internal problems of utilitarianism:

- How to rank all the different kinds of preference satisfactions (utilities)?
- Given such a ranking, how to make a useful utility calculation in real life?

One obvious objection to preference utilitarianism is that our conscious preferences do not always mirror our real preferences. Sometimes, to our own surprise, we do not at all become pleased when a preference of ours becomes satisfied. For instance, one may be perfectly convinced that one wants to read a good book alone at home in the evening, and so one satisfies this conscious preference. But when one is sitting there reading the – no doubt – good book, one discovers that this was not at all what one really wanted to do. How should preference utilitarianism take such facts into account? It has been argued that only preferences based on informed desires that do not disappear after therapy should count (Brandt 1979). Such amendments, however, make the epistemological problems of

utilitarianism grow even more exponentially than they did with the introduction of Mill's qualitative utilitarianism.

A peculiar problem for utilitarians is that they have to try to take account also of the fact that there are many people who intensely dislikes utilitarian thinking. Since the latter become deeply dissatisfied when utilitarians act on a utility calculation, the utilitarians have to bring even this fact into their calculations. One famous utilitarian, Henry Sidgwick (1838-1900), has argued that from a utilitarian point of view it is hardly ever right to openly break the dominating norms of a society. In a weaker form, the same problem appears every time a utilitarian intellectual shocks his contemporaries with a proposal for a radically new norm. How does he know that the immediate amount of displeasure that his proposal gives rise to among traditionalists, is outweighed by future preference satisfactions that he thinks will come about if his proposal is accepted? Has he even tried to find out? If not, then he either is inconsistent or has implicitly some non-utilitarian norm up his sleeves.

However, reflections by utilitarians can have quite an impact in spite of the fact that most consequence estimations have better be called consequence conjectures or consequence speculations. This is so because a utilitarian's views on what should be regarded as having equal utility, or being equal amounts of preference satisfaction, can contradict the traditional moral outlook of his society. With the claim that pleasures and pains ought to be considered independently of who has them, classical utilitarianism questioned values that were central to feudal societies. Modern preference utilitarianism, particularly in the hands of Peter Singer, has continued a utilitarian tradition of re-evaluation of contemporary common sense morality. In Singer's case, the moral reform proposals arise from the way he connects certain anthropological views with a specific evaluation of the difference between self-conscious long-term preferences on the one hand and desires that want immediate satisfaction on the other. He thinks along the following lines.

Assume that rats like human beings, (a) can have conscious feelings of pleasure and pain, and that they instinctively in each particular moment seeks immediate pleasure and tries to avoid immediate pain; but that they unlike human beings, (b) are completely unable consciously to imagine future states where they have feelings of pleasure and pain or can have an

interest satisfied or dissatisfied. Now, given these assumptions, what is from a preference utilitarian point of view the difference between painlessly killing a rat and killing a human being? Answer: at the moment of the deaths, with the rat disappears *only* the desire for the immediate pleasure satisfactions, but with the human being disappears *also* the desire for the satisfactions of all his hopes, wishes, and plans for the future. With the death of a healthy human being more possible satisfactions of *existing* preferences disappears than with the death of a healthy rat.

Singer claims that both rats and human beings have a *right to physical integrity* because they are able to suffer, but that – with some qualifications – we human beings also have a *right to life* because we normally anticipate and plan our future; to human beings applies 'the journey model of life'. Now the qualifications indicated. According to Singer's (rather reasonable) anthropology, fetuses, very small infants, and severely disabled people lack the anticipatory ability in question. But this implies, he argues, that in certain special circumstances painless abortion and infanticide can be justified. One possible kind of case would be the killing of very small infants whose future life in all probability would be full of suffering both for themselves and for their parents.

To many people these views mean an unacceptable *degrading of some forms of human life*, but this is only one side of Singer's reform coin. On the other side there is an *upgrading of some forms of animal life*. The latter has made him a central thinker in the animal liberation movement. From the premises sketched, he argues that the United Nations should declare that great apes, like human beings, have a right to life, a right to the protection of individual liberty, and a right not to be tortured. Why? Answer: because even great apes can form long-term interests. Not to accept such rights is, Singer says, an act of *speciesism*, which is as morally wrong as racism and sexism. Note, though, that all of Singer's talk of 'rights' are made within a utilitarian, not a deontological, framework.

According to Singer, while animals show lower intelligence than the average human being, many severely retarded humans show *equally low* mental capacity; therefore, from a moral point of view, their preference satisfactions should be treated as equal too. That is, when it comes to experiments in the life sciences, whatever the rules ought to look like, monkeys and human infants should be regarded as being on a par. From his

premises, Singer has also questioned the general taboo on sexual intercourse with animals.

Singer thinks we have to consider equal interests (preferences) equally and unequal interests unequally; whatever species the individual who has the interests belong to. All sentient beings have an interest in avoiding pain, but not in cultivating their abilities. But there are even more differences to take into account. That two individuals of the same species want the same kind of thing does not necessarily mean that they have an equal interest in it. Singer regards interests as being subject to a law of diminishing marginal utility returns. If two persons, one poor and one rich, both want to win 100 000 euros in a lottery, then if the poor wins the preference satisfaction will be higher (and therefore his interest is higher) than if the rich one wins. Similarly, a starving person has a greater interest in food than someone who is just a little hungry. Singer relies on the following reasonable anthropological principle: the more one has of something, the less preference satisfaction one gets from yet another amount of the same thing. Therefore, according to Singer, anyone able to do so ought to donate part of his income to organizations or institutions that try to reduce the poverty in the world.

So far, we have mainly described utilitarianism as if it tries to say how we should think when we don't know what *individual action* is the morally right one. Such utilitarianism is called *act utilitarianism*, but there is also another brand: *rule utilitarianism*. According to the latter, we have only to estimate by means of what rules we can create the greatest possible happiness in the world; and then act on these rules. This means that these rules take on the *form* of categorical general imperatives; the main principle of rule utilitarianism can be stated thus:

- Act only according to those rules by which you think that the greatest happiness for all sentient beings can be achieved.

At first, it might seem odd to judge a particular action by means *not* of the consequences of the action itself, but by the consequences of a corresponding rule. But there is a good reason for the change: the tremendous epistemological and practical problems that surround all utility calculations. These problems seem to diminish if it is the consequences of

rules, rather than the consequences of actions that should be estimated. Singer is a rule utilitarian, and rule utilitarianism was suggested already by John Stuart Mill. The famous preference utilitarian, R. M. Hare (1919-2002), proposes a two-level utilitarianism. He says that 'archangels' can stay on *the critical level* where act utilitarian calculations are made, that 'proles' cannot raise above the *intuitive level* where situations are judged by means of already accepted rules, but that most human beings can now and then try to approximate the archangels.

These considerations apart, there is also a sociological reason for rule utilitarianism. As soon as an act utilitarian accepts that society at large and many of its institutions need explicit norms, he has to accept some rule utilitarian thinking.

It seems impossible that any rule utilitarian would come to endorse Kant's commandment 'Do not commit suicide!', but they may well end up with proposing Kant's three other commandments: 'Develop your talents!', 'Do not make false promises!', and 'Help others!' Similarly, the principle that doctors should treat patients as autonomous individuals might be regarded as justified by both deontologists and utilitarians. A Kantian may say that this principle follows from The Categorical Imperative, which says that we should always treat a person as an end in himself and never simply as a means; and a rule utilitarian may say the principle follows from The Utility Principle, since it obviously has good consequences such as an increase in the trust in the health care system, which, in turn, creates more happiness and health or at least prevent distrust and ill-health. The fact that people who subscribe to different ethical paradigms can make the same moral judgment of many individual actions is perhaps obvious, but the point now is that such people may even be able to reach consensus about moral rules. In everyday life, both these facts are important to remember when a moral-philosophical discussion comes to an end.

In Chapter 2.4 we said that paradigms and sub-paradigms are like natural languages and their dialects. It is as hard to find clear-cut boundaries between a paradigm and its sub-paradigms, as it is to find discontinuities between a language and its dialects; and sometimes the same holds true between two paradigms and between two languages. Nonetheless, we have to make distinctions between different paradigms and different languages. Now, even though it is easy to keep religious and Kantian deontology

distinct form all forms of consequentialism, it takes a real effort to see the difference between the deontology of discourse ethics and the consequentialism of preference rule utilitarianism. Their similarity can be brought out if Habermas' principle of universalization (U) and the preference rule utilitarian formulation of the utility principle (R) are reformulated a little; especially if utilitarianism is restricted to persons. Compare the following statements:

- (U) a rule *is morally valid* if and only if: the consequences and side effects of its general observance for the satisfaction of each person's particular preferences are acceptable to all

- (R) a rule *is morally valid* if and only if: the consequences and side effects of its general observance maximizes the preference satisfactions of all persons.

The principle (U) does not make moral validity directly dependent on consequences but on acceptability, but then, in turn, this acceptability depends on what consequences the rule has. Conversely, there seems to be something odd with the principle (R) if it is not acceptable to most people. As we will see in Chapter 9.3, utilitarianism has had problems with how to justify the utility principle.

Let it here be added that many discourse ethicists will, just like Singer, give rights even to animals. They then distinguish between the communicatively competent persons spoken of in their basic principles, which are both *moral subjects* and *moral objects*, and animals, which at most are moral objects. Persons are both norm validators and objects for norms, but animals can only be objects for norms. Human beings, however, can be moral advocates for beings that lack communicative competence. According to discourse ethics, between moral subjects there are real duties, but between persons and animals there can only be quasi-duties.

From a commonsensical point of view, rules such as 'Do not make false promises!', 'Help others!', 'Don't lie!', 'Don't steal!', and, in medicine, 'Do not make false certificates!' and 'Do not make transplants without consent!', have one positive side and one negative. The positive side is that

when such rules are enforced people can normally trust that others will help them when needed, that they will not be given false promises, and so on; and such trust characterizes good societies. The negative thing is the inflexibility. Indeed, it seems as if we often have good moral reasons to make exceptions to our norms. We will return to this issue in Chapter 9.3.

*****

No form of consequentialism can ever eliminate the problem of judging consequences, since consequentialism holds that it is *only or mainly* the consequences of an action that determines whether it is morally right or wrong. But judging consequences may be important also for other reasons. One often has from an egoistic point of view good reasons to try to find out what consequences one's actions may have; at least if one wants to be a bit prudent. Also, consequences may be given a *subordinate role* in some other kind of ethical system; a fact we have already noted in relation to discourse ethics, and we will return to it later in relation to virtue ethics. This being noted, we will here take the opportunity to present one way of trying to estimate and compare the values of various medical interventions. Such values, be they moral or not, simply have to be considered in clinical medicine and in health care administrations. We will present the method of evaluating consequences by means of so-called 'Quality-Adjusted Life Years' (QALYs). Some optimists hope that one day it will be possible – on a broad scale – to allocate healthcare resources by means of the costs per QALY of different interventions. If Bentham would have been living, he might have tried to estimate the cost per pleasure increase of different interventions.

A life year is of course one year of life, but what then is a *quality-adjusted* life year; and what magnitudes can it take? One year of perfect health for one person is assigned the value 1 QALY, the year of death has the value 0, and living one year with various illnesses, diseases, and disabilities obtain QALY-values between 0 and 1. The values between 0 and 1 are determined by various methods. People can be made to answer questions such as 'how many years in a certain bad state would count as equal to perfect health for some definite number of years?' If ten years in a diseased state is regarded as equal to two years with perfect health

(= 2 QALY), then each diseased year is attributed the value 0.2 QALY. If a terminally ill patient gets 10 years of 0.2 QALY/year with a certain medical intervention, and 3 years of 0.8 QALY/year without it, then, on the assumptions given, he should not be given treatment, since 2.4 QALY is better than 2.0 QALY. Different patients may evaluate different things differently, which means that when it comes to interventions in groups, health care administrations have to work with average QALYs.

