Two Arguments Against Lying

CHRISTINE M. KORSGAARD

Department of Philosophy
The University of Chicago
Chicago, Illinois 60637, U.S.A.

ABSTRACT: Kant and Sidgwick are at opposite extremes on whether we may tell paternalistic lies. I trace the extremism to their views about ethical concepts. Sidgwick thinks fundamental ethical concepts must be precise. Common Sense morality says we may tell paternalistic lies to children but not to sane adults. Because the distinction between a child and an adult is imprecise, Sidgwick thinks this principle cannot be fundamental, and must be based on the (precise) principle of utility, which often mandates paternalistic lies to adults. Kant thinks that ethical concepts are ideals of reason, which do not fit the world precisely because the world is imperfect. We lie to children and the insane because they are irrational, but no one is perfectly rational. We must treat all persons with the respect due to rational agents, so the pressure of the theory is toward not lying to anyone. Decisions about where to draw the line must be made pragmatically and to some extent arbitrary. But fear of this is not a good reason to abandon ethical ideals for utilitarianism.

KEY WORDS: Autonomy, coercion, ideals, Kantianism, paternalism, precise and pragmatic concepts, slippery slope arguments, theoretical and applied ethics, Utilitarianism.

§1

In recent years philosophers have welcomed the development of a widespread interest in philosophical ethics. In their concern about the bewildering ethical questions generated by medical technology, legal practice, and the power and responsibility of the modern corporation, members of the professions and of the public have turned to philosophy, traditional repository of rigorous moral thought. This concern has provided philosophers with an opportunity to show that our subject is important and useful, and that we have knowledge on which others might draw. And so the profession has responded with the development of courses, textbooks, and a vast literature on the questions of "applied ethics."

Yet a gap between traditional ethical philosophy and the solution of ethical problems remains. Writers on applied ethics do not seem to draw very heavily on traditional theories, and certainly do not draw on their details. Often the "application" consists simply in borrowing a principle which the theory defends. And often that principle gives an answer which seems too facile and too extreme. Theorists, in turn, know that you can have a real mastery of the concepts and arguments of a complex ethical

Argumentation **2** (1988) 27—49.

theory and yet, when confronted by an ethical problem, find that you have no idea what resolution the theory provides. To some extent this gap is sociologically produced, for often different people are drawn to the different kinds of work. But it may also be that we have not thought enough yet about what sort of an activity "applying" an ethical theory is.

This paper is an attempt to begin to think about that question. The moral problem I will consider is whether and to what extent we may tell benevolent or paternalistic lies. I take this to be an important problem about which we need theoretical help, because while most of us agree that there is a general presumption against paternalistic lies, we also agree that some of them are legitimate. The question is which ones and why. I will look at what happens when two of the most important moral theories in our tradition — those of Sidgwick and Kant — are applied to this problem. In each case, I will approach the question of when paternalistic lies are permissible by asking why they are in general wrong, so that we can identify the cases in which the consideration that disallows them does not hold. Sidgwick and Kant were both deeply concerned about the practical side of ethics and wrote more than most traditional moral philosophers about casuistical questions. Yet their views exemplify the tendency of theory to go to extremes, for Sidgwick's theory mandates benevolent deception on a massive scale, while Kant's apparently allows none.[1] I will show how the difference in their conclusions can be traced to the different requirements they impose on ethical concepts. One of the tasks of theoretical ethics is to provide an account of the status of ethical concepts — where they come from and why we are justified in using them. I will argue that differences in the status of ethical concepts can make a difference in how we decide which objects to apply them to. If this is right, theoretical and applied ethics are inseparable, not only because the principles are borrowed from the theories, but because each theory must teach us how its own principles and concepts apply to the objects and events we encounter in our lives.

§2

A consequentialist is committed to the view that the rightness or wrongness of an action depends entirely on its consequences. The right action is the one with the best possible consequences realizable under the circumstances. The consequences considered may be either the direct consequences of the particular action, or the consequences of a general rule the action instantiates, a disposition it expresses, or an institution or practice it serves. In any case, the value of the consequences must be determinable before the rightness of the action can be ascertained. Consequentialist reasoning about what is right must start from a knowledge of which consequences are good.

On a consequentialist view, lies will be wrong only if they do more harm than good. Whether they do so is a question of empirical fact. But most consequentialists have found it so obvious that lies tend to be harmful that they do not bother to argue the point. Sidgwick says simply that ". . . it is generally a man's interest to know the truth . . . ." (ME 448)[2] But it is clear how the argument should go, for knowledge of the truth is necessary for success at the instrumental reasoning by which the consequentialist believes the good is pursued. So we achieve the best results when we can rely upon one other to be sources of true information.

