

VI. Good

1. The Good, Morally and Extramorally

Among all the things that exist, some stand out in our eyes by being good. But what is good? According to the classic definition, which goes back to Plato and Aristotle, the good is “that at which all things aim”.¹ If all things aim at the good, all human beings do so. Therefore, we can adapt the classic definition as follows: The good is what every human being aims at.

But this definition can again be met by the question which remains open: Is what every human being aims at the good? We would answer this question in the negative. What everybody aims at is not always the good. Everybody seems to aim at pleasure. But does that make pleasure the good? We cannot answer this question in the affirmative, because obviously there are bad pleasures, for example, the pleasure of the sadist. Thus, the concept of good, too, contains a surplus meaning, which is not rendered by the classic definition. What I have said about the concept of truth and about the concept of being also applies to the concept of good: It cannot be defined explicitly, but only implicitly. To put it differently: It can only be elucidated. To explain the concept of good is to make conscious what we already know of it in an undeveloped form.

In elucidating this concept, it is again advisable to begin with language. Like the terms “true” and “is”, “good” has *several* meanings, between which the classic definition – “at which all things aim” – makes no distinction. When we say “A glass of wine is good”, we do not mean the same thing as when we

¹ Aristotle, EN, Book 1, Chapter 1, 1094a2-3. Transl. Ross. Cf. Plato, Grg. 468b, 499c-500a. Cf. for more information Ferber, *Ho de diōkei*.

say “A will or an intention is good.” In the first instance, we mean that a glass of wine is a good means to an end, say, for our health or enjoyment. In the second instance, we mean that a will or an intention is good in itself. In the first instance, we invest the term “good” with a relative or instrumental meaning, in the second instance, with a (comparatively) absolute or moral meaning. In what follows, I will not discuss what is good merely in an instrumental or relative sense, but what is good in itself or in a moral respect. The discipline that examines the latter is called ethics.

Ethics, according to an apt definition by George Edward Moore (1873-1958) in his *Principia Ethica* (1903), is “the general enquiry into what is good”.² However, as we have confined ourselves to the morally good, I can narrow the definition down for our purposes as follows: Ethics is the general inquiry into what is morally good. By morality, I mean the kind of practical behaviour that corresponds to the theory of ethics.

What is morally good can be felt or told. When we tell what is good, we express ourselves in sentences that are not only descriptive but also evaluative. We judge human beings and their qualities, among other things, as good or bad, and their actions as right or wrong. Ethics, then, examines not only what is good and right, but also what is bad and wrong. What is good should be done; what is bad should be avoided. What is neither good and right nor bad and wrong, but indifferent, may be either done or avoided. That is why the language of morality contains not only evaluative sentences, but also some that command, forbid or permit something. For the language of morality, those sentences that either command or forbid something are particularly significant. They are also called normative sentences. But as

² Moore, PE, Chapter 1, § 2, 2.

values are not norms, so evaluative sentences are not normative ones. The value of human life, for example, is not the same as the norm not to destroy human life. It is, rather, the foundation of this norm, which we feel in our conscience. Conscience may be considered the awareness of the normative claim of values. In a similar way, the normative sentence, the proposition “Though shalt not kill” – or more exactly its content – is founded in the evaluative proposition “Human life is valuable.”

An examination of these evaluative and normative sentences is not in itself an evaluative or normative ethic. It does not tell us what is good or bad, right or wrong, to be done or not to be done. It only talks *about* the sentences we use to say that something is good or bad, right or wrong, or that we should do this and not do that. That is why this examination is also called *metaethics*. Metaethics is the study of moral language. It includes in particular two theories of evaluative or normative statements: cognitivism and emotivism.

2. The Metaethics of Moral Good

a) *Cognitivism*

The most obvious theory is cognitivism. According to this theory, moral sentences have the same status as those statements that we use to express an insight. It seems clear that we can tell as easily what is good or bad, right or wrong, as we can tell what is white and what is black. In both instances, we only need to open our eyes. That it is morally right to dress a bleeding wound, and morally wrong to let a person bleed to death, seems to be as clearly visible as the fact that snow is white and pitch is black.

This theory has several advantages. First, it agrees with our evaluative moral language. We speak in the indicative about

moral properties (“X is good or bad, right or wrong”), as well as about natural ones (“X is white or black”), and we attribute the values true or false to moral propositions as well as to assertions of facts. Further, this theory can easily explain the absolute validity that we attribute to *certain* moral values by stressing their reality and objectivity. In addition, realism in ethics and in epistemology is a basic attitude of common sense and a recurrent philosophy. It has been advocated by the majority of philosophers from Plato and Aristotle to G E Moore and others.³ Finally, it has the advantage that, unlike ethical scepticism, which claims that we can never recognise anything as good or bad, we do not immediately abandon it – at any rate outside moral philosophy.

A particularly clear formulation of ethical realism is found in *Memoirs from the House of the Dead* by Fyodor Mikhailovich Dostoevsky (1821-1881): “There are certain crimes which, from the beginning of the world, under every code of law, have always and everywhere been regarded as indisputably crimes and will continue to be so regarded while men are men.”⁴ Such a crime, for example, would be the murder of a whole people, or genocide. Conversely, we can say that some deeds are undeniably morally right and will remain so as long as human beings remain human beings. Thus, it is undeniably right for one human being or one nation to save another from starvation.

Cognitivism, then, leads to moral objectivism and realism. It recognises moral facts in reality. These are objective in so far

³ Cf. e.g. Plato, R., Book 4, 427d-434c; Book 6, 504a-506a; Book 7, 534b-c. Transl. Jowett. Aristotle, EN, Book I, Chapter 1, 1094a, 22-26. Moore, PE, Chapter 1, §10, 9-10.

⁴ Part 1, Chapter 1. Transl. Ronald Hingley and Jessie Coulson, Oxford 2008.

as they exist in themselves and not just for us, as does, for example, the fact that genocide is bad and saving people from starvation is good. The fundamental thesis of cognitivism can be formulated as follows: Moral sentences – or, more accurately, their content, the propositions – are true or false because they either agree or do not agree with moral facts.

Among these moral facts, we can distinguish two kinds: the basic and the derived. A basic or indeed “axiomatic” moral fact is conveyed by the first sentence of the Basic Law of the Federal Republic of Germany: “Human dignity is inviolable.” From this “axiomatic” proposition, it is possible to “derive” others, for example, that every human being has the right to life and physical integrity, and that the freedom of the individual is un-infringeable (Article 2.1).

But how do we recognise moral facts? Obviously, moral facts cannot be real facts of a physical nature, as, for example, the fact that snow is white and pitch is black. While even such physical facts are really hypothetical, moral qualities are downright invisible to us. We can see with our own eyes that snow is white and pitch is black, but we cannot read the moral quality of a face or an action directly from the outside. A face may seem friendly to us and yet hide unfriendly thoughts behind the smile. The features of a criminal, as a rule, are no different from those of “decent” people, as any visitor to a prison can confirm. An action like the transfer of a bleeding person from one car to another may be a rescue or a kidnap.

The qualities “good”, “bad”, “right” or “wrong” have no effect on our sensory organs, or at least not the same effect as the qualities “white” or “black”. Likewise, by just looking at Mr Smith or Mrs Jones, we cannot immediately tell that they possess inviolable dignity. As we do not perceive these moral qualities with our bodily eyes, cognitivism was able to conceive the idea that they are of a “supernatural” or metaphysical nature and

can only be “seen” with a “mind’s eye”. We do not see with our bodily eye that it is right to dress a bleeding wound, but wrong to let a person bleed to death, or that a human being possesses inviolable dignity. To “see” such things, we must “open our intellectual eye”.

The hypothesis of direct intellectual vision is by no means restricted to “seeing” mathematical or geometrical axioms, such as “The whole is greater than the part” (cf. p. 83). Rather, it has been transferred from “seeing” mathematical and geometrical axioms to “seeing” moral ones. As we cannot see moral facts either, even though they somehow seem to exist, the linguistic expedient of talking about non-sensory vision seemed appropriate. Cognitivism thus leads to intuitionism (from *intueor*: I gaze).

Regardless of the fact that the strange “opening of the mind’s eye” is a metaphorical auxiliary construction, it was precisely this intellectual intuition that supplied a decisive argument against objectivism: That intuition is an objective criterion is no truer of moral axioms than it is of the axioms of arithmetic and geometry. The intuitively plausible axiom that the whole is greater than the part, for example, is not valid for infinite quantities (cf. p. 84). Likewise, even an eye as sharp as that of Aristotle failed to “see” the inviolable dignity of the human being and the unfringeable freedom of the individual, believing as he did that some people were slaves by nature (cf. p.107). It takes an intuition that has grown and developed historically – that is, an acquired intuition – to “see” that axiom in its universal binding force.

But even an acquired intuition may come up against borderline cases in which it no longer sees clearly. Does a person who has been in a coma for the past two years still have inviolable dignity? Moral intuition, like mathematical intuition, may offer *prima facie* evidence (cf. p. 86), but by no means guarantees the

impartiality claimed by the objectivist. It can be corrected or even abrogated by other “intuitions”, as is, for example, the case with passive or – in the event of an incurable and unbearable illness – in certain circumstances active euthanasia.