Another way of determining what QALY number should in general be associated with a certain health state is to use standardized questionnaires; at the moment the EuroQol EQ-5D is ranked high. Here, the respondents are asked to tell whether they regard themselves as being good (= 1), medium (= 2), or bad (= 3) in the following five health dimensions:

- Mobility (M)
- Self-Care (S)
- Usual Activities (U)
- Pain/Discomfort (P)
- Anxiety/Depression (A).

Every answer is represented in the form of five numbers ordered as <M, S, U, P, A>, e.g., the value set <1, 3, 3, 2, 1>. There is then a ready-made table in which each such value set is associated with a single number between 0 and 1, called ‘the health state utility score’. One year with a utility score of 0.5 is reckoned as bringing about 0.5 QALY. If these measures are accepted, it is possible to compare a treatment A, which generates four years with a utility score of 0.75 (= 3 QALY), with another treatment B, which produces four years with a score of 0.5 (= 2 QALY), and immediately arrive at the conclusion that A is the better treatment; it is 1 QALY better.

QALY measurements of treatments can easily be combined with the costs of the same treatments. A ‘cost-utility ratio’ for a certain treatment is defined as the cost of the treatment divided by its number of QALYs. At the moment when this is being written, kidney transplantations (in the US) are estimated to cost 10.000 USD/QALY, whereas haemodialyses are estimated to cost 40.000 USD/QALY.

9.3 Knowing how, knowing that, and fallibilism in ethics

We can now chart the relationship between, on the one hand, utilitarianism and duty ethics, and, on the other hand, ethical systems that concentrate on evaluating rules or singular acts, respectively. It looks as follows:

	Utilitarianism	*Duty ethics*
Rule evaluating	J. S. Mill	I. Kant
Act evaluating	J. Bentham	W. D. Ross

Table 2: *Four classical moral-philosophical positions (Mill's position is debated).*

In Chapter 5, the distinction between knowing-that and knowing-how was presented from an epistemological point of view. We will now introduce the distinction in the present ethical context. To begin with, we will show its relevance for deontology and consequentialism. We will, when making our comments, move clockwise from Kant to Mill in Table 2.

Kant's categorical imperative and his four commandments lay claims to be instances of knowing-that of basic norms. But having such knowledge, i.e., *knowing-that about how to act* in contradistinction to knowing-that about states of affairs, does not assure that one is able to perform the actions in question. Sometimes there is no problem. Anyone who understands the imperative 'Do not make false promises!' will probably also thereby know how to act on it. But sometimes there can be quite a gap between having knowing-that of a rule and having knowing-how of the application of the same rule. It is easy to understand the rules 'Develop your talents!' and 'Help others!', but in many situations it is hard to know exactly how to act in order to conform to them. In the latter cases, trial-and-error, imitating, and some advice from others might be necessary before one is able to master the rules. Just hearing them, understanding them, and accepting them theoretically is not enough. The tacit dimension of knowledge is as important here as it is in craftsmanship. Its basis is the fact that perception, both in the sense of conscious perception and the

sense of informational uptake, mostly contains more than what a verbal description of the perceived situation contains. And, as we said in Chapter 5, there are four general ways in which know-how can be improved: (1) practicing on one's own, (2) imitation, (3) practicing with a tutor, and (4) creative proficiency.

A famous case of creative proficiency in the moral realm is the story of King Solomon. Two women claimed to be the mother of the same child, and Solomon had to decide which of them should be counted as being the real mother. He threatened to cut the child in two equal parts, and give the women one part each. One of the women then asked him to give the child to the other woman, whereupon Salomon instead gave it to her. He had suddenly realized that the real mother would prefer to save the life of her child more than anything else.

A similar creativity was showed by a surgeon who was refused to operate a young Jehovah Witness; the latter was suffering from an acute spleen rupture and internal bleeding, and needed blood transfusion, which this religion forbids. According to the rules, the surgeon would have to respect the young man's conscious decision not to be operated. Eventually, the patient became increasingly dizzy, and just as he was fading into unconsciousness, the surgeon whispered in his ear: 'Do you feel the wing-beats of death?' The young man's eyes opened wide in fright, and he requested the operation. Here, we might understand the surgeon's intervention as being based on rather paternalistic considerations, and not paying respect to the patient's autonomy. However, if the surgeon was in doubt whether the patient's stated view could be regarded as authentic, he might be seen as examining the patient's true emotional reaction. Even though we might question whether the emotional reaction should override the patient's ordinary views, the case illustrates ethical creativity proficiency.

The young man's eyes opened wide in fright, and he requested the operation. Although the surgeons intervention was based on rather paternalistic considerations – not paying respect to the patient's autonomy – the case illustrates surgical creativity proficiency.

Our remark to Kant's two imperfect duties applies with even more force to Ross' actual duties. Since he has put forward no knowing-that about how to combine different prima facie duties into one actual duty, it is only

by means of know-how that such a feat can be accomplished. And what is true of Ross' act duty ethics is equally true of the utility principle of act utilitarianism. Even though utilitarians can outline a mathematical representation (knowing-that) of what it means to calculate utility, only a god-like being, i.e., a person who is omniscient and has an almost infinite calculating capacity, could ever make a theoretical utility calculation that ends in a simple order: 'Now and here, you should do A!'. Act utilitarians have to make very crude estimations of consequences and utility values, and how to learn to make these in various situations must to a large extent be a matter of learning and improving a utilitarian kind of know-how.

What, then, to say about rule utilitarianism? Does it need know-how in some respect? Answer: yes, and for three reasons. It has one problem in common with act utilitarianism, one with act deontology (Ross), and one with rule deontology (Kant). First, rule utilitarians can no more than act utilitarians derive their duties by mere knowing-that; only people with a rather long experience can do the adequate estimations. Second, as soon as there are conflicts between rules, rule utilitarianism will encounter the same superposition problem as Ross. Third, some of the rules rule utilitarianism normally put forward have the same indeterminate character as Kant's imperfect duties have. In sum, rule utilitarianism is as much in need of knowing-how as the other moral positions mentioned are.

At least since the mid-1980s, many moral philosophers have stressed the need to make a very fine-tuned apprehension of the complexity of a situation before a moral judgment is passed. We see this as an implicit way of stressing the need for knowing-how even in matters of morals. One prominent such philosopher is Martha Nussbaum (b. 1947), who, among other things, claims that moral know-how can be gained by studying ancient literature. However, we will focus on a position whose most outspoken proponent is Jonathan Dancy (b. 1946). He defends what he calls '*moral particularism*' or 'ethics without principles'. According to him, situation determined knowledge can knock down every possible pre-given substantial moral principle such as the utility principle, Kant's commandments, and Ross' prima facie duties. This position, in turn, means that in principle moral knowing-how can always override moral knowing-that.

(Moral particularism in the sense mentioned might be seen as a *general* philosophical defense of what is sometimes called ‘case-based reasoning’ and ‘casuistic ethics’. It is not, however, identical with the Christian ethical theory that is called ‘situation ethics’; the latter allows for some principles.)

We will present the central thesis of moral particularism by means of an analogy with language. According to particularism, moral thinking is basically as little a matter of application of pre-given moral principles to singular cases as language understanding is basically a matter of applying words from a pre-given dictionary and rules from a pre-given grammar. Language and morals existed long before grammarians and law-makers entered the historical scene. Of course, language acts conform to some kind of pattern; therefore, it is always possible ex post facto to abstract word meanings and grammar, which may then be used when teaching a language. But dictionaries and grammar do not determine exactly or forever what sentence to use in a specific situation. Persons who speak a language fluently, and are able to find the proper words even in unusual and extraordinary situations, are not persons of principle; and neither is the morally good person. As dictionaries and grammar are at best crutches for the native speaker, moral principles are at best crutches for the morally sensitive person.

This does not mean that particularists are of the opinion that there are no moral reasons; their claim is that all reasons are context dependent. As a word can mean one thing in one sentence and quite another in another sentence, what constitutes in one case a reason for a certain action, can, it is claimed, in another case be no reason at all; or even be a reason for the opposite action. As the word ‘blade’ means one thing in ‘He sharpened the blade of his sword’ and another in ‘The blade of the plant had a peculiar green hue’, the fact that it knocks on your door is often a reason for you to open the door, but if you have decided to hide at home it is no reason at all; and if it is the big bad wolf that knocks, then the knocking is a reason to lock the door.

The opponents of particularism (i.e., *the generalists* of deontology and consequentialism) might want to object that the mistake is to think that ‘a mere knock’ is ever a reason to open; it is only the more specific

- ‘a knock by a friendly person’,
- or ‘a knock by a friendly person, when you don’t want to be alone’,

that is such a reason. To this the particularists retort: ‘But what if the big bad wolf stands just behind the friend who knocks when I have no special need to be alone?’ The generalists can then try an even more specific description such as:

- ‘a knock by a friendly person, when you don’t want to be alone, and you will not feel threatened when opening the door’.

This does not leave the particularists without an answer. Now they can say: ‘But what if the door has no handle since this was taken away when the door was painted ten minutes ago?’ Let us stop here. The general claim of the particularists is that they can falsify every proposal that a certain fact is – independent of the situation – a reason for a certain kind of action. In other words, particularists claim that there is no pre-given end to specification regresses of the kind exemplified above. When one speaks of reasons, one speaks in effect only of ‘default reasons’. (This exchange of ‘reasons’ for ‘default reasons’ is, by the way, quite parallel to the exchange of ‘causes’ for ‘component causes’ that we propose in Chapter 6.2.)

The particularists claim about specification regresses is, if correct, as devastating for an act duty ethics of Ross’ kind as it is for utilitarianism and classical deontology, since it implies that reasons cannot be added the way forces are added in classical mechanics. Therefore, there are no prima facie duties, i.e., there are no principles that surely *in each and every situation point in the same direction*. To take an example, according to moral particularism, there is neither an actual duty nor a prima facie duty that tells you: ‘Do not give make promises!’ Now, both Ross and Dancy accept that we might encounter situations in which it is morally right to make a false promise. The difference is this: Ross thinks that even in such situations there is a prima facie duty not to make false promises (which is overridden), but Dancy thinks there is not. When the situation is looked at in its very specificity, there is nothing pointing in the direction of not making a false promise; there is not even a rule saying ‘Do not give false promises, except when …!’; to the contrary, there is something in the situation that directly tells you to make a false promise.

Moral particularists believe, like the generalists, that a morally good person must be sensitive to moral reasons. What is at stake is the nature of moral reasons. The particularists thesis is that no verbally explicit moral rule can in its relative abstractness capture all the specificity that is characteristic of real life situations. Therefore, it is impossible to let linguistically formulated moral rules be the absolute end points of moral justifications. This does not imply that a person who wants to act morally is expected to simply gaze vacantly at the situation before him. He should rather look with an experienced eye, i.e., he should meet the situation with some know-how.

A last remark, the particularists' view that no linguistically formulated rule can capture everything that is of relevance for moral judgments is in no conflict with the following completely formal principle for moral consistency:

- all situations that are in morally relevant respects exactly alike should be judged in exactly the same way.

*****

The problem of how to apply the knowing-thats of deontological and consequentialist ethics is not the only problem that this kind of knowing-that has. Apart from *the application problem*, there is *the justification problem*, i.e., the problem of how to justify that the presumed basic norms really can count as valid; a problem that no substantive ethics can side-step. In relation to deontological ethics, we have already noted that it is hard to accept that God's commandments or Kant's categorical imperative justifies themselves. Now we will make some remarks on the same problem in relation to utilitarianism and the utility principle. Here comes a 'mnemonic doggerel' from Bentham:

> *Intense*, *long*, *certain*, *speedy*, *fruitful*, *pure*—
> Such marks in *pleasures* and in *pains* endure.
> Such pleasures seek if *private* be thy end:
> If it be *public*, wide let them *extend*

> Such *pains* avoid, whichever be thy view:
> If pains *must* come, let them *extend* to few.