But benevolent lies are aimed at achieving good results, so if there is even a *general* presumption against them, it must be because they are normally misguided. This in turn would have to be because of some general empirical tendency. I think that the argument many consequentialists would use to explain the presumption against benevolent lies is that an individual is normally in the best position to judge what is good or bad for herself. If you lie to someone for her own good, it is because you think that you know better than she does what it will be good for her to know. Perhaps you think that she should be spared painful information, or diverted from making a poor decision. Usually, according to this argument, you will be wrong as a matter of fact, because she is in the best position to judge what is good in her own case.

A tension in this argument emerges when one reflects on two of the assumptions on which it depends. One of the assumptions of the consequentialist approach generally is that we can have determinate knowledge about what consequences are good or bad. This account of the presumption against benevolent lies assumes that persons are, as a matter of fact, the best judges of what is good or bad for themselves. These assumptions do not harmonize well together. We might ask, with Socrates, why the science of the good is the only science in which everyone is supposed to be equally expert.[3]

The most important school of philosophical consequentialists, the utilitarians, of course have an answer to this question. They believe that what is good or bad for a person depends on some sort of psychological state to which the person has special access: what is good is either pleasure, or the satisfaction of desire, and what is bad is pain or frustration. Sidgwick defines pleasure as "feeling which the sentient individual at the time of feeling it implicitly or explicitly apprehends to be desirable," making it by definition something to which the individual has special access. (ME 131) If a person has a special ability to identify what is good in her own case, then there is some reason to believe she is in the best position to make choices for herself. Since she needs the truth in order to make her choices well, we can argue on utilitarian grounds that she has a right to the truth.[4]

But this kind of consideration does not show that the individual is best at predicting what will *bring about* good results in her own case. She may lack the appropriate technical knowledge. And among the many problems

involved in the assessment of pleasure and pain which Sidgwick describes, there are some that put the individual in an especially poor position to judge what she will find pleasant at a later stage.[5] For example, in certain emotional states you cannot properly imagine the pleasures and pains attendant upon emotional states of an opposite kind. (ME 144—145) And persons are unable to properly anticipate how much they will enjoy certain things later in life. (ME 147) So the general rule that persons can make the best choices for themselves will often fail, and will not support a strong presumption against benevolent lies. Accordingly, the utilitarian will see many benevolent lies not merely as excusable, but as straightforwardly right.

Sidgwick explicitly acknowledges this. He thinks that the morality of Common Sense is "unconsciously utilitarian", and that its long-term logical and historical tendency is towards a more thoroughgoing utilitarianism. (ME 423—457) And one of the cases he adduces to show this is the case of the Common Sense attitude to benevolent lies:

> . . . where deception is designed to benefit the person deceived, Common Sense seems to concede that it may sometimes be right: for example, most persons would not hesitate to speak falsely to an invalid, if this seemed the only way of concealing facts that might produce a dangerous shock: nor do I perceive that any one shrinks from telling fictions to children, on matters upon which it is thought well that they should not know the truth. But if the lawfulness of benevolent deception in any case be admitted, I do not see how we can decide when and how far it is admissible, except by considerations of expediency; that is, by weighing the gain of any particular deception against the imperilment of mutual confidence involved in all violation of truth. (ME 316)

Sidgwick here employs a slippery slope argument common in defenses of utilitariansim. We may lie to children and invalids because it is for their own good. So why shouldn't we lie to other persons when it is for their own good? That Sidgwick himself is prepared to endorse large-scale benevolent deception becomes clear later in *The Methods of Ethics*, when he discusses the question whether utilitarianism should be publicly promulgated. Sidgwick believes he has established that according to utilitarianism, an action done in secret may sometimes be right which, if found out, would be wrong, because of the bad effects of publicity. But he thinks that ordinary persons will be misguided by this truth, and that they are better off believing they may only do what may be done openly. Because of this, utilitarianism, if true, should be kept "esoteric" for utilitarian reasons. With an evident relish for his paradox he concludes:

> Thus the Utilitarian conclusion, carefully stated, would seem to be this; that the opinion that secrecy may render an action right which would not otherwise be so should itself be kept comparatively secret; and similarly it seems expedient that the doctrine that esoteric morality is expedient should itself be kept esoteric. Or if this concealment be difficult to maintain, it may be desirable that Common Sense should repudiate the doctrine which it is expedient to confine to an enlightened few. (ME 490)