Nor is the objectivity of this intuition alone open to doubt. There is also a possible argument against the concept of moral facts. Moral facts are not only facts, but norms. That the dignity of the human being is inviolable is a fact as well as a norm, that is, a ban on violating human dignity: If human dignity *is* inviolable, it follows that it *should* not be violated. If killing *is* morally wrong, then thou *shalt* not kill. If moral judgments are statements of facts, a constative proposition gives rise to a normative one.

Here it may be objected that it is not admissible to infer a normative proposition from a constative one. As this objection goes back to Hume’s *A Treatise of Human Nature* (1739-40), it is also called “Hume’s law”, which states that one cannot derive an “ought” from an “is”.⁵ In any valid deductive conclusion, the content must not go beyond that of the premises. A valid deductive conclusion is truth-preserving (cf. p. 61). However, if a normative conclusion is inferred from constative premises, the conclusion does not preserve the truth of the premises, but goes beyond their meaning. It adds something new, which was not included in the premises, namely, an obligation.

Cognitivism and intuitionism regard evaluative and normative sentences, such as “Killing is wrong” or “Thou shalt not kill”, as constative statements in linguistic disguise, which as such are true or false in so far as they agree or do not agree with moral facts. But then, according to “Hume’s law”, it is not possible to derive normative sentences from them. If the sentences

⁵ Hume, *Treatise*, Book 3, Part 1, Section 1, 469-470.

“Killing is wrong” or “Thou shalt not kill” state a fact, then “thou shalt” has no normative, but only constative, force. In that case, however, the ban on killing is not itself a ban, but the statement of a ban, and it seems that no valid norm can be derived from it. Like the derivation of an “ought” from an “is”, so on the basis of “Hume’s law” it seems “altogether inconceivable”⁶ how moral facts could exist. A moral fact would have to imply a norm. But it is “altogether inconceivable” how a fact could imply a norm. This obliterates the decisive reason for any moral realism, objectivism and cognitivism: Where there are no moral facts in reality, there are no objective moral facts either. Where there are no objective moral facts, there is nothing that can be objectively recognised.

b) *Emotivism*

As an alternative, we are offered Hume’s hypothesis that moral propositions such as “X is good or bad, right or wrong” have no cognitive content and only serve to describe our feelings. If, for example, I say that premeditated murder is wrong, this proposition renders neither a natural fact of the empirical world nor a metaphysical fact of an invisible world, but merely describes my internal experience. It describes a sense of revulsion or outrage that I feel in the face of premeditated murder. Our moral language, then, leads us into a constant deception. It pretends to describe real properties, but only our emotions are real. That is why this position is called emotivism. As these feelings are described by moral propositions, we can also call this emotivism descriptive emotivism. And because emotions are normally regarded as subjective, we can also talk about a descriptive moral subjectivism.

⁶ *Ibid.*, 469.

But here we can go one further step beyond Hume. A proposition such as “X is good or right, bad or wrong” need not have a descriptive function in spite of its descriptive form. As we have seen, there is no necessary connection between the form of a sentence and its function (cf. p. 41). A moral proposition such as “X is good or right, bad or wrong” need not be a description in spite of its descriptive form, but may also have an expressive purpose. According to this position, a proposition such as “Killing is wrong” is neither true nor false, because it can neither agree nor disagree with an internal or external fact. Rather, it has the same function as if “killing” were pronounced in a particularly indignant tone.⁷ That is how a mother teaches her child the first moral utterances, for instance, by saying: “Lying, Ugh.” This type of emotivism is also called expressive emotivism. In contrast to descriptive emotivism or subjectivism, according to the theory of expressive emotivism, moral propositions do not represent any moral facts at all – not even internal ones – but only express feelings. By so doing, they can also awaken feelings in other people and guide them to act.⁸

⁷ This view is held, for example, by Alfred Ayer (1910-1989), *Language, Truth and Logic*, Chapter 6: “If now I generalise my previous statement and say, ‘Stealing money is wrong,’ I produce a sentence which has no factual meaning – that is, expresses no proposition which can be either true or false. It is as if I had written ‘Stealing money!!’ – where the shape and thickness of the exclamation marks show, by a suitable convention, that a special sort of moral disapproval is the feeling which is being expressed. It is clear that there is nothing said here which can be true or false.” 107.

⁸ Cf. Ayer, *Language, Truth and Logic*: “It is worth mentioning that ethical terms do not serve only to express feeling. They are calculated also to arouse feeling, and so to stimulate action. ... In fact we may define the meaning of the various ethical words in terms both of the different feelings they are ordinarily taken to express, and also the different responses which they are calculated to provoke.” 109.

However, there is a counter-argument against both kinds of emotivism or subjectivism: We can obviously pass contradictory moral judgments and disagree, with good reasons, about moral propositions as well as about assertions of facts. For example, in a debate about whether or not abortion is reprehensible, a strong case can be made for either view. If moral judgments were only descriptions or expressions of feelings, they could not be contradictory, and there would be no point in looking for reasons to argue about whether or not they are right. Feelings can conflict. A man can love and hate a woman at one and the same time, and a woman can likewise love and hate a man. But in a logical sense, feelings cannot contradict themselves, as propositions can.

With regard to morally indifferent things, for example, smoking in the street, the constative form of a sentence like “Smoking is wrong” need not necessarily have a morally relevant content. Coming from an ordinary non-smoker, it may function as a simple personal expression, but if it is said by a fanatical non-smoker, it represents for him a moral fact. It also matters, then, *who* makes the statement in question. But if a morally relevant basic proposition, such as “Genocide is morally wrong”, can merely express feelings, the opposite, “Genocide is morally right”, would do the same. Since both are only expressions of emotions in disguise, there would be no point in arguing, with reasons, about which is right and which is wrong.

But when it comes to moral basic propositions of the kind I have mentioned, the emotivist conception runs counter not only to our theoretical basic convictions, which we express in descriptive and objectivist language, but also to the demand for generalisation that we attach to such moral basic propositions. If we describe an action as morally wrong, we are expressing an attitude that we expect others to share. “Murder is morally wrong” means not only that “murder is morally wrong for me”,

but also that “murder is morally wrong for anybody.” Conversely, it would seem unacceptable to us if somebody said: “It is wrong if I secretly kill my rich aunt, but it is not wrong for me, because I will profit from her death.” If it is wrong to kill my rich aunt, then it is also wrong for me.

This demand for generalisation is shown particularly clearly by the fact that moral basic propositions are socially sanctioned. If I do not observe them, I have to face a diversity of negative social sanctions, such as prison, a fine, withdrawal of social respect and other punishments. If moral basic propositions were only of a personal nature, it would be difficult to see why other people should be able to punish me for disregarding them. On the other hand, a mere expression – say, “Murder, how horrible” – is no more capable of socially sanctioned generalisation than an account of my personal feelings. Given the same facts, both the expressions and the accounts of our feelings can turn out very differently. I cannot expect other people to share my feelings. Nor can I expect my deviating expressions and emotions to be binding for other people. But neither have other people the right to punish me for my deviant feelings, or for my deviant moral expressions and emotions.

Emotivism, whether expressive or descriptive, can hardly justify the socially sanctioned demand for the generalisation of moral basic propositions, which distinguishes such basic propositions from mere exclamations and personal accounts of feelings. Nevertheless, expressive emotivism has the merit of drawing attention to the non-cognitive, expressive and action-guiding function that distinguishes moral basic propositions from merely descriptive ones. Moral basic propositions also serve to voice either commendation or condemnation, from which it is possible to derive norms as to what should be done and what should be avoided.

This gives rise to two demands that a satisfactory metaethical theory can be expected to fulfil: (a) It must take account of the cognitive and objective element in moral basic propositions and of the realistic language of morality; and (b) it must at the same time do justice to the emotive and subjective element – made up of commendation and condemnation – in moral basic propositions, so that norms can be derived from those propositions.

But the two demands seem to lead to a contradiction and to be incompatible. If moral basic propositions are cognitive and contain an objective element, they can be generalised. But then no norms can be derived from them. If moral basic propositions are emotive and contain a subjective element, it is possible to derive norms from them. But, in that case, are the moral basic propositions still able to be generalised and to be descriptive?

c) *Institutionalism*

To resolve this contradiction, we may regard basic moral propositions, such as “It is right to dress a bleeding wound, but wrong to let a person bleed to death”, as descriptions of *institutional* facts. Basic moral facts, then, do not exist in themselves, either in the physical world or in an invisible metaphysical world, as cognitivism suggests, nor are they merely subjective psychic facts, as descriptive emotivism assumes. But neither are they non-existent. Basic moral facts exist, but they are of an institutional nature. Accordingly, morality is neither something objective nor something subjective, but something essentially social, that is, an institution made by human beings. In so far as a moral institution, such as the ban on killing, exists regardless of whether or not I recognise it, it is not subjective, but objective. However, in so far as it is constituted by a language community, it is not objective in the strong sense of existing independently of a language community. It is objective only in an

intersubjective sense. It is valid among different people and, in the case of moral basic propositions, among all people, because it is supposed to set standards for all. (To avoid losing our way in a debate on exceptions, I will only deal with moral basic positions, such as the ban on killing in general, and ignore exceptions such as self-defence, killing in war, capital punishment, suicide, and passive or active euthanasia.)