The first three lines state a rather uncontroversial thesis. If your actions do not affect anyone at all apart from yourself, then feel free to seek your private pleasure and try to avoid pain. Only hardheaded ascetics who think there is something *intrinsically* wrong with pleasure can object; we will, just as Bentham, leave them out of account. The next three lines, however, are in much more need of justification. What makes Bentham able to move so swiftly from the first part of the verse, where only one's own pleasure seeking is at stake, to the second part, where one is also encouraged to help others to have pleasure? To the person who experiences them, pleasures come stamped as being in some sense positive, and pains as being negative. Therefore, if nothing else intervenes (e.g., deontological norms), the rule that I should seek pleasure and avoid pain for myself justifies itself. But why should a person care for the pleasures and the pains that he himself does not experience? In situations of compassion and empathy, one does in some sense experience even the sufferings of other people, but often the pleasures and pains of others are apprehended in a rather neutral way. Why care about them when this is the case? The utility principle puts one's own and all others' pleasures and pains on a par, but this is not the way they usually appear to us. A good argument in favor of the equalization is needed, but, a bit astonishingly, Bentham has none; and, even more astonishingly, neither has Mill, who in an oft-criticized sentence says:

> The sole evidence it is possible to produce that anything is desirable [= worthy of desire], is that people do actually desire it. If the end which the utilitarian doctrine proposes to itself were not, in theory and in practice, acknowledged to be an end, nothing could ever convince any person that it was so (*Utilitarianism*, Ch. 4).

But "the end which the utilitarian doctrine proposes" is to maximize the total amount of pleasure, and this is definitely not what most people "actually desire." Mill's move from each man's maximizing of his own

pleasure (egoistic hedonism) to each man's conformance to the utility principle (universalistic hedonism) is as logically unfounded as Kant's move from 'being capable of creating ends for oneself' to 'being an end in itself'. Mill can be accused of making two fallacies. First, he writes as if being 'subjectively desired' implies being 'objectively desirable' in the sense of worthy of desire. Second, he writes as if the fact that each man's happiness is objectively desirable should entail that it is always, even in cases of conflict with one's own happiness, objectively desirable to maximize the total amount of happiness. But what logically follows is only a trivial thing, namely that there is a kind of happiness of mankind that *can be defined* as the aggregated happiness of each person.

Justification is an enterprise not only in ethics. Let us illustrate the justificatory problem of utilitarianism by a detour back to Chapter 3.4 and the logical positivists' principle of verification. They claimed that only verifications (i.e., positive empirical observations) could justify one in regarding a non-formal scientific assertion as being true. But how then is the verification principle itself to be justified? Since it is meant to be both non-formal and justified, *it should be applicable to itself*. But this seems ridiculous. It would mean that we should regard the verification principle as true if and only if we can verify that it is true; but then we have nonetheless presupposed it, and there is no use in testing it. Similarly, if the utility principle is regarded as justified, it should be applicable to itself. But this seems ridiculous, too. It would mean that we should regard the utility principle as a basic norm if and only if we can show that following it would lead to a maximum of pleasure; but then we have nonetheless presupposed it, and there is no use in verifying it.

Within positivism, the verification principle is given no real justification, and neither is, normally, the utility principle within utilitarianism. The famous classical utilitarian Henry Sidgwick, however, has proposed a justificatory reform of utilitarianism. He claims that it is as possible in morals as in mathematics to have true intuitive non-empirical insights, and that it is by means of such an insight that the utility principle becomes justified. But to resort to intuitions has not ranked high in mainstream twentieth century philosophy, and Sidgwick did not have many followers. The justificatory problem of ethics is one reason why, as we mentioned at the beginning of this chapter, many twentieth century moral philosophers

have chosen to work only with presumed morally neutral meta-ethical problems. A new approach to the justificatory problems in moral and political philosophy was inaugurated by John Rawls (1921-2002) in his book *A Theory of Justice* (1971); here the notion of 'reflective equilibrium' is central.

Rawls does not use the term 'fallibilism', but his move does implicitly introduce fallibilism in ethical matters. He asks us to stop looking for self-justificatory moral principles and/or self-evident particular moral judgments. Instead, he claims, all we can reasonably strive for is that our considered principles and considered particular judgments cohere with each other. That is, they ought to balance each other or to be in *reflective equilibrium*. If they are not, then we have to change something. Let us take a simple example. Assume that someone who believed and defended for a long time the norm that it is absolutely forbidden to actively hasten a patient's death, ends up in a situation where he, after reflection, comes to the conclusion that he is – in this very special situation – allowed to do so in relation to an unbearably suffering patient's death. The patient intensely begs the physician to help him to die quickly, since all palliative treatments have failed. Then there is *dis*equilibrium between the physician's moral principle and his particular judgment, and if he is a reflective person he ought to reject or revise his principle, his particular judgment, or both in order to restore equilibrium.

Initially, in situations like the one above, we do not know whether to change a principle (or several), a particular judgment (or several), or perhaps both, which means that many conjectures may have to be tested before one can rest content and say that a new equilibrium has been found. Fallibilism enters the scene because no such equilibrium can be regarded as certainly stable. One day there may arise a completely un-thought of kind of situation, which makes us become very certain that, here, a particular moral judgment of ours have to contradict one of our earlier accepted principles. There is then disequilibrium. Our moral principle, or perhaps even our whole moral paradigm, contains an anomaly in much the same way as a scientific paradigm can contains anomalies caused by observations or measurements (see Chapter 3). Another day, a moral reformer (think, for instance, of Singer) may present a principle that so far we have not considered at all, and which contradicts many of our old

particular judgments. Then there would again be disequilibrium. Sometimes a whole new moral paradigm is proposed, e.g., utilitarianism at the end of the eighteenth century.

The method of reflective equilibrium consists in working back and forth among our considered judgments on both moral principles and particular moral judgments; revising any of these elements whenever necessary in order to achieve an acceptable coherence. Such a coherence means more than that our beliefs are merely logically consistent with each other. Our principles should be relevant for many of our particular judgments, and these judgments should be regarded as a kind of evidence for the principles. On Rawls' view of justification, one reflective equilibrium can be seen as being better than an older and superseded one, but nonetheless it can itself very well in the future be replaced by an even better reflective equilibrium. Even the 'principle-free' view of moral particularism can be evaluated in this way. When particularism is in reflective equilibrium, there is coherence between the view that there are no context independent moral reasons and all the considered particular judgements. Dancy is explicitly fallibilist and writes: "This [particularist] method is not infallible, I know; but then neither was the appeal to principle (Dancy 2005)."

Rawls' views on reflective equilibria are as possible to apply to rules and principles in the medical domain as to general moral principles. All those who work as health care professionals – doctors, nurses, and the rest of the staff – must try to find a reflective equilibrium between the rules of health care ethics and everyday medical practice.

After about 200 years of philosophical discussion of deontological ethics and utilitarian consequentialism, it is easy to summarize what is regarded as their main kind of anomalies, i.e., what kind of considered particular moral judgments each seems to be in conflict with. The Decalogue commandment 'You shall not kill!' is for many people contradicted by their particular judgments when in danger of being murdered or in case where someone begs for euthanasia. Kant's commandment 'Do not commit suicide!' is to many contradicted by what seems morally allowable when one is threatened by lifelong intense pains and sufferings. His commandment 'Do not make false promises!' seems in several situations not to be exactly fitting. Put more generally, what is wrong with

deontology is its inflexibility. And the same goes then of course for rule utilitarianism. Neither allows norms to take exceptions. Sometimes it does not seem morally wrong to break a rule, even though it would cause disaster if most people broke the rule most of the time.

Many find the utility principle of act utilitarianism contradicted by situations where they have to suffer for the happiness of all, even though they have done nothing wrong. Is it right, for instance, to force a person to donate one of his kidneys to another man if this would increase the total amount of pleasure? Examples of this kind can easily be multiplied. Also, when it comes to questions of justice, act utilitarianism is counter-intuitive. All act utilitarianism can say is that one is being treated justly if one is being treated as a means for the happiness of all. Since justice is normally regarded as justice towards (or between) persons, this view contains a complete redefinition of justice. It seems impossible for act utilitarians (but not for rule utilitarians) to bring personhood into their moral system. In short:

- Classical deontological ethics and rule utilitarianism cannot make sense of our considered particular judgments to the effect that, *sometimes*, we have to make exceptions to norms.
- Act utilitarianism cannot make sense of our judgments that, *in many situations*, we have some kind of moral rights as individual persons; rights which no utility principle can overrule.

Both these blocks of anomalies are highly relevant for medical ethics. Most physicians do now and then encounter situations that are quite exceptional from a moral point of view, and, as we will see in Chapter 10, many medical norms put the integrity of individual persons in the center. Normal patients shall be regarded as autonomous agents, and clinical research requires informed consent. But the utility principle does not automatically exclude even the killing of an innocent person. Here comes a common anti-utilitarian thought-experiment.

A patient suffering from injuries and several fractures arrives at an emergency room. With adequate treatment he is curable and able to be fully rehabilitated. This patient, however, happens to have the same tissue type as some other patients who are waiting for heart- and lung-transplants

as well as kidney transplants. On the assumption that it is for sure possible to keep it a secret, it would according to a utility calculation be right to take the organs from the injured patient and give them to the other patients. This would maximize happiness, but (rhetorical question) isn't it nonetheless wrong? It has to be added, however, that most act utilitarians are in fact against such interventions, since they regard it impossible to keep this a secret; accordingly trust in the health care system will erode and the overall negative consequences will become more significant than the positive.

The analogy between fallibilism in science and fallibillism in morals, which we noted in passing, contains yet another feature that it is good to be aware of. Neither in science nor in ethics is it possible to completely foresee the future. On the ruins of an old paradigm someone may one day be able to construct a wholly new and unforeseen version. Therefore, we are not trying to say that it is logically impossible for deontology and utilitarianism to recover from their present anomalies and stage a come back. Furthermore, wholly new moral paradigms might be created. Therefore, we are not claiming that it is logically impossible to create a moral paradigm that is neither a deontology, nor a consequentialism, and nor a virtue ethics.

*****

We would like to end this subchapter with some words on a minor school of thought called 'personalism'. It does not immediately fit into any of the three moral paradigms we are presenting. Mainly because it focuses on questions about *what has value and how to rank different kinds of values,* and does not bother to work out what the norms ought to look like. Like Kant, the personalists stress that persons have an absolute value, but unlike Kant they claim that to be a person means much more than to have a mind that has self-consciousness and can reason. All of them claim that to be a person essentially involves having an emotional life; some personalists also claim that personhood is necessarily social, i.e., that one cannot be a person without to some extent also caring for other persons. The basic personal relation is claimed to be, with an expression from Martin Buber (1878-1965), an 'I-Thou relation'. Such personalism contradicts

individualism, since it claims that true self-love involves love of others. Since many of the famous personalists have been religious thinkers, it should also be said that many personalists think there is an ‘I-Thou relation’ even between human beings and a higher being. Personalism has played quite a role in some discussions within nursing science and nursing ethics.

The most famous philosopher who has thought along personalist lines is probably the German Max Scheler (1874-1928). He claims that there are four distinct classes of values (and disvalues) that can be ranked bottom up as follows: (1) sensory values, i.e., sensory pleasures in a broad sense; (2) vital values, i.e., things such as health, well-being, and courage; (3) spiritual values, e.g., justice, truth, and beauty; (4) holy values such as being sacred. Scheler argues (but other personalists contest it) that in cases of conflict one has always to act so as to realize the higher value.

9.4 Virtue ethics

A virtue is a habit and disposition to act and feel in a morally good or excellent (virtuous) way. Etymologically, it might be noted, the term is quite sexist. It comes from the Latin ‘vir’, which means ‘man’ in the masculine sense, i.e., the virtuous man is a thoroughly manly man. The corresponding ancient Greek word, ‘aretē’, does not, however, have such an association. It means skill or excellence in a non-gendered sense.

In order to see what is specific to virtue ethics, one has to see what deontology and consequentialism have in common. As before, we will use Kantian ethics and utilitarian ethics as our examples. The contrast between these ethical views and virtue ethics contains some intertwined features. Whereas the central principle in both Kantianism and utilitarianism tells what a morally right *action* looks like (Kant’s categorical imperative and the utility principle, respectively), the central principle of virtue ethics tells us what kind of *man* we ought to be. It simply says:

- Be a morally virtuous man!