Of course, this is a recommendation for concealment rather than directly for deception, but it seems clear enough that deception will be involved; the "enlightened few" will almost certainly have to voice moral opinions they do not hold. Sidgwick thus favors wholesale deception of the public concerning important moral truths by an intellectual elite.[6]

Suppose that Common Sense, confronted by this unattractive result, digs in its heels and insists that the difference between a child and an adult is morally significant. While we may deceive a child for its own good, we may not do that to an adult. Sidgwick will reply that the distinction between a child and an adult is morally arbitrary because it is a difference of degree, a matter of placement on a continuum. There seems to be no non-arbitrary place on the continuum to draw the line. If we do draw the line at some selected place, we will get irrational results: I may lie to you for your own good the day before your eighteenth birthday, say, but not the day after. Yet all that has happened is that you are two days older, and that doesn't seem to be a good reason for the difference. This violates one of the fundamental axioms of ethics, which is that we must not treat different cases differently without a good reason. (ME 209, 379) Moral issues cannot be settled by appeal to arbitrary distinctions like this, but must be settled by the principle of utility, because that principle will give us a clear justification for lying in some cases and not in others. Sidgwick's commitment to this kind of argument is explicit in a discussion of the principle that everyone should be equally free. He argues:

> . . . it seems obviously needful to limit the extent of its application. For it involves the negative principle that no one should be coerced for his own good alone; but no one would gravely argue that this ought to be applied to the case of children, or of idiots, or insane persons. But if so, can we know *a priori* that it ought to be applied to all sane adults? since the above-mentioned exceptions are commonly justified on the ground that children, etc., will manifestly be better off if they are forced to do and abstain as others think best for them; and it is, at least, not intuitively certain that the same argument does not apply to the majority of mankind in the present state of their intellectual progress. Indeed, it is often conceded by the advocates of this principle that it does not hold even in respect of adults in a low state of civilisation. But if so, what criterion can be given for its application, except that it must be applied wherever human beings are sufficiently intelligent to provide for themselves better than others would provide for them? and thus the principle would present itself not as absolute, but merely a subordinate application of the wider principle of aiming at the general happiness or well-being of mankind. (ME 275)

This passage exemplifies a kind of slippery slope argument that Sidgwick employs frequently in his debunking of Common Sense morality. This kind of slope is especially slippery because it has two sources. One is the usual utilitarian source: Sidgwick claims that we coerce children because they will be better off, and if this is a reason in one case, it will be a reason in another. But the other source of slipperiness is that the moral principle is operating with what I will call a pragmatic, as opposed to a

metaphysically precise, concept. If a concept is metaphysically precise, its application is determined wholly by features of the thing we apply it to. If it is pragmatic, its application is indeterminate — there are borderline cases — and it can only be made determinate artifically. A concept which divides a continuum is often pragmatic, and must for practical purposes be to some extent arbitrarily set. This need not mean that we have no good reason for drawing the line at the point that we do (although it might), but it will mean our reasons are considerations other than features of the objects to which the concept is applied. The distinctions between the child and the adult, the deeply troubled and the mentally ill, the "idiotic" and the slow, are in this way pragmatic. These are divisions of continuums, and where exactly we should draw the line is unclear. To some extent the way that these distinctions are made does depend on features external to the persons in question — such as the kinds of work and the kinds of care that are available in the societies and cultures in which the distinctions are employed.

We all recognize that the concepts with which our positive laws operate are in this way pragmatic. For instance, we are content to define a legal adult as anyone over eighteen or twenty-one. This is not because although some people are "really" adults by that age and some are not, we must have the same law for everyone. Perhaps there is a temptation to think that there must be a certain degree of something — rationality, intelligence, self-control — that qualifies you as a metaphysically real adult, and that the ideal legal system would identify that and assign legal majority only when it is present. But rationality, intelligence, and self-control are themselves matters of degree, and the degree of them required for adulthood will in turn have to be determined pragmatically. This sort of consideration teaches us to resist the temptation to look for metaphysical precision in the concepts employed in the positive laws, and to accept the artificiality and cultural mutability of the concepts they employ.