The term “institutional fact” was introduced by J R Searle in his book *Speech Acts*.⁹ Institutional facts are objective and not just a matter of feeling. Nevertheless, they cannot be reduced to real facts. Examples of such institutional facts are “Mr Smith married Miss Jones; the Dodgers beat the Giants three to two in eleven innings; Green was convicted of larceny; and Congress passed the Appropriation Bill.”¹⁰ Unlike a real – that is, merely physical or psychic – fact, an institutional fact comes into being as a result of constitutive rules. These rules are structured as follows: “X stands for Y in the context of community C.” They are called constitutive because they constitute X as Y. But as they constitute X as Y in the context of the language community LC, they are also semantic rules: They give X a certain meaning Y in the context of the language community LC. A real physical action X – in the context of the language community LC – is given the meaning Y, which may be a marriage, a victory, a theft or a ratification. Institutional facts, then, are real facts, interpreted in a specific way. In institutional facts, the real world and the semantic world enter a certain association. This association, with regard to institutional facts, is of a normative nature. Therefore, Miss Jones, by marrying Mr Smith, accepts some ob-

⁹ Searle, *Speech Acts*, Chapter 2, Section 7, 50-53.

¹⁰ Searle, *ibid.*, 50-53.

ligations towards Mr Smith, as does Mr Smith towards Miss Jones by marrying her.

If we enter for Y a normative or evaluative concept – that is, the meaning of a normative or evaluative expression – such institutional facts can also contain norms or values. Among such institutional facts of a moral nature, I count the moral basic facts, for example, that it is morally right to dress a bleeding wound, but wrong to let a person bleed to death, or that genocide is morally wrong, but preventing death by starvation morally right.¹¹ A certain physical action (or omission) X – for example, dressing a wound or allowing a person to bleed to death, genocide or supplying food – is turned by constitutive rules into Y, that is, into a morally right or wrong one. As the physical action X represents a physical fact, we can say that physical facts are turned into institutional facts by constitutive rules.

But psychic facts, for example, Mr Smith's jealousy over Miss Jones, can also turn into institutional facts. Jealousy is generally attributed a negative value, being regarded as a "vice". Therefore, it is possible to derive from it the norm of not being jealous. Conversely, the lack of jealousy is generally attributed a positive value and regarded as a "virtue", so that it is possible to derive from it the norm of magnanimity. But since "virtues" and "vices" represent internal actions and are not immediately visible from outside, the social sanctions are also less obvious. A so-called inchoate offence, for example, the desire to kill, does not produce sanctions until it is articulated or until preparations for the action become known. Likewise, jealousy

¹¹ The institutional understanding has been applied to the law by Donald Neil MacCormick and Ota Weinberger, cf. *Institutional Theory*, 1986, esp. Introduction, 1-30, and Chapter 2, "The Law as an Institutional Fact", 49-76. I expand the institutional understanding to metaethics. Further: Ferber, *Moral Judgments*, 372-392.

and magnanimity, envy and lack of envy, etc., are not assigned either a negative or a positive value until they become visible. But then internal facts or facts of consciousness also provoke external reactions from other people.

Semantic facts, too, can turn into institutional facts. By ignoring an individual murder and formulating the abstract proposition “Murder is morally wrong”, we turn a semantic proposition into an institutional one. We no longer refer to an individual murder, but to the meaning of the statement “Murder is morally wrong”, so that the meaning of the proposition itself becomes the referent. Likewise, the first article of the German Basic Law, “Human dignity is inviolable”, is an institutional fact of a semantic nature. It does not refer to the inviolable dignity of Mr Smith or Miss Jones, but to the inviolable dignity of the human being in an abstract sense. From this institutional fact, it is possible to derive the norm that the dignity of the human being should not be violated. It is possible to do so because the proposition itself contains an in-built norm. “Human dignity is inviolable” also means: “Human dignity ought not to be violated.” The word “is” in normative contexts has a normative function despite its indicative form.¹²

Thus, facts from three different worlds – the physical, the psychic and the semantic – can become institutional facts if values and norms are built into them. We can add to facts of all three kinds a normative, evaluative and institutional interpretation. If these values and norms are of a moral nature, the institutional facts become moral ones. But if basic moral facts are of an institutional nature, we can assert the existence of moral facts without contravening “Hume’s law” of the impossibility of

¹² For the normative “is”, cf. Ferber, “*Normatives ‘ist’ und konstatives ‘soll’*”, 185-199. No English translation.

deriving an “ought” from an “is”. For such institutional facts contain values and norms from the outset.

On the one hand, the institutional understanding of moral facts can explain the extent to which moral facts are objective and at the same time generalisable. As facts, they are objective and valid for everybody, albeit only in the weak sense of objective, that is, as intersubjective or objective in the context of the language community LC. Today this comprises, for moral basic facts, the official language of almost all states and is codified in the Convention of Human Rights. There is almost no state and hardly an individual who would dare to claim officially that genocide or murder (with the exceptions I have mentioned) is morally permitted. The language community LC here embraces almost the entire community of human beings.

On the other hand, this interpretation can also explain the extent to which moral basic facts contain a subjective element and moral norms can be derived from them. They contain a subjective element in so far as they are made by a specific language community by means of constitutive rules; and it is possible to derive norms from them, since they contain norms from the outset. Thus, from the institutional fact that it is morally right to dress a bleeding wound, but wrong to let the person bleed to death, we can derive the norm that a bleeding wound should be dressed and the person should not be allowed to bleed to death.

The institutional understanding of moral facts also explains how far moral facts are made and sanctioned by human beings. Institutional facts are obviously made by human beings. From the outset, they contain norms that entail sanctions if they are not followed. The sanctioning is particularly noticeable where moral facts are institutionalised by law. They are backed by the state as the sole legitimate bearer of physical violence. The ban on killing, with the exceptions I have mentioned, for example, is laid down by law. Disregard of it results in a graduated range

of fixed punishments – depending on whether it is a question of murder, manslaughter, death caused by design or death caused by negligence – such as a fine, prison and, in some states, even the death penalty.

But not all institutional facts of a moral nature are legally sanctioned. Nor are all institutional facts, which are sanctioned by law, moral. For example, the law authorising the murder of mentally handicapped and sick people, brought into force by Hitler through a secret directive, was certainly not moral. At the time of National Socialism, those who perpetrated something of the kind did not find themselves in conflict with the law. But they were surely despised by most of their fellow humans and, after the fall of the Third Reich, wanted by the authorities. Some facts of a basic moral nature that are not, or are only moderately, sanctioned by the state – for example, the expectation that we should behave towards our fellow humans in a friendly and helpful way – are sanctioned by human beings by praise for friendly and helpful behaviour and censure for unfriendly and unhelpful behaviour. Here, the sanctions are not applied by the state, but they are nevertheless of a social nature, consisting in certain positive or negative responses of other people, such as praise or blame, recognition or rejection, support or obstruction.

Finally, the institutional view of moral facts shows the extent of the demands that these facts can make on me beyond the pursuit of my own interests. The demands of morality as a social institution do not always correspond to what I want. Sometimes I would rather not be “moral”. Nor is morality as a social institution what God or a metaphysical or supernatural authority expects from me. To believe that, we would have to be able to assume that such an authority, or God, exists. Morality is primarily what *a* – or *the* – community of human beings demands from me. Since it is not my will, but the will of others, that is

behind morality as a social institution, it can demand that I perform some actions that are not in my interest, but in the interest of others.

The institution of the ban on killing demands that I obey it even if it were not in my interest to do so, for example, if I could gain an advantage by killing somebody. But it is obviously in the interest of another person not to be murdered. Conversely, I have to dress a bleeding wound, even if the life of the person concerned means nothing to me. But obviously it is in the interest of the person bleeding that the wound be dressed. The situation is similar with regard to other moral basic propositions, for example, the proposition that one should not commit incest, steal or cheat. Therefore, we may say that the institution of morality, at least as far as the basic demands are concerned, is what *a* – or *the* – community of human beings expects from me, regardless of whether or not I as an individual profit from it. This should not be misunderstood to mean that morality forbids us as a matter of principle to pursue our own interests. Rather, it means that morality restricts the pursuit of our own interests by demanding that we also consider the interests of others. That is why morality can make demands on me even beyond the pursuit of my own interests.

As a rule, our basic moral feelings are embedded in this institutional framework. They are not simply subjective, but usually well socialised; that is, they internalise the constitutive rules and thereby the will of a language community. We are repelled by a premeditated murder because that is how we were trained to feel by our parents, teachers and fellow humans, and because other people feel the same. If we had been socialised three thousand, or perhaps just three hundred, years ago, many or even most of us would be outraged by the murder of a member of our own family, tribe or nation, but perhaps indifferent

to, or even satisfied by, the murder of a member of another family, tribe or nation.

The same applies to discrimination against people because of their race, gender, disability or sexual orientation. We regard discrimination of race, gender, disability or sexual orientation as reprehensible because that is how we were socialised and how we internalised the corresponding constitutive rules. If we had lived in another epoch, many or even most of us would have seen nothing reprehensible in slights to people of a different colour, to women, to disabled or gay people. Once a language community has fixed these institutional facts, we can apply descriptive propositions to them and in the process recognise that murder, torture or racial or gender discrimination is reprehensible, as is discrimination because of disability or sexual orientation. But even then, we do not recognise facts that exist as such, but rather institutional facts made by human beings.