Virtue ethics was already part of Chinese Confucianism (Confucius 559-479BC), and it was the dominant ethics in Ancient Greece. Its

philosophical father in the Western world is Aristotle. During the nineteenth and the twentieth century, its influence in Western philosophy was low, but there were nonetheless critics of deontology and utilitarianism. Some famous late twentieth century Anglo-American virtue ethicists are Philippa Foot (b. 1920), Bernard Williams (1929-2003), and Alasdair MacIntyre (b. 1929).

The difference stated between virtue ethics and deontology/consequentialism does not mean that only virtue ethics can talk about virtuous persons. In Kantian ethics, a virtuous person can be defined as someone who always wants to act in conformity with Kant's categorical imperative; and in utilitarianism, a virtuous person can be defined as someone who always tries to act in conformity with the utility principle. Conversely, neither does the difference stated mean that virtue ethics cannot at all formulate an act-focusing principle. It becomes, however, a rather empty tautology that does not distinguish virtue ethics from deontology and consequentialism:

- Act in such a way that your action becomes morally right.

This virtue ethical imperative can neither fulfill the testing function that the first formulation of Kant's categorical imperative fulfills, nor can it have the regulative function that the utility principle has. Nonetheless, stating it makes a sometimes misunderstood feature of virtue ethics clear. A virtuous person is someone who out of character or habit chooses the morally right action, but his actions do not become the morally right ones *because* he is a virtuous person. It is the other way round. A person is virtuous because he performs the morally right actions. Compare craftsmanship. A good physician is good because he can cure people; the treatments do not cure because they are prescribed by a good physician. No virtuous person can stamp actions as being virtuous the way a baptizing priest can give a child a certain name. Aristotle writes:

> The agent [the virtuous man] also must be in a certain condition when he makes them [the virtuous actions]; in the first place he must have knowledge, secondly he must choose the acts, and

> choose them *for their own sakes* [italics inserted], and thirdly his action must proceed from a firm and unchangeable character (*Nicomachean Ethics* 1105a32-36).

For reasons of practical expediency, we must often take it for granted that virtuous-regarded persons have acted in the morally right way, but this does not mean that such persons act in the right way by definition. Another way to put this point is to say that for a virtuous person the primary target is to perform virtuous acts, but in order to become able to hit the target continuously, he should try to become a virtuous person. The fact that experiments in social psychology have shown that character traits are more situation dependent as was once widely thought, does not mean that they are completely insignificant. Rather, it means that virtuous persons should try to take even this situation dependency into consideration.

The misunderstanding that a virtuous person can *define*, not only find out and perform what is to be reckoned as the morally right action, may have many sources. Here is one possible: it is noted that a virtuous person often in everyday life functions a bit like a judge in a law court, but it is forgotten that juridical decisions can always be questioned. Let us expand. Court judges are meant to apply pre-written laws, but since the written law is often not detailed enough to take care of the cases under scrutiny, the judges (exercising their knowing-how) have often so to speak to define where the boundary between legal and illegal actions should be drawn. But his is not the absolute end of the procedure. If the court in question is not the Supreme Court, then the verdict can be appealed; and even if it should be the highest instance, there is nothing logically odd in thinking that the verdict is wrong.

In the last subchapter we described how know-how enters Kantianism and utilitarianism. In virtue ethics know-how becomes much more prominent because of the vacuity of its basic knowing-that, the norm: Act in such a way that your action becomes the morally right action! Kant’s categorical imperative, his four commandments, and all forms of the utility principle convey some substantial knowing-that, but the norm now stated does not.

We have earlier described the difference between hypothetical and categorical imperatives or rules. Know-how can exist in relation to both,

but now we are talking only about know-how in relation to categorical imperatives. In Chapter 5, on the other hand, we were concerned only with know-how in relation to goals that one is free to seek or not to seek. This distinction between knowing-how in relation to hypothetical and categorical imperatives can be found already in Aristotle. He calls the former kind of know-how 'techne' and the latter one 'phronesis'. In modern literature, they are often called 'technical skill' and 'practical wisdom', respectively. Technical skill is know-how in relation only to the choice of means, whereas practical wisdom is know-how in relation to questions of what one has categorically to do. (Theoretical wisdom, 'sophia', is simply having infallible knowing-that, 'epistéme'.)

If moral particularism (see the last subchapter) is right, then a further distinction is needed. There are two kinds of phronesis (moral knowing-how); we will call them '*principle-complementary*' and '*principle-contesting*' phronesis. The former kind is needed when a pre-given, but vague, moral principle (moral knowing-that) shall be applied; it is needed both in order to know if the principle is at all applicable, and in order to know how it should be applied. Such know-how can by definition never contest the principle, and this is the only kind of phronesis that deontological and consequentialist ethical systems require. The other kind of phronesis is the one required by particularism, i.e., a moral know-how that can contest every moral knowing-that. Virtue ethics contains both kinds of phronesis. A practically wise man can both apply rules in extraordinary situations and take account of particular situations where all existing moral principles have to be overridden. In this undertaking, it should be noted, consequences may sometimes have to be considered and sometimes not. Moral principles are neither necessary nor sufficient for becoming a virtuous person, but experience is necessary. In an oft-quoted passage Aristotle says:

> while young men become geometricians and mathematicians and wise in matters like these, it is thought that a young man of practical wisdom cannot be found. The cause is that such wisdom is concerned not only with universals but with particulars, which become familiar from experience, but a young man has no

> experience, for it is length of time that gives experience (*Nicomachean Ethics* 1142 a).

The fact that virtue ethics contains the principle-contesting kind of phronesis, does not imply that virtue ethicists find all substantive moral rules superfluous. It means only that they regard all such rules as rules of thumb, as default rules, or as rules for normal circumstances; as such, by the way, the rules can be in need of principle-complementary phronesis. As computers need default positions in order to function in general, human beings seem to need default norms in order to function in society. The question whether moral rules are useful in education and in practical moral thinking is distinct from the question whether moral justification ends with verbally stated substantial universal principles. It is only principles of the latter kind that the particularists' and the virtue ethicists oppose.

We have remarked that as soon as act utilitarians and act duty ethicists enter politics or institutional acting, even they have to accept discussions of rules and make room for some kind of rule utilitarianism and rule duty ethics, respectively. Exactly the same remark applies to virtue ethics. Virtue ethics cannot rest content with merely stressing the particularist aspect of individual acts; it has to say something about rule creation, too. As a person may ask what action he has categorically to do, a law-maker may ask what rules he has categorically to turn into laws. There are two kinds of phronesis: *act phronesis* and *rule phronesis*, respectively. At least modern societies require phronesis two times, first when laws and rules are made, and then when the laws and the rules are applied. The societies of Confucius and Aristotle did not contain as strict a division of labor between law-makers and law-appliers as our modern societies do. Rule phronesis is a virtue of politicians and legislators, and act phronesis is a virtue of judges, policemen, bureaucrats, and rule followers of all kinds.

For Aristotle, the rules followed in personal acting are not rules directly showing how to act, but rules for what kind of personal character traits one ought to develop in order to be a virtuous man. The character traits in question are connected both to certain ways of feeling and to certain ways of acting. Often, a virtuous character trait is contrasted with two contrary opposite and vicious character traits; one of these bad character traits is seen as expressing itself in an excess and the other in a deficiency of a

certain kind of action or feeling. The virtuous man is able to find the golden mean between two bad extremes. The classical example is the virtuous soldier. He is brave, which means that he has found a golden mean between cowardice (too much fear) and foolhardiness (too little fear). But what it means to be brave in a certain situation cannot be told beforehand, i.e., no one can be brave without know-how. Here comes an Aristotelian list, even though it lays no claims to be completely true to Aristotle's (sometimes hard-translated) examples in *Nicomachean Ethics*; our list is merely meant to convey Aristotle's general approach in a pedagogic way.

Vice (defect)	**Virtue (mean)**	**Vice (excess)**
Foolhardiness (too little fear)	Courage	Cowardice (too much fear)
Insensibility (caring too little about pleasure)	Temperance	Self-indulgence (caring too much about pleasure)
Stinginess (giving too little)	Generosity	Prodigality (giving too much)
Shamelessness (too little shame)	Modesty	Bashfulness (too much shame)
Humility (too little integrity)	Pride	Vanity (too much integrity)
Apathy (too little emotion)	Good Temper	Irascibility (too much emotion)
Surliness (being too negative towards others)	Friendliness	Flattery (being too positive towards others)

These rules of conduct ('Be courageous!', etc.) have to be regarded as rules of conduct for normal situations in one's own community. One should not urge enemies to be courageous, and one should not be generous and friendly towards enemies. And even in one's own community there may be situations where it is adequate to lose one's good temper and become angry. Aristotle is quite explicit on this.

No virtuous mean does in itself take any degrees, only the deviations do. The virtuous mean is in this sense like the zero point on the scale for

electric charges. There are degrees of both negative and positive electric charge, but there are no degrees of not being electrically charged.

Hippocrates thought in the same way about doctors. They ought to have the general virtues and then some for their profession specific virtues; one might say that Hippocrates put forward an applied virtue ethics. Especially, he warned against unbridled behavior, vulgarity, extortion, and shamelessness, as well as being insatiable. He even put forward rules for how doctors ought to dress. According to the Hippocratic recommendations, doctors should be moderate when it comes to financial matters, and they should neither be greedy nor extravagant. Moderation is the golden mean between these two extremes. There is quite a structural similarity between his balance thinking about diseases and his balance thinking about morally good behavior. Here comes a Hippocratic list of vicious and virtuous ways of being.

Extreme	**Golden Mean**	**Extreme**
Ignorant	Modest	Pretentious
Tempting	Friendly	Coquettish
Weak	Robust	Domineering
Inferior	Humble	Pompous
Nonchalant	Devoted	Fanatic
Slow-witted	Self-command	Impulsive
Soiled (corrupt)	Decent	Artful (too cunning)
Ignorant	Careful	Finical (too keen on details)
Cynical	Empathic	Hypersensitive

The relationship between moral rules and moral know-how in virtue ethics can be further clarified by reflections on how moral knowledge can develop. If there is a deontological norm such as ‘Do not make false promises!’, then there is, as Kant explicitly noted, three *logically different* ways in which one can conform to it. One can:

- act *in accordance with it*, i.e., act without caring about the rule, but nonetheless happen to act in such a way that one conforms to it
- act *on it*, i.e., know the rule and consciously conform to it, but doing this for reasons that are morally indifferent

- act *for it*, i.e., know the rule and consciously conform to it because one thinks that this is the morally right thing to do.

According to Kant, it is only persons that *act for* the categorical imperative that act morally, and can be regarded as being good (virtuous) persons. In Aristotle one meets Kant's tripartite logical distinctions as a distinction between three necessarily consecutive stages in the moral development of human beings. Small children can only be taught to act *in accordance with* good moral rules by means of rewards and punishments. Today, being more aware of the imitating desires and capacities of children, we might add that even small children can learn to act in accordance with moral standards by imitating good role models. They need not be educated the way dogs are trained.

When children have become capable of understanding rules and of consciously acting *on* such, they enter the second stage. They should then, says Aristotle, be taught to heed morally good rules 'as they heed their father'. Still, they lack moral insight, but they can nonetheless regard it as natural and necessary to follow the rules in question; normally, they even develop some technical skill in doing so.

When entering the third stage, two things happen. First, the technical skill becomes perfected, which means that the rules are left behind in the way stressed by particularists. Second, the skill becomes transformed into phronesis. That is, persons now have an insight that they categorically ought to do what they are doing because this is morally good. The first (rule-trespassing) change makes Aristotle differ from Kant, but the second (insight-creating) change parallels Kant's move from 'acting on' rules to 'acting for' moral rules.