But many people are resistant to the idea that moral laws might operate with concepts of this kind as well. There is a temptation to think that a moral principle can only operate with a metaphysically precise concept, one whose application is at least in principle perfectly clear. This gives rise to a slippery slope argument that is a commonplace in moral debate. For example, some people believe that we cannot settle the question of the morality of abortion until we find a metaphysically precise distinction between mere fetuses and persons. As in the case of adult and child, there is no reason to believe that there is a precise distinction, or that one adopted would not itself be in part determined by pragmatic considerations, for the difference is one of degree. Now there are people who will use the fact that the mere-fetus/person distinction is one of degree as an argument against abortion. They create a slippery slope from killing fetuses to killing infants to killing still older persons, and say where will it

all end. The only real difference here is a difference in the degree of dependence, and any setting of a certain amount of independence as a criterion of personhood will be arbitrary. This kind of argument is also sometimes used against vegetarianism. The ultimate indeterminateness of the plant/animal distinction can be used to create a slippery slope. These arguments are similar to the argument that we may be paternalistic towards adults because we are paternalistic towards children and the only real difference is one of degree. The philosophical instinct behind these arguments is that moral principles can only operate with metaphysically precise concepts or natural kinds and not with pragmatic concepts.

These slippery slope arguments are notoriously weak. The trouble is that which way the slope runs depends on which end of it your intuitions are most securely fastened to. Imagine someone who argues that the benefit to a child is not a sufficient reason for coercing him, for at that rate we might sometimes coerce adults. Or imagine someone arguing that there is no reason we should not euthanize a seriously disabled and disfigured infant, just as we would not hesitate to abort a seriously disabled and disfigured fetus. (You just raise up the lower end of the slope and the ball slides back into the other fellow's lap). These arguments are not as commonly made as their conservative counterparts, but they are no worse — and no better. That is the trouble — the slippery slope produced by the pragmatic concept can prove nothing, but at best is a heuristic device that shows us where we must look for some principles that will help.

Now Sidgwick is deeply concerned with this sort of problem. He believes that ethics should proceed from self-evident principles, and that one of the criteria of a self-evident principle is that "the terms of the proposition must be clear and precise." (ME 338) Because pragmatic distinctions either are not clear and precise, or are only arbitrarily so, they destroy the self-evidence of the principles that employ them. (ME 293n) In a discussion of whether government is justified by consent of the governed as expressed in voting, Sidgwick remarks that if it is voting that legitimates government then even children must be allowed to vote. And, he complains:

> . . . if to avoid this absurdity we exclude children, an arbitrary line has to be drawn . . . (ME 298)

This destroys the self-evidence of the principle. Sidgwick believes that an arbitrary and pragmatic character infects many of the concepts used in Common Sense moral principles, and he directs his account of Common Sense morality to freeing "the common terms of Ethics, as far as possible, from objection on this score." (ME 339) Because he thinks that his efforts to do this are unsuccessful, he thinks that only the principle of utility can resolve moral problems. This comes out clearly in Sidgwick's recurring reflections about the concept of "veracity" itself. Discussing the "Intui-

tional" view that veracity is right regardless of consequences, Sidgwick says:

> . . . we find that in the common notion of different kinds of actions, a line is actually drawn between the results included in the notion and regarded as forming part of the act, and those considered as its consequences. For example, in speaking truth to a jury, I may possibly foresee that my words, operating along with other statements and indications, will unavoidably lead them to a wrong conclusion as to the guilt or innocence of the accused, as certainly as I foresee that they will produce a right impression as to the particular matter of fact to which I am testifying: still, we should commonly consider the latter foresight or intention to determine the nature of the act as an act of veracity, while the former merely relates to a consequence. We must understand then that the disregard of consequences which the Intuitional view is here taken to imply, only relates to certain determinate classes of action (such as Truth-speaking) where common usage of terms adequately defines what events are to be included in the general notions of the acts, and what regarded as their consequences. (ME 97)

But later Sidgwick *denies* that common usage adequately defines which consequences are to count as rendering a communication veracious:

> . . . we found no clear agreement as to the fundamental nature of the obligation; or as to its exact scope, i.e. whether it is our actual affirmation as understood by the recipient which we are bound to make correspondent to fact (as far as we can), or whatever inferences we foresee that he is likely to draw from this, or both. (ME 355)

This sort of problem becomes especially acute when a gradual change in our linguistic practices renders it uncertain what inferences the recipient of our words will draw from them.