As a rule, then, our basic moral feelings are not only subjective, but neither are our basic moral insights strictly objective. Rather, moral emotions and cognitions are inserted into the institutional framework of a community – a framework that is both objective and subjective. It is objective in so far as it exists intersubjectively and regardless of whether or not I recognise it. It is subjective in so far as it has come into being through constitutive rules, from which more norms can be derived. But since the institution of basic morality has largely solidified, it has the appearance of something objectively given. It is so deeply rooted that its human and social origin has been forgotten. This appearance of objectivity is in fact necessary for the moral basic rules to be universally recognised and, to a certain degree, to be actually effective.

3. Normative Ethics

a) *The Concept of the Good as the Foundation of Morality*

We also, however, consider it morally better, or more correct, not to discriminate against people because of their race or gender disability or sexual orientation. And we relied even more on an intuitive understanding of what is morally right or wrong when we described dressing a bleeding wound as morally right. Just as we all somehow know the meaning of “is” and “is not”, we also know the meaning of morally “good” and “bad” or “right” and “wrong”. It is this unconscious knowledge that must be made conscious. Since the concept of the good also has an action-guiding function, we may hope that the answer to the question “What is morally good?” will also help us find an answer to the question “Why should we be morally good?” or “Why should we act in the morally right way?”

To obtain an answer to that question by clarifying the concept of the good, it is not enough to give a cause as to why we should act well, that is, the right way. Such a cause could be an inner trigger or motive. Such a motive could be, for example, fear or hope, in particular, hope of a reward for a good deed, or fear of punishment for a bad deed. We would, of course, prefer to be rewarded rather than punished. If we were always rewarded for moral behaviour and always punished for immoral behaviour, moral behaviour would be synonymous with what we want in our own interest. The kind of behaviour which pursues our own interests cleverly is also called wise in the sense of prudent.

Obviously, it is prudent to behave so that we are rewarded and imprudent to behave so that we are punished. Often moral behaviour is prudent, and goodness the best policy in an instrumental sense. But this is not always the case. Occasionally, moral behaviour is imprudent. At times, our goodness is the

most formidable weapon of our enemies. Often, our goodness is not rewarded but exploited and punished. Again and again, it is precisely wickedness that gets the reward. We all know that the “baddies” are sometimes rewarded and the “goodies” sometimes punished, even though we cannot put it quite like Shakespeare: “Some rise by sin and some by virtue fall.”¹³

Therefore, in reply to “Why should we be good and act in the right way?” we cannot invoke a personal motive – for example, our own interest – and say: “We should be good, that is, act in the right way, because it is in our own interest.” Rather, we look for a reason to be good or to act in the right way. But this reason should be independent of any concern over rewards and punishments.

Morality as a social institution is an institution with social sanctions. Otherwise it would be as good as ineffectual. But rewards and punishments do not account for the validity of morality. As much as social sanctions are the motives for our moral behaviour, as little are they the reason why we should behave morally. Beyond a certain level of social evolution, we have internalised the institution of morality so deeply that it commands us to be moral even if we are neither rewarded for moral behaviour nor punished for immoral behaviour.

Moral behaviour resembles a heavy steamer continuing its voyage long after the engines of self-interest are cut. Only then do we believe that we are acting in a truly moral way, once we have abandoned our self-interest and ceased to expect anything in return.

The definition of the good tries to give a reason why we should be moral and act morally, even without a reward. Such a reason can become an indirect cause guiding our actions. The

¹³ *Measure for Measure*, II, i, Escalus.

reason for morality becomes the cause if we adopt it and allow it to determine our actions. It is this reason that we seek in the concept of the moral good.

b) *The Good as Utility*

Therefore, in what follows, I will take the concept of the good and the bad as my starting point, ignoring such motives as rewards and punishments. From our institutionalist position, I will restrict the question “What is morally good?” to “Which institutional facts are morally good?” To find an answer, we will first consider the consequences of the morally good and the morally bad, or evil. For this, we start with the following elucidation of Plato: “The corrupting and destroying element is the evil, and the saving and improving element the good.”¹⁴

Here Plato does not yet distinguish between what is good in itself and what is good only as a means to an end. He believes that all that is good saves and improves, and all that is bad corrupts and destroys. Therefore, the morally good also saves and improves, while the morally bad also corrupts and destroys. To put it more simply: The morally good is *useful*, the morally bad *harmful*. It is morally wrong not to dress a bleeding wound, because that harms the person who would otherwise bleed to death. But it is morally right to dress the wound, because that benefits the bleeding person. Genocide is even more morally wrong, because it leads to the destruction of a whole nation. But supplying food is morally right, because it preserves life. Morality, then, is generally useful and life advancing, immorality harmful and life obstructing. Indeed, morality as a social institution would have found it hard to establish and consolidate itself if, in contrast to immorality, it were not beneficial, at least, for

¹⁴ Plato, R., Book 10, 608e. Transl. Jowett.

the life of a community, though not always for the life of each individual.

The definition “The good is life enhancing and the bad life obstructing” provides the following answer to the question of why we should be good and act in the right way: Morally right actions are useful and life enhancing; morally wrong actions are harmful and life obstructing. Therefore, the reason for morality is an extramoral value, namely, usefulness or the ability to enhance life. Accordingly, those institutional facts that are useful and life enhancing are morally right, and those that are harmful and life obstructing are morally wrong.

We do not, however, want simply to live. We want to live happily, and we do not want to live unhappily. If we were asked to describe the difference between a happy life and an unhappy life, we might answer that a happy life is full of pleasure and an unhappy life full of pain. According to this view, morality is not only life enhancing, and immorality is not only life obstructing, but morality actually leads to a happy or pleasurable life and immorality to an unhappy or painful one.

This sounds strange because the term “moral” today has acquired a secondary meaning as the enemy of pleasure. Nevertheless, it is a view that has been asserted time and again from Greek antiquity to this day. It is called “eudemonism” (from *eudaimonia*: happiness) and “hedonism” (from *hēdoné*: pleasure). An influential newer version of this theory is found in Mill’s *Utilitarianism* (1863), where both eudemonism and hedonism are bracketed under the concept of “utilitarianism” (from *utilitas*: usefulness):

The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended

pleasure, and the absence of pain; by unhappiness, pain, and the privation of pleasure.¹⁵

As Mill has actions in mind, we can also talk about “act utilitarianism”: A single action – for example, dressing the wound of a person about to bleed to death – is morally right if it promotes happiness or pleasure; it is wrong if it causes the opposite of happiness or pleasure, that is, unhappiness or pain. Utilitarianism can further be extended to include rules. Therefore, we can also talk about “rule utilitarianism”: A rule – for example, “Thou shalt not kill” – is morally right if it generates happiness or pleasure and morally wrong if it causes misfortune or suffering.

Among these rules, we may include not only regulative moral precepts, such as “Thou shalt not kill”, but also constitutive moral rules, such as “A certain action X is regarded as morally right in the context of the language community LC.” Since such rules are absorbed in institutional facts, we may, within our institutionalist position, expand rule utilitarianism into institutional utilitarianism. Institutional utilitarianism, like rule utilitarianism but unlike act utilitarianism, can fall back on the moral tradition of the basic rules and institutions discussed so far, the consequences of which are to a large extent known to us from experience. This is simpler than having to consider afresh before each individual action what the consequences would be. The existing moral institutions were established in a lengthy process, and we may assume that they are not completely wrong. But that is no reason why all existing institutional facts should be morally right.

Institutional utilitarianism supplies us with a criterion for determining when institutional facts are morally right or wrong.

¹⁵ Mill, *Utilitarianism*, Chapter 2, 9-10.

Institutional facts are morally right if they generate happiness, pleasure or utility and morally wrong if they generate unhappiness, pain or harm. Thus, for example, the institutional fact that “human dignity is inviolable” is morally right, because it generates happiness, pleasure or utility, but the opposite is morally wrong, because it generates unhappiness, pain or damage. Further, the institutional fact that “magnanimity is good, envy is wrong” is morally right, as a rule, because magnanimity leads to happiness, pleasure or utility and envy to unhappiness, pain and harm – not only for the object of the envy, but also for the envious subject. As has been said, envy is the only vice that has no pleasure in it.

But utilitarianism does not claim that those actions or institutions alone are good that enhance my happiness, my pleasure and my utility. It does not claim that good is what benefits me. It also thinks of others. Nor does modern utilitarianism distinguish the different social levels. It does not say that good is what benefits my class only, and at the most, indirectly, the others. Modern utilitarianism is democratic. It believes that the good is what maximises the happiness, pleasure or utility of as many human beings as possible. In that process, according to a formula by Jeremy Bentham (1748-1832), “everybody [is] to count for one, and nobody for more than one”, because everybody’s basic interests are equal to everybody else’s. The good then, in a catchy phrase, is what brings “the greatest possible happiness to the greatest possible number of people”. The good is what maximises the happiness, pleasure or utility of the greatest possible number of people.

By this criterion, I can also explain why discrimination based on race, gender, sexual orientation or disability is morally wrong and non-discrimination right. Imagine the great pleasure food can give to a hungry person and the appalling pain experienced by somebody starving to death. We feel a similar pain if,

for example, we are rejected when looking for a job or an apartment, or if we are disadvantaged or unfairly treated in some other way because of our race, gender, sexual orientation and disability.