Aristotle's developmental thinking has counterparts in contemporary philosophy and psychology. In Chapter 5.4, we presented the five stages of skill acquisition that the Dreyfus brothers have distinguished: the novice stage, the advanced beginner stage, the competence stage, the proficiency stage, and the expertise stage. They think that these distinctions apply to moral skill and moral maturity too. Since Aristotle's stages allow for having grey zones in-between them as well as for having sub-stages, one can try to relate the Dreyfus' stages to Aristotle's. We would then like to make the novice and the advanced beginner stages sub-stages of Aristotle's

second stage, for since the novices are instructed by means of rules, Aristotle's first stage is already left behind. The competence stage seems to be an intermediary between Aristotle's second and third stage, whereas the proficiency and the expertise stages seem to be sub-stages of the third stage.

At the beginning of the 1970s, a developmental psychologist, Lawrence Kohlberg (1927-1987), claimed to have found empirically a law of moral development. It gave rise to much discussion. Like Aristotle, Kohlberg found three main levels, but unlike Aristotle he divides each of them into two stages, and the last level is in one respect not at all Aristotelian. Kohlberg's main views are summarized in Table 3 below. Level 1 is called 'the pre-conventional level', level 2 'the conventional level', and level 3 'the post-conventional level'. At the pre-conventional level we (human beings) have no conception of either informal moral rules or formal rule-followings. At the conventional level we have, but we simply take the rules for granted. Finally, at the post-conventional level, we have arrived at the insight that there are morals, and at the last sub-level we have even realized that morals can be discussed.

Stages of moral consciousness:	Main idea of the good and just life:	Main kinds of sanctions:
(1a) Punishment-obedience orientation	Maximization of pleasure through obedience	Punishment (as deprivation of physical rewards)
(1b) Instrumental hedonist orientation	Maximization of pleasure through tit-for-tat behavior	Punishment
(2a) Good-boy and nice-girl orientation	Concrete morality of gratifying interaction	Shame (withdrawal of love and social recognition)
(2b) Law-and-order orientation	Concrete morality of a customary system of norms	Shame
(3a) Social contract orientation	Civil liberty and public welfare	Guilt (reaction of conscience)
(3b) Universal ethical principles orientation	Moral freedom	Guilt

Table 3: *Kohlberg's three levels and six stages of moral development.*

In relation to this table, Kohlberg says something like the following. Children from 0-9 years live on the pre-conventional level where, first (1a), they only strive for immediate hedonistic satisfaction, but later (1b) learn that by means of tit-for-tat behavior one can increase one's amount of pleasure. They will act in accordance with morals either by chance or because adults reward them for moral behavior and punish them for immoral behavior; punishment is here taken in such a broad sense that a mere negative attitude counts as punishment.

Children and teenagers in the age of 9-20 are normally living on the conventional level. They do now perceive that there is something called 'being good' and 'being bad', respectively. At first (2a), they only apprehend it in an informal way when they concretely interact with other people, but later (2b) they can connect 'being good' and 'being bad' to the conformance of impersonal rules. Something in human nature makes people ashamed when they regard themselves as having done something bad. The living on this level is deontological in the sense that role conformance and rule following are regarded as being important quite independently of their consequences.

Some people may forever stay on the conventional level, but many adults proceed to the post-conventional level. First (3a) they realize that at bottom of the informal and formal rules that they have earlier conformed to, there is an implicit or explicit social contract between people. Then, perhaps, they also come to the conclusion (3b) that good social contracts ought to be based on universal ethical principles. Here, Kohlberg is much more a Kantian and/or utilitarian than an Aristotelian. On level 3, people react with a feeling of guilt if they think that they have done something that is seriously morally wrong.

(It has been argued that 'ontogeny recapitulates phylogeny', i.e. that the biological development of an organism mirrors the evolutionary development of the species. Similarly, Kohlberg tends towards the view that the moral development of the individual recapitulates the historical moral development of societies. He says: "My finding that our two highest stages are absent in preliterate or semiliterate village culture, and other evidence, also suggests a mild doctrine of social evolutionism (Kohlberg 1981 p. 128).")

Leaving societal moral development aside, what to say about Kohlberg's schema? Well, our aim here is mainly to show that questions about moral development emerge naturally, and that there are reasons to think that discussions of them will continue for quite a time. We will only add three brief remarks and then make a historical point.

First, does the schema contain as many levels and stages that it ought to contain? From an empirical point of view, Kohlberg later came to doubt that one should speak of stage (3b) as a *general* stage of development; not many people reach it. From a moral-philosophical point of view, however, he has speculated about the need for a seventh level, and from such a perspective Habermas has proposed within level 3 a more dialogical stage (3c), which fits discourse ethics better as an end point than (3b) does. One might also ask whether all kinds of emotional sanctions have been taken into account. Is there only shame and guilt? Some philosophers have argued that even remorse is a kind of 'morally punishing' emotion – and a better one.

Second, what about the last stage? It seems to make no difference between deontology and consequentialism; it only focuses on what these have in common, namely a stress on the existence of universal ethical principles. However, this stage differs from both old-fashioned deontology and consequentialism in bringing in fallibilism; on stage (3b) people are prepared to discuss their own moral norms.

Third, what about the place afforded to virtue ethics? It seems to have no place on any stage above (2a). The stress on know-how typical of virtue ethics matches Kohlberg's stress on the informal aspect of the moral interaction that characterizes this stage. But then there is only explicit rule-following and/or discussions of such rules. The Dreyfus brothers have argued that the kind of phronesis that virtue ethics regards as central has to be made part of the description of the last stage.

Kohlberg's first empirical investigations gave rise to a very intense debate about gender and moral development. According to his first results, on average, girls score lower on moral development than boys do. One of Kohlberg's students, Carol Gilligan (b. 1936), then argued that this was due to the fact that Kohlberg, unthinkingly, favored a principle-seeking way of reasoning to the detriment of a relation-seeking way. Both are necessary, both can be more or less developed, and none of them can be

given a context independent moral priority over the other. The first way is the primary one in relation to questions of justice, and the second way in relation to moral problems in caring. According to Gilligan, given traditional sex roles, faced with moral problems boys focus on principles and justice, whereas girls focus on face-to-face relations and caring. If both kinds of moral issues are given their moral-developmental due, then, Gilligan argues, there is no longer any sex difference with respect to moral development. Her writings became one seminal source of what is now known as 'the ethics of care'. In this school of thought one stresses (and investigates) the moral importance of responding to other persons as particular individuals with characteristic features.

In our exposition of virtue ethics, we have so far made a distinction between Aristotle and modern virtue ethicist thinking in only one respect: the latter needs a notion of 'rule phronesis'. But we think there are three more respects in which modern virtue ethics has to differ from the traditional one. First, fallibilism has to be integrated. Aristotle was an infallibilist, but as we have seen even expert know-how can miss its target (Chapter 5.4-5). Even if the expert in front of a contesting non-expert sometimes has to say 'I cannot in detail tell you why, but this is simply the best way to act!', it may turn out that the non-expert was more right than the expert.

Second, according to thinkers such as Socrates, Plato, and (but to a lesser extent) Aristotle, there cannot arise any real conflicts between acting morally right and acting in the light of the happiness of one's true self. Such apparent conflicts arise, they claim, only when a person lacks knowledge about what is truly good for him. Socrates thought it was best for his true self to drink the hemlock. But this view is hard to defend. The dream of what might be called an 'unconflicted psychology' has to be given up even among virtue ethicists. When in their original culture, an elderly Eskimo was asked by his son to build an igloo in which he could 'travel on his own', he was expected to be aware that he had become a burden to his tribe, and follow its duties and sacrifice himself. According to classical virtue ethics, the old Eskimo might be said to act also in his own individual self-interest, but there is no good reason to cling to this view. Virtue ethicists have to accept that conflicts can arise between what their moral know-how tells them to do, and what their true self-interest

wants them to do. The existence of such conflicts is for duty ethicists and utilitarians such a trivial and elementary fact that for many of them its denial immediately disqualifies classical virtue ethics for consideration. Even though a deadly sick elderly utilitarian patient, who wants to live longer, may for reasons of maximizing happiness come to the conclusion that younger patients with the same disease should have his place in the operation queue, he would never dream of saying that his choice furthers his own self-interest.

Third, we have to repeat that experiments in modern social psychology have shown that character traits are not as situation independent as classical virtue ethicists assumed.

In what follows, we will take it for granted that modern virtue ethics, just as duty ethics and utilitarian ethics, accepts the existence of conflicts between acting morally right and acting in the light of one's true self. This means that all three have a motivational problem:

- what can in a situation of conflict make a man act morally instead of only trying to satisfy his self-interest?

We will not present and discuss any proposed solutions to this problem, only put virtue ethics on a par with deontology and consequentialism. The latter might argue that morality has the sort of authority over us that only a rule can provide, but virtue ethicists can then insist that situations and individuals can come in such a resonance with each other that, so to speak, a moral demand arises from the world. The situation simply demands that self-interest has to surrender.

Our little plea for a modern fallibilist virtue ethics has structural similarities with our earlier strong plea for a fallibilist epistemology. Before we bring it out, let us here a second time (cf. p. 00) quote Thomas Nagel about the inevitability of trying to find truths and trying to act right:

> Once we enter the world for our temporary stay in it, there is no alternative but to try to decide what to believe and how to live, and the only way to do that is by trying to decide what is the case and what is right. Even if we distance ourselves from some of our thoughts and impulses, and regard them from the outside, the

> process of trying to place ourselves in the world leads eventually to thoughts that we cannot think of as merely "ours." If we think at all, we must think of ourselves, individually and collectively, as submitting to the order of reasons rather than creating it (Nagel 1997, p. 143).

When we focused on truth-seeking, we linked this quotation to Peirce's view that we should "trust rather to the multitude and variety of [the] arguments than to the conclusiveness of any one[; our] reasoning should not form a chain which is no stronger than its weakest link, but a cable whose fibers may be ever so slender, provided they are sufficiently numerous and intimately connected." Having now focused on right-act-seeking and introduced the notion of phronesis, we can say that Peirce's view is a way of stressing the importance of phronesis in epistemology. The methodological rules taught in a science should be looked upon as default rules.

9.5 Abortion in the light of different ethical systems

We have presented the three main ethical paradigms and some of their sub-paradigms, and we have shown by means of the notion of 'reflective equilibrium' how discussions about what paradigm to accept can make cognitive sense. Next we will present what one and the same problem may look like from within the various paradigms. Our example will be one of the big issues of the twentieth century, abortion. We leave it to the reader to find structural similarities between this and other cases. Some aspects of the problem of abortion will probably, because of the rapid development of medical technology, soon also appear in other areas.

The problem we shall discuss is how to find a definite *rule* that speaks for or against abortion in general or at some date. In relation to a single case of a woman who wants a forbidden abortion, a prima facie duty ethicist may then nonetheless come to the conclusion that the rule is overridden by other prima facie duties, an act utilitarian that it is overridden by some act utilitarian considerations, and a virtue ethicist that his practical wisdom requires that he makes an exception to the rule. But this is beside the rule discussion below.

In all deontological ethical systems put forward so far, the rule for abortion (pro or against) is derived from some more basic norm. Let us start with a religious duty ethicist who believes in the sanctity of human life and regards the norm 'You shall not kill human beings!' as being absolute, and one from which he claims to be able to derive the rule that abortions are prohibited. From a pure knowing-that point of view he has no problem, but from an application point of view he has. The application cannot be allowed to be a matter of convention, because deontological norms should be *found*, not created. That is, the norm presupposes that there is in nature a discontinuity between being human and not being human. By definition, where in nature there are only continuities, every discontinuity must be man-made and in this sense conventional. To take an example: the line between orange and yellow colors is conventional, and such a kind of line cannot be the base of a deontological norm. So, what does the first part of the developmental spectrum for human beings look like? Where in the development '(egg + sperm) → zygote → embryo → fetus → child' is there a discontinuity to be found between life and non-life? Here is a modern summary of prototypical stages and possible discontinuities (Smith and Brogaard 2003):

a. the stage of the fertilized egg, i.e., the single-cell zygote, which contains the DNA from both the parents (day 0)
b. the stage of the multi-cell zygote (days 0-3)
c. the stage of the morula; each of the cells still has the potential to become a human being (day 3)
d. the stage of the early blastula; inner cells (from which the embryo and some extraembryonic tissue will come) are distinguished from outer cells (from which the placenta will come) (day 4)
e. implantation (nidation); the blastula attaches to the wall of the uterus, and the connection between the mother and the embryo begins to form (days 6-13)
f. gastrulation; the embryo becomes distinct from the extraembryonic tissue, which means that from now on twinning is impossible (days 14-16)
g. onset of neurulation; neural tissue is created (from day 16)
h. formation of the brain stem (days 40-43)

i. end of the first trimester (day 98)
j. viability; can survive outside the uterus (around day 130 [should be 147])
k. sentience; capacity for sensation and feeling (around day 140)
l. quickening; the first kicks of the fetus (around day 150)
m. birth (day 266)
n. the development of self-consciousness

First some general comments on the list. Discontinuity i (the end of the first trimester) is obviously conventional; and discontinuity j (viability) is conventional in the sense that it depends on medical technology. Several medieval theologians discussed an event not mentioned in the list above, the date for 'ensoulment' of the fetus, i.e., the day (assumed to be different for boys and girls) at which the soul entered the body. Some ancient philosophers, including Plato, thought that it was not until the child had proven capable of normal surviving that it could be regarded as a human being. This view might have influenced the Christian tradition of not baptizing a child until it is six months old. Children who managed to survive the first six months of their lives were at that time assumed to have good chances of becoming adults. Aristotle took quickening to be the decisive thing; and so did once upon a time the British Empire. British common law allowed abortions to be performed before, but not after, quickening.

The Catholic Church takes it to be quite *possible* that the fertilized egg is a living human being, and that, therefore, abortion might well be murder. The coming into being of the zygote marks a real discontinuity in the process that starts with eggs and sperms, but to regard the zygote as possibly a human seems to erase another presumed radical discontinuity, that between human beings, other animals, and even plants. If single cells are actual human beings, what about similar cells in other animals and in plants? And even if it is accepted that the human zygote is a form of human life, it cannot truly be claimed that an embryo is a human *individual* before the sixteenth day, because up until then there is a possibility that it may become twins (or more).

Faced by facts like these, some religious deontologists take recourse to the philosophical distinction between *actually* being of a certain kind and

having the *potentiality* of becoming a being of this kind. Then they say (in what is often called ‘the potentiality argument’): neither a sperm nor an egg has in itself the potentiality of becoming a human being, but the fertilized egg and each of its later stages has. Whereupon they interpret (or re-interpret?) the norm of not killing as saying: ‘You shall not kill living entities that either actually or potentially are human beings!’ The main counter (reductio in absurdum) argument goes as follow: if from a moral point of view we should treat something that is potentially H (or: naturally develops into H) as already actually being H, then we could treat all living human beings as we treat their corpses, since naturally we grow old and die; potentially, we are always corpses.

Even if, the comments above notwithstanding, a duty ethicist arrives at the conclusion that life begins at conception, abortions may still pose moral problems for him. How should he deal with extra-uterine pregnancy? If the embryo is not removed, then the woman’s life is seriously at risk, and gynecologists are not yet able to remove the embryo without destroying it. The duty to sustain the life of a fetus (a potential human being) is here in conflict with the duty to help an actual human being in distress, or even in life danger. The Catholic Church solves the problem by an old moral principle called ‘the Principle of Double Effect’. It says that if an act has two effects, one of which is good and one of which is evil, then the act may be allowed if the agent intends only the good effect. In the case at hand, it means that if the abortion is intended only as a means to save the life of the mother it can be accepted, even though a potential human being is killed; a precondition is of course that there is no alternative action with only good effects, i.e., an action that saves the life of both the fetus and the mother.

Let us take the opportunity to say some more words about the principle of double effect. It has won acceptance even among many non-religious people. For instance, think about whether to provide painkillers (e.g., morphine) to terminally ill patients. Morphine has two effects. It relieves the patient from pain, but it also suppresses the respiratory function, which may hasten the patient’s death. Here, also, the intention of the action might be regarded as crucial. If the intention is solely to relieve the patient’s pain, the action might be acceptable, but if the intention is to shorten life then, surely, the action is unacceptable.

Let us next look at another well-know principle, the principle of respecting the autonomy of the individual, and regard it as a deontological norm. What was earlier a problem of how to find a discontinuity that demarcates living from non-living human beings becomes now a problem of how to find a discontinuity that demarcates individuals from non-individuals. The principle can give to a woman (or the parents) an absolute right to choose abortion only if neither the embryo nor the fetus is regarded as an individual. Note that if an embryo and/or a fetus is regarded as a *part* of the pregnant woman on a par with her organs, then it cannot possibly be regarded as an individual human. But here a peculiar feature appears. The duty to respect others does not imply a duty to take care of others. Therefore, even if a fetus is regarded as an individual human, it seems as if abortion is acceptable from the autonomy principle; only a deontological principle of caring could directly imply a prohibition on abortion. In a much discussed thought experiment (J. J. Thomson 1971), the readers are asked to think that they one day suddenly wake up and find themselves being used as living dialysis machines for persons who have suffered renal failure. Isn't one then allowed just to rise, unplug oneself, and leave the room? But what if one has consented to be such a dialysis machine?

The arguments pro and con abortion that we have now presented in relation to classical deontology can also be arguments pro and con corresponding prima facie duties.

Some clergies do not consider abortion as a sin in itself, but use consequence reasoning in order to show that to allow abortions will probably lead to more sinful actions, i.e., actions that break some deontological norm. Already in the seventeenth century a Jewish rabbi, Yair Bacharach (1639-1702), who thought that fetuses are not human beings, argued that to allow abortions might open floodgates of amorality and lechery.

In rule utilitarianism we find a series of different arguments both for and against abortion. Here, the structure of the arguments is different, since there is no basic moral notion of human being or personhood. Utilitarians can very well accept moral rules that rely on completely conventionally created boundaries in the zygote-embryo-fetus development, if only the rules in question are thought to maximize utility.

One basic utilitarian argument for abortion is that an illegalization of abortion produces negative consequences such as deterioration of the quality of life of the parents' and their already existing children, or gives the (potential) child itself a poor quality of life. Such considerations become increasingly strong when the fetus has a genetic or chromosomal anomaly, which will result in a severe disease or mental handicap. In utilitarianism, however, there is no direct argument to defend that the mother or the parents should be sovereign in the abortion decision. Since the consequences can influence not only the family, but also the society at large in psychological, sociological, as well as economic respects, it is also a matter of utility calculations to find out who should be the final decision maker.

New medical technology has introduced quite new consequential aspects of abortion. A legalization of selective abortion or selective fetus reduction based on information about what sex, intelligence, risks for various diseases and disabilities the potential child has will probably also affect how already existing people with similar features experience their lives.

As soon as someone reaches the conclusion that abortion cannot in general be prohibited, he has to face another question: should there be a time limit for how late during pregnancy abortions can be allowed? One might even have to consider Singer's view that abortion can, so to speak, be extended to infanticide. Now we have to look at the development list again. Does it contain any discontinuities of special importance for the new question? And for the utilitarians there are. Especially, there is stage k, sentience. When there is a capacity for sensation and feeling, there is a capacity for pleasure and pain, which means that from now on an abortion may cause pain in the fetus. Abortion pills such as mifepristone might be painless, whereas prostaglandin abortions can cause the fetus to die through (on the assumptions given) painful asphyxiation. For Singer the discontinuity where self-awareness can be assumed to arise (around the third month after birth) is of very special importance.

In many countries where abortion is legal, the woman can decide for herself as long as the fetus is less than twelve weeks old; in some countries the limit is eighteen weeks, and with special permission from authorities it can be allowed up to twenty-two weeks. These limits are not only based on what properties the embryo as such has, they are also based on facts such

as the frequency of spontaneous abortion during different weeks, and the vulnerability of the womb before and after twelve weeks of pregnancy; after twelve weeks the womb is rather vulnerable to surgical intervention. If one day all abortions can be made by means of medical intervention a limit like this have to be re-thought.

Stage j (viability) in our list, which seems unimportant to deontologists because it is heavily dependent on the development of medical technology, receives a special significance in utilitarianism. Why? Because it may make quite a difference to the parents' experiences and preference satisfactions if the aborted fetus could have become a child even outside the uterus. But such experiences can also undergo changes when people become used to new technologies. Here again we meet the utilitarians' problems with actual utility calculations. It depends on the result of his utility calculation whether he should put forward the norm: abortions after the moment of which a fetus becomes capable of surviving outside the womb should not be allowed.

We have defined a modern virtue ethicist as a person who:

(i) accepts moral particularism
(ii) accepts a conflicting psychology
(iii) accepts moral fallibilism
(iv) realizes that characters can be situation-bound
(v) accepts that he has to discuss the moral aspects of certain social rules.

Such a virtue ethicist cannot rest content with merely saying that virtue ethics leaves all rules behind. What does the problem of abortion look like to him? In one sense it is similar to that of the rule utilitarian, i.e., he ought to think about the consequences, but he can do it having some rights of persons as default norms. Modern virtue ethics is so to speak both quasi-deontological, since it accepts default norms, and quasi-consequentialist, since one of its default rules is take also consequences into account. But there is still something to be added: the virtue ethicist is careful as long as he has had no personal encounter with people who have considered abortions, having aborted, and who have abstained from abortion. The more experience of this kind he has, the better. Why? Because through

such experience he may acquire tacit knowledge that influences what abortion rule he will opt for.

9.6 Medical ethics and the four principles

Due to the fact that neither classical deontology nor consequentialism have managed to come up with norms that provide reasonable and comprehensive guidance in the medical realm, the American philosophers Tom Beauchamp and Jim Childress have, with great resonance in the medical community, claimed that the latter in both its clinical practice and research ought to rely on four *prima facie principles* (in Ross' sense). Relying on our comments on moral particularism in the last two subchapters, we would like to reinterpret these principles as *default principles*. But this change is more a philosophical than practical. It means, though, that we think that these rules fit better into virtue ethics than into deontological and utilitarian ethics. The four principles are rules of:

1) Beneficence. There is an obligation to try to optimize benefits and to balance benefits against risks; a practitioner should act in the best interest of the patient and the people. In Latin, briefly: *Salus aegroti suprema lex*.

2) Non-maleficence. There is an obligation to avoid, or at least to minimize, causing harm. In Latin, briefly: *Primum non nocere*. Taken together, these two principles have an obvious affinity with the utility principle; taken on its own, the latter has affinity with so-called 'negative consequentialism'.

3) Respect for autonomy. There is an obligation to respect the agency (cf. Chapters 2.1 and 7.5), reason, and decision-making capacity of autonomous persons; the patient has the right to refuse or choose his treatment. In Latin, briefly: *Voluntas aegroti suprema lex*. This principle has affinity with Kant's imperative never to treat people only as means, and it implies sub-rules such as 'Don't make false promises!', 'Don't lie or cheat!', and 'Make yourself understood!' When respect for autonomy is combined with beneficence, one gets a sub-rule also to

enhance the autonomy of patients; it might be called ‘the principle of empowerment’.

4) Justice. There are obligations of being fair in the allocation of scarce health resources, in decisions of who is given what treatment, and in the distribution of benefits and risks. It should be noted that justice means treating equals equally (horizontal equity) and treating unequals unequally (vertical equity).

These principles are as prima facie principles or default principles independent of people’s personal life stance, ethnicity, politics, and religion. One may confess to Buddhism, Hinduism, Christianity, or Islam, or be an atheist, but still subscribe to the four principles. They might be regarded as the lowest common denominator for medical ethics in all cultures. In specific situations, however, the four principles well may come in conflict with religious deontological norms, secular deontological norms, ordinary socio-political laws, and various forms of utilitarian thinking. How to behave in such conflicts lies outside the principles themselves; they do not provide a method from which medical people can deduce how to act in each and every situation.