Institutions and institutional facts that discriminate are morally wrong also because they do not contribute to the greatest happiness, or greatest pleasure, of the greatest number and may actually cause the majority of people unhappiness or pain. Thus, the criterion of utilitarianism reveals not only when institutions are good, but also when they are fair. Fairness here means equal treatment of people of different race, gender, ability and sexual orientation. Fairness is what maximises the happiness, pleasure or utility of the majority or, in the ideal case, of all concerned. With happiness, pleasure or utility as its purpose, morality as understood by utilitarianism is built on an extramoral value.

If we ask for proof of the thesis that the good is happiness, pleasure or utility, we receive the following answer: Everybody strives for happiness, pleasure and utility. The seventeen-year-old Cécile in Françoise Sagan's (1935-2004) novel *Bonjour Tristesse* confesses candidly: "My love of pleasure seems to be the only consistent side of my character." It is because everybody strives to be happy that everybody regards happiness, pleasure and utility as good. That is why eudemonism, hedonism and utilitarianism, in both theory and practice, are recognised by everybody. Mill writes:

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. No reason can be given why the general happiness is desirable, except that each person, so far as he believes it to be attainable, desires his own happiness. This, however, being a fact, we have not only all the proof which the case admits of, but all which it is possible to require, that happiness is a good: that each person's happiness is a

good to that person, and the general happiness, therefore, a good to the aggregate of all persons.¹⁶

“All the proof which the case admits of” is no proof at all in the strict sense. It does not prove the thesis by deriving it as a conclusion from certain premises. Rather, it confers on the thesis itself the status of a first premise of an axiomatic character: Just as it is evident that an object is visible if we actually see it, it is equally evident that we all want our own happiness, our own pleasure and our own utility. What we want we regard as good. Thus, we all regard our own happiness, our own pleasure and our own utility as good. To that extent, Mill shares the classic definition of good as “that at which all things aim”. But from the axiom of utilitarianism, he draws the conclusion that the purpose of the utilitarian theory is also right in that everybody regards universal happiness as desirable, that is, as good, and that everybody desires universal happiness.

As we have seen, evidence is only a *prima facie* criterion for the truth of an axiom. Therefore, it is also only a *prima facie* argument in favour of the axiom of utilitarianism. On reflection, some objections arise. They concern both the axiom of utilitarianism and the validity of the conclusion drawn from it.

α) Let us assume that this axiom is evident and we all want our own happiness, pleasure or utility and therefore regard our own happiness, pleasure and utility as good. But it does not follow that we all also want the happiness, pleasure or utility of others and also regard the happiness, pleasure or utility of others as good. And it follows even less that we all want, or regard as good, the greatest possible happiness, pleasure and utility of the greatest possible number of other people. But if moral actions and institutions are supposed to promote “the greatest

¹⁶ Mill, *Utilitarianism*, Chapter 4, 52-53.

happiness of the greatest number”, they will not only serve my happiness, but also the happiness of others. As we have seen, moral institutions can demand actions from me that go beyond my own interests. It is not immediately obvious how, by wanting my own happiness and regarding it as good, I also want and regard as good the “happiness of the greatest number”.

β) The conclusion of utilitarianism may be false and its axiom, at the same time, true. The axiom that we all want – and regard as good – our own happiness, pleasure or utility seems evident. But opinions differ as to what is our own happiness, pleasure or utility. If we replaced the general terms “happiness”, “pleasure” or “utility” with the concrete ideas that people mean by them, we would end up with very different things: Those dying of thirst long for a glass of water; those who are very hungry long for a slice of bread; those who are freezing long for a warm coat; those without a home long for a roof over their heads; those confined to a dark prison long for sunshine and freedom; those who are lonely long for a human companion; those surrounded by too many people long for solitude. But it is also possible to want to give pleasure to others. The axiom of utilitarianism seems so evident only because it does not say exactly what people want. If we get to the heart of the matter, it says very little or nothing.

Another way of putting it would be that we all desire something desirable. But this change of wording shows that the axiom of utilitarianism is not an empirical hypothesis that can be falsified through experience. It is a conceptual thesis that is true only on the basis of the meaning of the words used. Based on the meaning of the words “desire” and “desirable”, it is true that we all desire desirable things. But as true as this axiom is, it is also trivial. And as soon as we formulate it as an empirical hypothesis, it actually becomes false. We do not desire happiness, pleasure or utility directly: We desire good things; and we

obviously desire different good things according to our needs. Only once we abstract from the individual goods, and ask ourselves what we desire to achieve through them, can we say in hindsight that we desire happiness, pleasure or utility. Therefore, the axiom of utilitarianism only seems to be directly evident. In fact, it rests on acquired evidence. But this acquired evidence can also be called into question through reflection. Seen close up, it is actually false.

γ) So we see that the consistency of the utilitarian theory does not follow from the axiom and that the axiom of the utilitarian theory is not directly evident. But even if the axiom of utilitarianism were directly evident and if the purpose of the utilitarian theory followed from it, we could ask: Is my happiness, pleasure or utility morally good? Is even the greatest possible happiness, pleasure or utility of the greatest possible number morally good?

I must answer this question in the negative, just as I already answered the question of whether everything that everybody aims at is morally good. The utilitarian definition of good – like the classic definition of good as “that at which all things aim” – does not distinguish between the morally good and the extramorally good. Even if happiness, pleasure and utility are good, this does not mean that they are morally good from the outset. This is true of both my happiness and the happiness of others. For it is also possible to aim at giving others a happiness which is regarded as immoral, for example, by aiding and abetting murder. Likewise, our own happiness is by no means always morally good – say, if it rests on the unhappiness of others.

As we have seen, the morally good is not something that we can perceive through an external experience, as we can, for example, colours or sounds. Nor is it something that we can perceive directly within us through an internal experience, as we can, for example, pleasure or pain. What is morally good is no

real predicate and has no real existence. It is determined through constitutive semantic rules and therefore has only a semantic existence. The semantic rules concerned spring from the basic will of a community and are absorbed by the institutional facts of that community. After the event, however, they can be internalised by the individual. Thus, the concept of the morally good obtains a dimension that is not exhausted by the eudemonistic, hedonistic and utilitarian definitions. On the other hand, it is a mistake of eudemonism, hedonism and Mill's utilitarianism to explain the concept of the morally good through other concepts, such as happiness, pleasure or utility, which do not automatically contain this moral dimension. We can call this mistake, with Moore, the "naturalistic", and in the language of the previous chapter, the "realistic fallacy".¹⁷ It consists in the immediate transition from a real predicate – happiness, pleasure or utility – to a semantic one – that of the morally good.

This "realistic" fallacy is the ontological counterpart of the logical fallacy of inferring a normative statement from a constative one (cf. p. 179). As little as I may infer a normative statement from a constative one, as little may I infer a semantic predicate from a real one. Not only does the content of a normative statement go beyond that of a constative one, but the content of a semantic predicate also exceeds that of a real one. We need a rule that tells us *how far* a real fact that promotes happiness, pleasure or utility is morally good. In contrast, any rules that aimed at happiness, pleasure or utility would turn the comparatively absolute character of moral obligation into something

¹⁷ Moore, PE, Chapter 1, § 10: "But far too many philosophers have thought that when they named those other properties they were actually defining good; that these properties, in fact, were simply not 'other,' but absolutely and entirely the same with goodness. This view I propose to call the 'naturalistic fallacy' and of it I shall now endeavour to dispose."

relative by making it dependent on an extramoral condition, that is, the condition of the happiness, pleasure or utility of the greatest possible number. Moore was right to place Joseph Butler's dictum as a motto at the head of his *Principia Ethica*: "Everything is what it is and not another thing" (cf. p. 88). Thus, the morally good, too, is what it is and no other "thing", for example, happiness, pleasure or utility.

But even if happiness, pleasure or utility were the condition of the rightness of our moral institutions, it would be difficult to define positively what this condition may be. Happiness can mean different things to different people, so that it is not easy to compare the happiness of different people. This is all the more true if we leave it to people to decide for themselves what to see, and look for, as their happiness. Different people look for different things when they look for happiness. Often they do not really know what they are looking for. They then resemble drunks looking for their houses with the vague idea that they have one.¹⁸ But sometimes they do not look for a house, but for a castle in the air. Happiness, to borrow a definition used by Kant in another context, is an "ideal of imagination".¹⁹ The "greatest happiness of the greatest number" is even more an "ideal of imagination". Happiness cannot be pursued directly but completes an activity like "an end which supervenes as beauty does on those in the flower of their age."²⁰ Real, profound happiness, however, as far as accessible, may be, in Spengler's words, "presence without thought".

¹⁸ This saying is attributed to Voltaire.

¹⁹ Kant, Groundwork, Section 2, 418. Transl. Gregor.

²⁰ Aristotle, NE, 1174b33. Book 10, chapter 4, Transl. Ferber.

c) *The Good as a Rule*

Therefore, it is advisable to look for a more correct answer to the question “What is morally good?” We already used this answer in an undeveloped form, when we characterised moral basic propositions as generalisable. But utilitarianism also assumes this answer, likewise in an undeveloped form, in so far as it sees the good in “the greatest happiness of the greatest number”, formulating the rule that every human being must be counted as one human being and no human being as more than one human being. And we can find this answer, again in an undeveloped form, in the Sermon on the Mount, where Jesus commands: “Therefore all things whatever ye would that men should do to you, do ye even so to them.”²¹

This rule is called the *Golden Rule*. It can be phrased positively and negatively. The wording in the Sermon on the Mount is positive. The negative wording that has become established is “Don’t do to others what you don’t want others to do to you.” The Golden Rule was developed by Kant. In his view, the morally good is not happiness, nor the “greatest happiness of the greatest number”. We are told what is morally good by rules. Among the relevant rules there is a main one. It is the rule of generalisation. Kant calls it the “categorical imperative”. He discusses it in many different ways in his *Groundwork of the Metaphysics of Morals* (1783), but it remains a single command:

²¹ Matthew, VII, 12. Transl. King James Bible. Cf. Luke 6, 31. In a negative wording, the Golden Rule is already found in Tobit, 4, 15 of the Old Testament: “What you hate, do not do to any one” and in Confucius, Lunyu 15, 14: “Tze-kung asked: saying, ‘Is there one word which may serve as a rule of practice for all one’s life?’ The Master said, ‘Is not reciprocity such a word? What you do not want done to yourself, do not do to others.’” Transl. Arthur Waley, *The Analects*, London 2000.

There is therefore but one categorical imperative, namely, this: Act only on that maxim whereby thou canst at the same time will that it should become a universal law.²²

Kant calls this imperative “categorical” because, in contrast to a hypothetical imperative, it is not relative or conditional, but absolute and unconditional. It is unconditional because it is not tied to any conditions. In particular, it is not tied to the condition of happiness, pleasure or utility. It is free of the consequences that could result for me and others from following it. We can foresee these consequences to a great extent, but not invariably. Nevertheless, we must abide by this rule and simply wait for what may come: “One must be good and expect the rest.”²³

The categorical imperative tells us that any actions are good only if they are carried out according to maxims that can be generalised. A maxim is a subjective principle. Therefore, the categorical imperative demands that we act only on those subjective principles that are generalisable. A subjective principle is generalisable if we all are able to adopt it without willing something that we cannot will. Subjective principles, which everybody is able to adopt, can be right at the intersubjective level. Therefore, the categorical imperative commands us to act only on those subjective principles that can be right at the intersubjective level.

Within our institutionalist position, we may also expand this rule of the generalisability of our actions to institutional facts: The only morally right institutional facts are those that all other people can adopt as far as possible. Thus, the criterion of gener-

²² Kant, *Groundwork*, Section 2, 491. Transl. Gregor.

²³ Kant, *Observations*, 19. Transl. Frierson/Guyer.

alisability, which was developed above for moral basic propositions, also becomes the reason why they are morally right.

By this criterion, we can explain why, for example, the institutional fact of dressing a bleeding wound is good, while letting a person bleed to death is bad, or saving people from starvation is good, but genocide bad. These institutional facts are clearly generalisable: I can will a law that commands us to dress a bleeding wound, but forbids us to let a person bleed to death, or a law that orders us to save people from starving to death, but forbids genocide. However, I cannot will a general law that forbids us to dress a bleeding wound, but allows us to let a person bleed to death. Even less can I will a law that forbids saving people from starvation, but permits genocide. If I willed such a law, I would implicitly will something that I cannot explicitly will. After all, I myself could one day be in danger of bleeding or starving to death, or indeed of being murdered. But generally, I can no more will to bleed or starve to death than I can will to be murdered. Given such a law, I would find myself in a conflict of my will.

By the same criterion, we can also explain why discrimination against people because of their race, gender, sexual orientation and disability is morally wrong and non-discrimination is right. Non-discrimination is right because this institutional fact is generalisable, while discrimination is not. Let us assume that there is a general law that allows discrimination against a race, gender, sexual orientation and disability. This would mean that, just as I would be allowed to discriminate against other people because of their race, gender, sexual orientation and disability, other people would be allowed to discriminate against me because of my race, gender, sexual orientation or disability. But if I willed a law that allowed me to be discriminated against, I would will something that I cannot will. I cannot will to be discriminated against.

As little as I can deny the metalogical axioms of identity and contradiction, as little can I will the “moral axioms” to be abolished. I cannot deny the metalogical axioms of identity and contradiction because, in order to deny them, I have to affirm them (cf. p. 92). But in order to abolish “moral axioms”, such as the ban on killing or discrimination, I would have to will something that I cannot will. In the first instance, I would be caught up in a contradiction of my language and theoretical reason; in the second instance, in a conflict of my practical reason or, to put it differently, in a conflict of my reasonable will.

This position, which claims to be able to tell what is right and wrong on the basis of the generalisation rule, is not called utilitarian. It is called “deontological” (from *to déon*: obligation) because these commands impose an obligation on us, regardless of any useful consequences they may have. Likewise, the main command, the obligation to generalise, exists regardless of any useful consequences it may entail. In fact, there is no necessary connection between my moral or immoral behaviour and the outcome. That is why there is also no necessary connection between the concepts of morality and happiness or between the concepts of immorality and unhappiness. Moral behaviour often, but not necessarily, leads to happiness; immoral behaviour often, but not necessarily, leads to unhappiness. Happiness can be an addition to moral behaviour and “blossom unexpectedly”. But it need not be so. Conversely, unhappiness can be an addition to immoral behaviour. A merit of the deontological position is that, unlike the hedonistic, eudemonistic or utilitarian positions, it sees no conceptual link between morality and happiness, but instead makes this connection dependent on the way of the world. Another merit of the deontological position is that it does not commit the “naturalistic” or “realistic fallacy”. It explains the concept of the morally good not by the real conse-

quences that can be experienced internally or externally, but by a rule of our will.

However, the generalisation rule – regardless of Kant’s different wordings and the metaphysical edifice of ideas in which he embedded it – is also open to serious objections. Three remarks about the generalisation rule concerning institutional facts must suffice here:

α) First, the generalisation rule, which determines what is good or bad, starts with a prior understanding of what is morally good or bad. Not every moral principle that can be generalised is morally good by definition. What if we wanted a general law that obliged everybody to get up early in the morning? Would that make getting up early a morally relevant action?

The generalisation rule alone, then, cannot give us the criterion to determine whether some institutional facts are morally right or wrong. Morally indifferent institutional facts could also be transformed into duties for everybody. To be able to serve as a criterion, the generalisation rule needs certain initial guidelines about what is *prima facie* morally good and what is wrong. After all, it is this rule that makes the criterion of moral institutional facts the foundation of morality.

It is a particular characteristic of these guidelines that the only actions to be institutionalised are those that have direct or indirect consequences for the vital interests of other people. Morality, as defined here, is first of all a social morality. Getting up early would be morally relevant only if the vital interests of other people were directly or indirectly affected by it. They would be affected if they suffered an undeserved disadvantage owing to my getting up late, for example, if I were a duty doctor who arrived late at the scene of an accident, violating the moral command to save life.

So it is only if we are guided by a prior understanding of which institutional facts are moral that the generalisation rule

provides a criterion as to when an institutional fact is morally right. This shows us that the generalisation rule cannot determine on its own which facts are moral, but can do so only in connection with a prior understanding of what is morally good. This involves taking the consequences of the good and the bad into account, and is basically utilitarian. Thus, getting up early is a morally relevant action if it has either useful or harmful consequences for the vital interests of other people.

β) Second, the generalisation rule assumes that, just as my actions can have either a positive or a negative effect on the vital interests of other people, the actions of others must be able to affect my vital interests, either positively or negatively. Usually, none of us is so far remote from other people as to hope or fear nothing from them. If we had nothing to hope or fear from other people, we would be able to generalise our subjective principles without a conflict of our will. Thus, the generalisation rule does not apply independently of all experience; it is only valid under specific conditions, in particular, the condition of a certain uniformity of people and their circumstances. But these conditions are such that we can regard them as largely fulfilled by most institutional facts of a moral nature. None of us is protected from others to such a degree that we could not be killed, robbed, defrauded, abused or subjected to other kinds of disadvantages.

γ) Finally, the generalisation rule does not exist in isolation from all consequences either, but it takes account of the consequences that certain institutions of a moral nature may have. The generalisation rule is predicated on the fact that I cannot want the consequences of its abolition. I cannot want an institution that allows killing without restraint, because I myself do not want to be killed. However, the generalisation rule considers not only the consequences that its abolition could have for me, but also the consequences that it could have for others. It

abstracts from my vital interests and takes account of the vital interests of all others by equating mine with theirs.

The categorical imperative, then, is in principle a hypothetical imperative that makes the vital interests of all people the condition of morality. It is a general hypothetical imperative, which could be worded as follows: Act solely on that principle which considers not only your vital interests, but those of *all* other people. And this means that as much as Mill and Kant may differ in their reasoning, they agree on the aim. Mill himself put this as follows:

To give any meaning to Kant's principle, the sense put upon it must be, that we ought to shape our conduct by a rule which all rational beings might adopt with benefit to their collective interest.²⁴

Thus, neither the prerequisites nor the aims of the utilitarian and the deontological positions are as far apart as they seem to be. Both are guided by a prior understanding of the good that takes the consequences into account. Both aim at useful consequences, not only for me, but also for all other people.

Nevertheless, the deontological explanation of morality is clearly preferable to the utilitarian, for it makes it clear that the morally good is not necessarily connected with the concept of happiness, pleasure or utility, but presupposes a specific will. The specific will, in this interpretation, is the will of a human community. Constitutive moral rules and the institutional facts corresponding to them are facts for a human community. One aspect of an institutional fact is that it applies to every member of a specific language community LC. The criterion of morality suggested here indicates only that the language community LC must not be restricted to a specific group LC₁, LC₂, LC₃, etc. –

²⁴ Mill, *Utilitarianism*, Chapter 5, 78-79.

for example, rich, white, men, etc. – but should, as far as possible, include all people.