Normally, the principles are not hierarchically ordered, but it has been argued (Gillon 2004) that if any of the four principles should take precedence, it should be the principle of respect for autonomy. Sometimes they are presented as (with our italics) ‘four principles *plus attention to scope*’ (Gillon 1994). Obviously, we can neither have a duty of beneficence to everyone nor a duty to take everyone into account when it comes to distributive justice. What scope do then principles one and four have? In relation to principle three one can ask: who is autonomous? Those who subscribe to the four principles have to be aware of this ‘scope problem’. In the terminology we have earlier introduced, we regard this ‘scope problem’ as a ‘phronesis problem’.

Often, in the medical realm ethical reasoning is performed by persons working in teams. This adds yet another feature to the complexity we have already presented. We have spoken of one man’s phronesis and moral decision when confronted with a specific situation. But in teams, such decisions ought to be consensual. For instance, a ward round can contain a

chief physician, specialist physicians, nurses, and assistant nurses. When, afterwards, they discuss possible continuations of the treatment, even moral matters can become relevant, and the differences of opinion can be of such a character that literal negotiations with moral implications have to take place. Then, one might say that it befalls on the group to develop its collective phronesis. In clinical ethical committees, which can be found in almost every Western hospital, similar kinds of situations occur. In such committees, which are not to be confused with research ethics committees (see Chapter 10), representatives from various specialties and professions convene in order to solve especially hard and critical situations.

9.6.1 The principles of beneficence and non-maleficence

As is usually done, we will remark on the first two principles simultaneously. One reason for bringing them together is the empirical fact that many medical therapies and cures have real bad side effects; surgery and chemotherapy might cure a patient in one respect, but at the same time cause some handicap in another. Sometimes one dangerous disease is treated with another dangerous disease; for instance, in 1927 the Austrian physician Julius Wagner-Jauregg (1857-1940) received the Nobel Prize for treating syphilis with malaria parasites.

But the two principles belong together even from a philosophical point of view. If there are several possible beneficial, but not equally beneficial, alternatives by means of which a patient can be helped, then to choose the least beneficial is, one might say, a way to treat the patient in a maleficent way. As in utilitarianism degrees of pleasures and pains naturally belong to the same scale, here, degrees of beneficent and maleficent behavior belong together. If a treatment has both desired and undesired effects, it is important to balance these effects and try to make the treatment optimal. The same holds true of diagnostic processes; there can be over-examinations as well as under-examinations, and both are detrimental to the patient.

According to the Hippocratic Oath, each physician is expected to testify:

- I will apply dietetic measures for the benefit of the sick according to my ability and judgment; I will keep them from harm and injustice.

This has been regarded a basic norm in medicine since ancient times. In more general words: physicians should try to do what is supposed to be good for the patient and try to avoid what may harm or wrong him. The ambition to avoid making harm is weaker than the ambition ‘first of all do not harm’ (‘primum non nocere’), a saying which is often, but falsely, thought to come from Hippocrates. The latter rule did first see the light in the mid nineteenth century. It was occasioned, first, by the rather dangerous strategy of bloodletting, and, later, the increased use of surgery that followed the introduction of anesthesia. Some patients died as a result of bloodletting; several physicians recommended that as much as 1200 ml blood should be let out, or that the bloodletting should continue until the patient fainted.

While minimizing-harm might appear to be an obvious ambition, the history of medicine unfortunately provides several examples where one has not been careful enough. In Chapter 3.3, we told the sad story about Semmelweis and the Allgemeine Krankenhaus in Vienna. The current perception of what constitutes good clinical practice does not automatically prevent medical disasters. Therefore, in most modern societies, physicians have to be legally authorized before they are allowed to treat patients on their own.

Proposed cures may be counter-productive, especially when a treatment initially connected to a very specific indication is extended to other more non-specific indications. Cases of lobotomy and sterilization provide examples. Lobotomy was invented by António Egas Moniz (1874-1955) to cure psychiatric anxiety disorders. At the time of the invention, no other treatments were available, it was applied only in extremely severe cases, and used only on vital indication. Moniz’s discovery was regarded as being quite a progress, and he was awarded the Nobel Prize in 1949. Unfortunately, it was later used in order to cure also disorders such as schizophrenia and neurosis; it was even applied in some cases such as homosexuality, where today we see no psychiatric disorder.

Before the birth control pills were introduced in the 1960s, it was difficult to prevent pregnancy in mentally handicapped individuals. Therefore, many of these individuals were sterilized. The reason was of course that a mentally disabled person is assumed not to be able to take the

responsibility of parenthood, and that children have a right to have at least a chance to grow up in a caring milieu. Sterilization meant that the mentally handicapped didn't need to be kept in asylums only in order to stop them from becoming pregnant. Today we use birth control pills and other contraceptive medication for the same purpose. Sterilization, however, was also used in connection with eugenic strategies, which are very controversial because mental handicap is not always inherited; and when it in fact is, it is neither monogenetic nor a dominant trait. This means that (according to the Hardy-Weinberg law of population genetics, presented 1908) a sterilization of all mentally handicapped individuals has almost no effect in a population-based perspective. Nevertheless, in several Western countries in the 1940s and 1950s, thousands of mentally disabled individuals, as well as others with social problems, were sterilized on eugenic grounds.

It is easy to be wise after the event; the challenge is to be wise *in* the event. Therefore, doctors have to consider whether their actions will produce more harm than good. It seems to be extremely rare that individual physicians deliberately harm patients, but, as Sherlock Holmes said about Dr Roylot (in 'The Speckled Band), when doctors serve evil they are truly dangerous. The ideal of beneficence and non-maleficence is the ideal for physicians, nurses, and health care staffs. In many countries this fact is reflected in the existence of a special board, which assesses every instance in which patients are put at risk by unprofessional behavior. Also, omitting taking care of patients in medical need is usually understood as severe misconduct.

In the Hippocratic Oath, a physician is also expected to testify that:

- I will neither give a deadly drug to anybody who asked for it, nor will I make a suggestion to this effect. Similarly I will not give to a woman an abortive remedy. In purity and holiness I will guard my life and my art.

This part of the oath has to be re-thought in the light of the beneficence and non-maleficence principles. We have already commented upon abortion. What about euthanasia? Is it defensible to help terminally ill patients suffering unbearable pain to die? What is to be done when no

curative treatment is available, and palliatives are insufficient? Doctors are sometimes asked to provide lethal doses of barbiturate. Do physicians harm a patient that they help to die? Do they harm someone else? As in the case of abortion, there are many complex kinds of cases. Within palliative medicine, a terminally ill patient might be given sedation that make him fall asleep, during sleep no further treatment or nutrition is given, and because of this he dies within some time. Here, again, 'the principle of double effect' might be brought in. One may say that the intention is only to make the patient sleep, which is a good effect, but then inevitably there happens to be also a bad effect, the patient dies. Usually, this is not understood as euthanasia.

The original Greek meaning of 'euthanasia' is 'good death', but the word acquired an opposite negative ring due to the measures imposed by the Nazis during the years 1942-1945. They called euthanasia the systematic killing of people not regarded as worth living (often chronically ill patients and feeble-minded persons). Today, euthanasia is defined as:

- a doctor's intentional killing of a person who is suffering 'unbearably' and 'hopelessly' – at the latter's voluntary, explicit, and repeated request.

Euthanasia is then distinguished from 'physician-assisted suicide', which can be defined as:

- a doctor's intentional helping/assisting/co-operating in the suicide of a person who is suffering 'unbearably' and 'hopelessly' – at the latter's voluntary, explicit, and repeated request.

The Nazi misuse of the word euthanasia has given emphasis to a slippery slope argument against euthanasia: if 'good euthanasia' is legalized, then also 'bad euthanasia' will sooner or later be accepted. And, it is added, when this happens, the trust in the whole health care system will be undermined. In Holland and Belgium, where euthanasia was legalized in 2002 and 2003, respectively, nothing of the sort has so far happened. In a democratic setting, the risk of entering a slippery slope might be very small, but there are still very important concrete issues to discuss. If

society legalizes euthanasia, under what conditions should it be accepted and who should perform it?

In 1789, the French physician Joseph Guillotine (1738-1814) suggested a less harmful method of execution than the traditional ones. Hanging had come to be regarded as inhumane to both the criminal and the audience. The guillotine was introduced with the best of intentions, but during the French Revolution it became a handy instrument of mass executions. Dr Guillotine dissociated himself from its use and left Paris in protest. Against this background, it is worth noting that the American Medical Association has forbidden its members to participate in capital punishment by means of drug injections. Even though chemical methods (e.g., sleeping medicine in combination with curare or insulin) might appear more human and less harming than execution by hanging, gas, or the electric chair, it might be discussed whether it is at all acceptable for a physician to participate in these kinds of activity. In brief: does the principle of non-maleficence imply that physicians ought to refuse to participate in capital punishments?

9.6.2 The principle of respect for autonomy

Patient autonomy might be defined as a patient's right to take part in medical decision-makings that lead to decisions by which he is more or less directly affected. If he is the only one who is affected, then he should be allowed to make the decision completely on his own, but this is rather uncommon. Autonomy means more than integrity and dignity. To respect a patient's integrity or dignity is to respect his wishes, values, and opinions even though he is in the end not allowed to be part of the final decision. Autonomy takes degrees, it might be strong (as in normal adults), weak (as in small children), or entirely absent (as in unconscious patients). Integrity, on the other hand, is usually regarded as non-gradable. Wishes might be important and should be respected as far as possible even in patients that lack autonomy and in dead patients. Such wishes may be manifested orally or in so-called 'advance directories'. Since, advance directories are rather uncommon, relatives who knew the patient may have to interpret his wishes in relation to whether to withhold or to withdraw a treatment or to donate organs or tissues.

If we speak about the patient's right to take part in decision making, it is necessary to distinguish between, on the one hand, the patient's right to

decline examination, treatment, or information and, on the other hand, his entitlement to exercise his positive rights. Usually, it is easy in principle to respect the autonomous patients' right to say no, i.e., to respect their negative rights, but it might nonetheless sometimes be hard in practice. The competent Jehovah's Witness, who rejects life supporting treatment if it includes blood transfusion, is the classical illustrative case. The positive rights, however, are complicated even in principle since their realization involve costs and have repercussions on queuing patients.

Sometimes a patient requests examination and treatment where there seems to be no clear medical need. In such cases, to respect autonomy means initiating negotiations. If a patient with a two day old headache visits his GP and asks for a computer tomography (CT) of his skull, the GP might well challenge the reason for conducting this examination. If, after having examined the patient, the GP finds that the symptoms more probably derive from the muscles of the neck, he should suggest physiotherapy – and ask if this might be an acceptable alternative. Unless the GP suspects that a brain tumor or other pathological change might be the cause, the autonomy principle by no means implies that he should grant the patient's wish. This would be much more than to respect the autonomy of the patient; it would mean that the GP subordinates himself to the patient. A doctor has to make reasonable estimations of probabilities, and react (Africa apart) according the saying, 'if you hear the clopping of hooves outside the window, the first thing that comes to mind is not a zebra'.

Were all patients presenting a headache referred to a CT of the skull, radiologists and CT-scanners would be swamped. The skilled GP should be able to distinguish between a patient with a tension-based headache and a brain tumor. The reason why the GP does not simply agree is thus partly his skill and knowledge. Patients have to accept a special kind of autonomy, the 'professional autonomy' of the physicians. Doctors should consider the patients' problem with empathy and with respect of their autonomy, while at the same time respecting *his* own professional autonomy. Ideally, the negotiations between doctors and patients should be based on mutual trust and aim at reaching consensus. This view is referred to as 'patient-centered medicine', in contrast to 'physician-centered medicine'. The latter is characterized by the doctor setting the agenda, and

since the doctor knows more about medicine, he is also assumed to know independently of the patients' wishes what is best for the patient.