With these reservations, we can accept the generalisation rule as a rule of thumb and define it as follows: The morally right, or good, institutional facts are those that affect the vital interests of other people and that can in principle be adopted by all of us without wanting anything that we cannot want. Conversely, the morally wrong, or bad, institutional facts are those that affect the vital interests of other people and that cannot be adopted by all of us because in so doing we would have to want something that we cannot want.

Therefore metaethical institutionalism does not lead into metaethical moral relativism. Moral relativism we may define with Plato in the following way: “Whatever in any city is regarded as just and admirable, *is* just and admirable, in that city and for so long as that convention maintains itself.”²⁵ When, for example, a city regards slavery as just and admirable, then it *is* just and admirable in this city.

Metaethical institutionalism marks with the generalisation rule the dividing line between moral and immoral institutional facts. So it is an immoral institutional fact that some human beings are treated as slaves because it cannot in principle be adopted by all of us without wanting something that we cannot want.

In this way, the proposed institutionalism tries to give also a synthesis of morality and legality. With the generalisation rule, institutionalism gives a moral foundation of legal obligations. Moral obligations which are not embedded by laws of the state or the community very often – if not always – remain inefficient. This can be observed when, as in war, laws of the state or

²⁵ Tht. 167c. Transl. Levett.

the community, for example, not to steal or not to treat other people as slaves, are out of force. On the other hand, legal obligations without moral foundation are not yet moral. It is not yet a moral achievement not to steal only because the state or the community that I live in has forbidden stealing and I fear the sanctions. It becomes a moral act only when I do not steal, because I think it is right not to steal although I do not have to fear any sanctions.

The synthesis of morality and legality has been called by Hegel *Sittlichkeit* in distinction to *morality*.²⁶ Thus, institutionalism also tries to redefine what Hegel called *Sittlichkeit*, that is, customary or institutional morality in distinction to (personal) morality. *Sittlichkeit* is the deliberate and free acceptance and observation of the prevailing institutions of the community that I live in in so far as they are legitimated by the generalisation rule. Morality is the morality of my personal consciousness and may go beyond customary or institutional morality.

So it is a legal obligation not to steal money from my demented mother or father but to support them financially if necessary. It is a still widely accepted institutional moral obligation of *Sittlichkeit* to give her or him a gift at Christmas. But it is a personal moral obligation to pay them a visit every week.

The vital interests of other people can also be understood in a narrower and in a wider sense. In the narrower sense, they refer to a bare, undamaged life; in the wider sense, to a free, equal and happy one. Thus, the ban on killing, for example, is morally right because it concerns the interest of other people in bare life and because it can be adopted by all without a conflict of the will. The ban on discrimination is morally right because it

²⁶ Hegel, *Philosophy of Right*, Part 3, § 142.

touches on other people's interest in a free and equal life and because, again, we can all adopt it without conflict.

However, the generalisation rule is only a rule of thumb because its application requires "power of judgment sharpened by experience", as Kant says about moral laws in general.²⁷ The power of judgment is a faculty of common sense. One of its tasks is to determine which individual cases fall under a given rule. Aristotle illustrates this by the following extramoral example:

This is why some who do not have knowledge, and especially those who have experience, are more practical than others who have knowledge; for if a man knew that light meats are digestible and wholesome, but did not know which sorts of meat are light, he would not be promoting his health, but the man who knows that chicken is wholesome is more likely to be promoting his health.²⁸

It is the power of judgment that decides which meat benefits our health or at least does not upset our digestion.

Similarly, it could be said that if we know that discrimination should be avoided, but perhaps not when a person of another colour, gender, sexual orientation or disability will feel discriminated against, we shall not achieve a great deal with this general knowledge. We are more likely to achieve something if we also know when a person of another colour or gender will feel discriminated against or hurt. Some people feel humiliated by the mere mention that they are gay, while others are proud of it.

A further task of the power of judgment is to decide how to apply the rule to the individual case. In so doing, it must follow

²⁷ Kant, *Groundwork*, Foreword, 389. Transl. Gregor.

²⁸ Aristotle, *EN*, Book 6, Chapter 7. Transl. Ross, slightly altered by Ferber.

a principle that mediates between the two. We can call this the principle of the appropriate. It may be paraphrased in a diversity of ways. Plato, for example, who does not yet know the generalisation rule, explains this principle as “...the graceful, the opportune, the right, and all that has its seat in the middle between two extreme ends”.²⁹ This “graceful and opportune” is not the ultimate good, that is, the idea or ideal of the good, which Plato also calls “the exact itself”, that is, the standard of measure.³⁰ But it shows how “the exact itself” is to be realised in empirical conditions, in which a certain inaccuracy occurs.

What is morally right is essentially expressed today through the generalisation rule, even though we cannot capture the ideal of the good by following it. But the generalisation rule must also be applied with the power of judgment. That is why its use, again, leads to a certain inaccuracy. We may try to avoid the extremes of discrimination against women and disadvantaging men, but we will not always be able to avoid both at the same time. Likewise, we will apply the rule against lying with judgment. An experienced physician will sometimes withhold the whole truth from a patient without actually lying.

The poet Ingeborg Bachmann (1926-1973) wrote: “People can face the truth.” But the truth can be traumatic. Therefore, we must tell it as appropriate, for example, waiting for the right moment, avoiding the extremes of deception and discouragement, and trying not to hurt any feelings.

The principle of the appropriate cannot supply a prescription for dealing with all the isolated incidents of life. Time and again, our power of judgment has to mediate afresh between the

²⁹ Plt. 284e. Transl. Skemp with modifications by Ferber. For this principle, cf. Ferber, *Propädeutische Lektüre des Politicus*, 63-74.

³⁰ Plt. 284d. Transl. Ferber.

generalisation rule and the individual incidents, paying attention to the demands of the appropriate. The use of this rule allows a certain modification, so that what is appropriate in one situation need not be appropriate in another. Here, the principle of Aristotle applies: “Such things depend on particular facts, and the decision rests with the perception [of the concrete situation].”³¹ The power of judgment is not actually the foundation of the right moral judgment, but when it operates appropriately, it completes that judgment.

4. Minimum and Maximum Morality

The generalisation rule does not explain all institutional facts of a moral nature, but only those that are necessary for a minimum or institutional morality. By “minimum morality”, I mean a morality that adheres to moral basic propositions, such as those forbidding killing and discrimination. I had to take these basic propositions as my starting point because they represent, as it were, the “primal or *ur*-phenomenon” of morality.

Morality begins with the ban on killing (and incest). As a result of our profound interest in life itself, the ban on killing carries exceptional weight. Thus, many states today protect even the life of a murderer. Modern morality is aptly described by the statement that all human beings are created equal and, as human beings, have the same rights. We find this not only in the American Declaration of Independence (1776; cf. p. 107), but also in the Declaration of Human and Civil Rights (1789) of the French National Assembly: “Men are born and remain free and equal in rights. Social distinctions may be founded only upon the general good” (Article 1). If human beings are born

³¹ EN, Book 2, Chapter 9, 1109b23-24. Transl. Ross.

and remain free and equal in their rights, no human being must be discriminated against because of an accidental human characteristic, such as race, gender, disability or sexual orientation. Since we also have a profound interest in a free and equal life, the ban on discrimination has likewise been given a very wide remit.

In contrast to such a minimum morality, there is a maximum morality, which commands us to love our neighbours and even our enemies. That is the morality of the Gospel and in particular of the Sermon on the Mount.³² The command to love our neighbours and our enemies can be generalised without embroiling us in a conflict of the will: If all people love their enemies, all my enemies also love me. This is something that I can obviously will without finding myself in a conflict of the will. But, however much we praise those who love – that is, sincerely do good to – their enemies, we do not blame those who do not love their enemy. Demands of the kind that call upon us to love our neighbours and our enemies are maximum demands. We could also call them ideals. However, as ideals, they rise above the concept of morality outlined here, although for smaller communities, for example, communities of Christians and other communities trying to live according to the demands of the Sermon on the Mount, they are binding.

Conversely, mere prudential rules, which only command us to act in our own interest, fall below the concept of morality, as I have sketched it, and belong in the field of eudemonism. Eudemonism tries to show how we can be happy or at least not unhappy. It does not prescribe, in a generally binding way, what we are to do and not to do. Rather, it gives recommendations that we may follow in order to reach a specific goal – happiness

³² Cf. Matthew 5.43-44, 22.38; Luke 6.27-30. Transl. King James Bible.

or well-being. Such recommendations may either keep within the bounds of morality or break them.

An example of advice keeping within the bounds of morality would be Kant's recommendation of "regimen, frugality, courtesy, reserve, etc., which experience teaches do, on the average, most promote well-being".³³ To this recommendation of an austere, aging philosopher, I would add Maxim Gorki's (1868-1936) statement: "Long live the man who knows not how to be frugal to himself."³⁴

In contrast, the prudential rules of unrestrained selfishness that a leader is advised to observe by Niccolò Machiavelli (1469-1527) in *The Prince* (1532) do not obey morality: "Injuries ought to be done all at one time, so that, being tasted less, they offend less; benefits ought to be given little by little, so that the flavour of them may last longer."³⁵

In our reflections, we have tried to follow a middle way between the demands of maximum morality, on the one hand, and the prudential rules of unlimited selfishness, on the other. The result is a generalisable minimum institutional morality, which only provides a few guidelines – for example, the "moral axioms" mentioned above – within which we can realise our vital interests and pursue our happiness. Nobody should do less than what is demanded by such a minimum morality. But everybody may do more.