Since autonomy has degrees, the autonomy principle cannot possibly altogether forbid paternalistic behavior on the part of physicians and health care workers. The word paternalism comes from the Latin 'pater', which means father, and today it refers to the decision making role of the father in traditional families. Parents of small children must decide what is in the best interests of their children, and now and then there are structurally similar situations in medicine. When there is a considerably reduced autonomy, as in some psychotic patients, there might be reasons not even to respect the patient's negative right to refuse treatment. This might be called 'mild paternalism'. All paternalistic actions must be for the benefit of the patient; some paternalistic actions can be defended by an assumption that the patient will approve of the action when he recovers from his disease or disorder; this is so-called 'presumed informed consent'. Of special concern are patients suffering from Alzheimer's disease. When the disease has progressed, the patients are often in a state in which they do not know their own best interest, and here weak paternalism is the rule. Not being paternalistic would amount to cynicism.

During the last decades of the twentieth century, the doctor-patient relationship became both formally (in laws) and really 'democratized', i.e., the doctor's paternalism was decreased and the patient's autonomy increased. The patient-centered method is an expression of this change. In the Hippocratic period, medical knowledge was secret, and not supposed to be made freely available; during the last centuries it was public in theory but hard to find for the laymen. Today, internet has changed the picture completely. Rather fast, and with a very small effort, many patients check a little on the internet what their symptoms can mean before they meet their GP. And after having been diagnosed, some of them seek contact with other patients with the same diagnosis in order to discuss treatments and side effects or they ask another doctor for a second opinion.

9.6.3 The principle of justice

Modern medical research has brought about many new medical technologies within preventive and curative medicine, as well as very advanced high tech examinations, and our knowledge of diseases, their

etiology, and their pathogenesis has grown. In other words, health care systems are today able to do and offer much more than only, say, three or four decades ago. During the 1970s and 1980s, many health care systems were in a phase of rapid expansion from which quite a number of new specialties emanated. Many modern hospitals and health care centers were built, and an increasing number of physicians and nurses were educated to meet the needs.

At the same time, quite naturally, the patients' expectations rose. During the 1980s, however, the costs of the expansion became a public issue much discussed by politicians. To put it briefly and a bit simplified, it became clear that the costs of the health care system had reached a limit. This insight put to the fore priority settings. Problems of what a fair health care distribution and a fair cost distribution should look like came to be regarded as very serious. Should the healthcare system be based on public finances, private finances, or a mixture of the two? Should patients be provided with health care through a random principle? Should health care be provided depending on the patients' verbal capacity, financial situation, and social influence? Should it be based on some egalitarian principle? Let us take a brief look at some of these options.

By means of thought experiments, it might be easy to conceive of situations where priority settings based on a random principle makes sense. For example, assume that Smith and Thomson (who are of the same sex and age, and have the same position in society, etc.) have suffered from a heart attack, and have received the same diagnosis and the same prognosis both with and without treatment. Furthermore, they arrive at the same emergency clinic at the same time, where, unhappily, it is only possible to treat one of them with the high tech means available to the cardiology intensive care unit. Here it seems adequate to flip a coin. If Smith wins, the doctor might have to say to Thomson: 'Sorry, you were out of luck today. Please, go home and try to take it easy the next couple of weeks. You may die or you may survive, and if you survive you are welcome back should you suffer a new heart attack'. Reasonable as it may seem, one might nonetheless ask if this would be fair to Thomson. In one sense Smith and Thomson are treated equally (when the coin is flipped), and in another they are treated unequally (only one is given the treatment). If both Thomson and Smith had voted for the political party that had introduced the random

system, it might seem fair to treat them in the way described, but what if the procedure has been decided quite independently of Smith's and Thomson's views? Would it be fairer to treat neither Thomson nor Smith at the clinic, and offer both of them low tech oriented care and painkillers?

In thought experiments one can imagine people to be exactly similar, but in real life this is hardly ever the case. Different patients suffer from different diseases, and some of them have had the disease a long time and others only a short time. Furthermore, some diseases are life threatening and others not. Some patients are young and others old, some are male and some are female, and they come from different ethnic cultures. Also, they have different social status, different incomes, and pay different amounts of taxes. Some are employed and some unemployed, some are refugees and some have lived in the society in question for generations. How to apply a random principle in the midst of such variations? In several countries where priority rules have been developed, it is said that those most in need of a treatment should be prioritized and assessed in relation to the severity of the disease. Sometimes, however, so-called VIPs (famous politicians, movie stars, sports heroes, etc) become (especially on operation lists) prioritized before other patients who have been waiting for a long time.

Brainstormers have argued that the state should at birth supply each member of the society with a certain amount of money, which should be the only money he is allowed to use for his health care. Furthermore, during the first seventy years he is not allowed to use this sum on anything else but health care, but then he is free to use what remains in any way he wants. This means that when the money runs out, it is impossible to receive more health care services. If you are unlucky and suffer from diabetes or a chronic disease early in your life, then the health care money may run out when you are quite young, whereas if you are lucky you can have much money to spend on possible later diseases. Even in this proposal, individuals are in one sense treated equally (they receive the same sum at birth) and in another unequally (people who suffer the same disease may nonetheless later in life not be able to buy the same treatment). The proposal is in conflict with the kind of solidarity based justice that says that it is fair that the healthy citizens subsidize the sick ones. Others claim that such solidarity justice is theft, and that the only fair way of

distributing health care resources is to let only people who can pay for it receive it. In the case of Thomson and Smith described, the reasonable thing would then be to sell the high tech treatment by means of an auction where Thomson and Smith can bid.

A special problem of justice in priority settings appears in relation to wars and war-like situations. Is it fair to extend to enemies a principle such as ‘patients who are most in need of treatment should be given treatment first’? The extreme case is a suicide bomber that survives together with many injured people. Assume that he is the most injured, and that to save his life would imply a long and complicated operation that would take resources away from the innocently injured, and cause them further suffering; some might even die. What to do? Justice seems to be justice in a pre-given group, and then one can ask what this group looks like for physicians. Does he belong to mankind as a whole, to the nation to which he belongs, or to some other community? Is perhaps the question of justice wrongly formulated? Perhaps an analogy with counsels for the defense in law courts is adequate? The judge and the jury are obliged to come to a fair decision (and a possible punishment), but the counsel is not. Society has prescribed a division of labor according to which the task of the counsel is to work solely in the interest of his clients. Perhaps the world community should place physicians in a similar situation in relation to prospective patients?

Reference list

Apel K-O. *Towards a Transformation of Philosophy*. Routledge. London 1980.

Aristotle. *Nicomachean Ethics*. (Many editions)

Beauchamp TL, Childress JF. *Principles of Biomedical Ethics*. Oxford University Press. Oxford 2001.

Bentham J. *An Introduction to the Principles of Morals and Legislation*. (Many editions)

Bernstein RJ. *Beyond Objectivism and Relativism*. Basil Blackwell. Oxford 1983.

Bok S. *Lying. Moral Choice in Public and Private Life*. Vintage Books. New York 1979.

Brandt, RB. *A Theory of the Good and the Right*. Oxford University Press. Oxford 1979.

Broad CD. *Five Types of Ethical Theory*. London. Routledge 2000.

Capron AM, Zucker HD, Zucker MB. *Medical Futility: And the Evaluation of Life-Sustaining Interventions*. Cambridge University Press. Cambridge 1997.

Crisp R, Slote M (eds.). *Virtue Ethics*. Oxford University Press, Oxford 1997.

Curran WJ, Casscells W. The Ethics of Medical Participation of Capital Punishment by Intravenous Drug Injection. *New England Journal of Medicine* 1980; 302: 226-30.

Dancy J. *Ethics Without Principles*. Oxford University Press. Oxford 2004.

Dancy J. Moral Particularism. *Stanford (online) Encyclopedia of Philosophy* (April 11, 2005)

Doris JM. *Lack of Character: Personality and Moral Behavior*. Cambridge University Press. New York 2002.

Dreyfus H, Dreyfus S. What is Moral Maturity? Towards a Phenomenology of Ethical Expertise. In Ogilvy J. (ed.). *Revisioning Philosophy*. Suny Press. New York 1986.

Edelstein L. *The Hippocratic Oath: Text, Translations and Interpretation*. John Hopkins Press. Baltimore 1943.

Fulford KWM, Gillet G, Soskice JM. *Medicine and Moral Reasoning*. Cambridge University Press. Cambridge 1994.

Fletcher J. *Situation Ethics: The New Morality*. John Knox Press. Louisville, Ky. Westminster 1997.

Fesmire S. *John Dewey: Moral Imagination. Pragmatism in Ethics*. Indiana University Press. Bloomington 2003.

Gewirth A. *Reason and Morality*. The University of Chicago Press. Chicago 1978.

Gilligan C. *In a Different Voice: Psychological Theory and Women's Development*. Harvard University Press. Cambridge Mass. 1982.

Gillon R. Lloyd A. *Health Care Ethics*. Wiley. Chichester 1993.

Gillon R. Medical Ethics: Four Principles Plus Attention to Scope. *British Medical Journal* 1994; 309: 184.

Gillon R. Ethics needs principles – four can encompass the rest – and respect for autonomy should be 'first among equals'. *Journal of Medical Ethics* 2003; 29: 307-12.
Habermas J. *Moral Consciousness and Communicative Action*. Polity Press. Cambridge 1990.
Hare RM. *Moral Thinking*. Oxford University Press. Oxford 1981.
Helgesson G, Lynöe N. Should Physicians Fake Diagnoses to Help their Patients? *Journal of Medical Ethics* (forthcoming).
Holm S. The Second Phase of Priority Setting. *British Medical Journal* 1998; 317: 1000-7.
Kjellström R. Senilicid and Invalidicid among Eskimos. *Folk* 1974/75; 16-17: 117-24.
Kant I. *Groundwork of the Metaphysics of Morals*. (Many editions.)
Kohlberg L. *The Philosophy of Moral Development*. Harper & Row. San Francisco 1981.
Kuhse H, Singer P. *Should the Baby Live?* Oxford University Press. Oxford 1985.
Lynöe N. Race Enhancement Through Sterilization. Swedish Experiences. *International Journal of Mental Health* 2007; 36: 18-27.
McGee G. *Pragmatic Bioethics*. Bradford Books. Cambridge Mass. 2003.
MacIntyre A. *After Virtue*. Duckworth. London 1999.
Mill JS. *Utilitarianism*. (Many editions.)
Norman R. *The Moral Philosophers. An Introduction to Ethics*. Oxford University Press. Oxford 1998.
Nord E. *Cost-Value Analysis in Health Care. Making Sense of QALYs*. Cambridge University Press. New York 1999.
Nussbaum, MC. *Love's Knowledge: Essays on Philosophy and Literature*. Oxford University Press. New York 1990.
Nussbaum M. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press. Cambridge 2001.
Rawls J. *A Theory of Justice*. Harvard University Press. Cambridge Mass. 1971.
Ross WD. *The Right and the Good*. Oxford University Press. Oxford 2002 .
Sharpe VA, Faden AI. *Medical Harm: Historical, Conceptual, and Ethical Dimension of Iatrogenic Illness*. Cambridge University Press. Cambridge 1998,
Singer P. *Practical Ethics*. Cambridge University Press. Cambridge 1993.
Smith B, Brogaard B. Sixteen Days. *The Journal of Medicine and Philosophy* 2003; 28: 45-78.
Smith CM. Origin and Uses of Primum Non Nocere – Above All, Do No Harm! *Journal of Clinical Pharmacology* 2005; 45: 371-7.
Statman D. (Ed.) *Virtue Ethics. A Critical Reader*. Edinburgh University Press. Edinburgh 1997.
Swanton C. *Virtue Ethics*. A Pluralistic View. Oxford University Press. Oxford 2003.
Thomson JJ. A Defense of Abortion. *Philosophy & Public Affairs* 1971; 1: 47-66.

Umefjord G, Hamberg K, Malker H, Petersson G. The use of Internet-based Ask the Doctor Service involving family physicians: evaluation by a web survey. *Family Practice* 2006; 23: 159-66.

World Medical Association, Declaration of Geneva 1948. See e.g. < www.cirp.org/library/ethics/geneva >