Those who do more than is demanded by this minimum morality produce meritorious or "supererogatory" (from *superero-*

³³ Kant, *Groundwork*, Section 2, 418. Transl. Gregor.

³⁴ From the story "The Clock", Chapter 8. Transl. Ted Crawford, *The Clock*, in *British Socialist*, August 1912, 378-384, slightly altered by Ferber.

³⁵ Machiavelli, *Prince*, Chapter 8, 271. Transl. Marriott.

gare: paying over the odds) works.³⁶ Whoever not only pays lip service to the ideals of the Sermon on the Mount, but actually lives by them, can be said to do more than is normally expected. A general practitioner who moves to a distant region to develop basic medical services for a poor population, even though he could have a higher income and a more comfortable life as a specialist in the city, performs such a praiseworthy deed. Nobody is blamed for failing to act in this way, but those who do earn special merit. The concept of good, like the concept of truth, is an ideal concept. An ideal concept is never matched completely by reality.

The Gospel rightly says: “No one is good except God alone.”³⁷ This means that no human being, but only God, would completely fulfil this ideal concept. A human being can only try to get closer to God and the ideal concept of the good.

If the ideal concept of the good cannot be fully realised, it can even less be theoretically exhausted by our explanation of the morally good by means of the generalisation rule. Nevertheless, the generalisation rule is a minimum condition that must also be fulfilled by a maximum morality. It is a necessary, but by no means sufficient, condition for the maximum morality of loving one’s neighbour and one’s enemy.

Measured by minimum morality, and even more by the prudential rules of sheer selfishness, this love is unreasonable. Tolstoy’s Levin, for example, says: “Reason discovered the strug-

³⁶ Cf. the parable of the good Samaritan, Luke 10, 35: “And on the morrow when he departed, he took out two pence, and gave them to the host, and said unto him, Take care of him; and whatsoever thou spendest more (*quodcumque supererogaveris*), when I come again, I will repay thee.” Transl. King James Bible.

³⁷ Mark, 10, 18; Luke 18, 19. Quoted by Kant, Groundwork, Section 2, 408: and so He says of Himself, “Why callest thou me good? none is good, save one, that is, God.” Transl. King James Bible.

gle for existence, and the law that requires us to kill all who hinder the satisfaction of our desires. That is the deduction of reason. But the love of one's neighbour reason could never discover, because it's unreasonable."³⁸

Levin seems to be using the word "unreasonable" for "imprudence", and "reason" for "prudence". Reason, understood in such terms, is rather instrumental: It contents itself with finding the appropriate means to an end set by the struggle for survival. As we believe today, it does not necessarily follow from the struggle for survival that our opponents must be killed. We can just as well spare them. That is the case with stable or species-preserving evolutionary strategies, in which the opponents exercise restraint in using their deadly weapons. But as Nietzsche says in one of his lucid moments: "One has regarded life carelessly, if one has failed to see the hand that – kills with leniency."³⁹

Nevertheless, Levin, representing thousands who have done more for others than is customary with minimum morality, can only find in this ideal the missing meaning of life (cf. p. 32). But this ideal can only be believed and therefore no longer doubted. Tolstoy's great novel *Anna Karenina* ends with the simple words: "But my life now, my whole life apart from anything that can happen to me, every minute of it is no longer meaningless, as it was before, but it has the positive meaning of goodness, which I have the power to put into it."⁴⁰

³⁸ Tolstoy, *Anna Karenina*, Part 8, Chapter 12. Transl. Constance Black Garnett with small alteration by Ferber, London 2010.

³⁹ Nietzsche, *Beyond Good and Evil*, Chapter 4, § 69, 86. Transl. Zimmern.

⁴⁰ Tolstoy, *Anna Karenina*, Part 8, Chapter 19. Transl. Constance Black Garnett, London 2010.

5. The Generalisation Rule as an Axiomatic Demand of Practical Reason

We have explained the minimum condition of the good by the conflict of our will that follows from our failure to keep to it. If an individual said: “I want to do only what benefits me”, we would be able to tell him that he wants something that he will be unable to want as soon as he generalises the subjective principle of his will. For then the others could also do only what is to their advantage without taking account of his vital interests. But what if he asked us: “Why should I generalise the subjective principles of my will if I suffer no disadvantage by not generalising them?”

At this point, we can give him no further explanation. We would only be able to do so by introducing premises that cannot be supported by empirical evidence, for example, by arguing that all human beings are numerically one. Then the possible conflict between my vital interests and those of the others would have been overcome. Whatever I would do for the others, I would also do for myself. Whatever the others would do for me, they would also do for themselves. In either case, the rules of prudence would coincide with those of morality.

But the rules of morality are characterised by the fact that they do not always coincide with those of prudence. The moral demands, and in particular the generalisation rule, came into being because we are not all one and because between my will and the will of the others, there is a potential conflict. Just as we have to accept the metalogical axioms as institutions of the human language community (cf. p. 93), so we must also accept the generalisation rule as a superior axiomatic metainstitution above the institutional facts of morality. We must accept the metalogical axioms because without them we cannot talk meaningfully, and the metainstitution of the generalisation rule be-

cause we want to overcome the potential conflict between my will and that of the others.

While sceptics, in relation to the metalogical axioms, become embroiled in a contradiction of theoretical reason, in relation to the generalisation rule, they only face a discord of practical reason or the will. The difference is that a contradiction of theoretical reason is a logical contradiction, while a discord of practical reason or the will is a real conflict. This real conflict of having to will something that I do not will can occur at any time. But it usually does not occur until others will something that I do not will. Therefore, its occurrence or non-occurrence depends on empirical conditions.

Since not wanting the generalisation rule does not lead to a logical contradiction, the generalisation rule cannot be proven, like the metalogical axioms, for example, by the inconsistency of the attempt to abolish it. But neither can it be proven, like an empirical law, by claiming that, as experience teaches, moral behaviour always pays. On the contrary, honesty may be a costly thing.

Rather, the generalisation rule itself is a normative axiom for which neither logical nor empirical reasons can be given. It is an attribute of morality that it can make demands on me that go beyond the pursuit of my interests. However, an attempt to explain what goes beyond the pursuit of my interests through those same interests is predestined to fail. It leaves a gap in the argument, or a stain that cannot be erased. Thus, Mill's proof of utilitarianism by the axiom "Everybody strives for happiness, pleasure and utility" leaves a blemish in the picture of classic utilitarianism.

Kant realised that the generalisation rule, or categorical imperative – in his view "the supreme principle of morality" –

cannot be explained any further and we can only “comprehend” its “*incomprehensibility*”.⁴¹ Admittedly, we are not faced here with the “incomprehensibility” of a supernatural “fact of reason”, but rather with the “incomprehensibility” of a certain form of life. It is a form of life that *wills* this rule, regardless of what stage of development it has reached.

I am tempted to call this form of life the form of life of human beings *as* human beings. If I were asked: “Why should I generalise the subjective principles of my behaviour if I incur no disadvantage by not generalising them?”, I would only be able to answer: “You should still generalise the subjective principles of your behaviour.” And if I were asked further: “Why should I will something for which you cannot give me a reason?”, then I would only be able to answer, adapting a saying of Wittgenstein (cf. p. 53): “Here we can only describe and say: Such is moral life.”

This has brought us to a provisional ending of our introduction to key concepts of philosophy. Even though we have not yet seen the sun, we have worked a small part of our way out of the cave. Let us rest here for a while before we continue to move “from here to there”⁴² – perhaps to “the end of the journey”.⁴³

What all these concepts had in common was that we assumed them, but were unable to grasp them completely by our elucidations. That was particularly noticeable in connection with the last three concepts, truth, being and good. Their content went beyond any explicit definition. Key concepts of phi-

⁴¹ Kant, Groundwork, Section 3, Concluding Remark, 463. Transl. Gregor.

⁴² Basic formula frequently used by Plato, Phdr. 250e. R.529a. 619c. Tht. 176a-b.

⁴³ Plato, R., Book 7, 532c.

osophy are the meanings of key words of philosophy. The fact that we were unable to elucidate fully the meanings of these key words shows that they cannot be exhausted by our elucidations. What we grasped through them were only aspects of these concepts, as they appeared to us because of the “weakness” of our “arguments”.⁴⁴ This book was intended as a record of a small walk through philosophy and I am fully aware of its shortcomings. But the experience of philosophical inadequacy has been expressed more happily by a poet. Therefore, at the provisional end of our philosophical walk, let me return to Rilke’s poem “The Walk” (written in 1924, published posthumously) and reproduce it in its entirety in plain English:

Fixed on the sunlit hill, my gaze
runs ahead of the road I have scarcely entered on.
So does what we were unable to grasp
grasp us, full of appearance, from the distance –

and transform us, even if we fail to reach it,
into what, though hardly sensing it, we are:
a sign waves in reply to our sign ...
But we feel only the headwind.

⁴⁴ Plato, Ep. VII 343a.