> In the case of formulae imposed by law — such (e.g.) as declarations of religious belief . . . a difficulty is created by the gradual degradation or perversion of their meaning, which results from the strong inducements offered for their general acceptance; for thus they are continually strained and stretched until a new general understanding seems gradually to grow up as to the meaning of certain phrases; and it is continually disputed whether we may veraciously use the phrases in this new signification. A similar process continually alters the meaning of conventional expressions current in polite society. When a man declares that he 'has great pleasure in accepting' a vexatious invitation, or is 'the obedient servant' of someone whom he regards as an inferior, he uses phrases which were probably once deceptive. If they are so no longer, Common Sense condemns as over-scrupulous the refusal to use them where it is customary to do so. But Common Sense seems doubtful and perplexed where the process of degradation is incomplete and there are still persons who may be deceived . . . (ME 314—315)[7]

Sidgwick also emphasizes our doubts about whether the use of such devices as *suppressio veri* and *suggestio falsi* is within the scope of veracity. (ME 316—317, 448—449) The implication of these remarks is that the very concept of veracity is a pragmatic one, which therefore cannot play a role in a self-evident principle.

For Sidgwick, a large part of the appeal of the principle of utility is that it enables us to settle questions in a definite way. It enables us to apply the concepts of right and wrong firmly to cases by linking them, at least

in principle, to solid empirical facts, facts about pleasure and pain. It bypasses the use of pragmatic distinctions like those between veracity and falsehood, adults and children. *The Methods of Ethics* is littered with questions like: "We seem to condemn either extreme: yet what clear and accepted principle can be stated for determining the true mean?" (ME 249) "Where then is the limit to be fixed?" (ME 334) "But if the rule does not hold for an extreme case, where can we draw the line?" (ME 308) and the answer is always "We do not seem able to obtain any clear and generally accepted principle for deciding this point, unless the Utilitarian formula be admitted as such." (ME 348) Utility is the central concept of ethics because it is the only one with the precision requisite to render its principle self-evident.[8]

§3

The Kantian begins from an altogether different theory of value. Kant distinguishes those things which are unconditionally good — things which are good in themselves and so in any and all conditions — from things which are only conditionally good — that is, things whose goodness depends upon external conditions. According to Kant, the only unconditionally good thing is a good will. (G 393/9) Everything else which is good has only a conditional value. In the argument that culminates in the second formula of the Categorical Imperative, the Formula of Humanity, Kant tells us that "All objects of inclinations have only a conditional worth, for if the inclinations and the needs founded on them did not exist, their object would be without worth." (G 428/46) In other words, the things we desire have value, because we want and need them, not the reverse. Our desire is a condition of their value. Our wanting them is not enough to make them good, however, for obviously many of the things we want are not good. Even if we want them we will not judge them good unless they are conducive to our happiness. And:

> It need hardly be mentioned that the sight of a being adorned with no feature of a pure and good will, yet enjoying uninterrupted prosperity, can never give pleasure to a rational impartial observer. Thus the good will seems to constitute the indispensable condition even of worthiness to be happy. (G 393/9)

When an object is conditionally good, we must determine whether the condition of its goodness is met. Say the initial condition is that someone desires that a certain state of affairs be realized. Then we can raise the further question why it should be good that this person's desire should be satisfied. If we say it makes him happy, we can ask why it is good that he should be happy. In this way we give rise to a regressive argument, a search for the unconditioned condition of the goodness of the thing. In

Kant's view, this condition will always be the presence of a good will. A good will confers value on the objects of choice of the person who has it, for

> . . . rational nature exists as an end in itself. Man necessarily thinks of his own existence in this way; thus far it is a subjective principle of human actions. Also every other rational being thinks of his existence by means of the same rational ground which holds also for myself; thus it is at the same time an objective principle from which, as supreme practical ground, it must be possible to derive all laws of the will. The practical imperative, therefore, is the following: Act so that you treat humanity, whether in your own person or in that of another, always as an end and never as a means only. (G 429/47)

By "humanity" Kant means the power of free rational choice, for "the capacity to propose an end to oneself is the characteristic of humanity" (MMV 392/50) and in fact "rational nature is distinguished from others in that it proposes an end to itself." (G 437/56) According to Kant the choice of an end is always an act of freedom. (MMV 386/42) We regard the objects of our own free rational choices as good (this is a "subjective principle of our actions") and so we must regard the objects of the free rational choices of others as good as well. This makes humanity, or the power of free rational choice, the limiting condition of the rationality of choice itself. That is to say: for a choice to be rational, it must be consistent with the supreme and unconditional value of the power of free rational choice.[9]

One of the examples that follow the Formula of Humanity is the case of a lying promise. A man is in need of money and "borrows" some, falsely promising to pay it back although he knows he will not be able to. Kant explains what is wrong with the action this way:

> For he whom I want to use for my own purposes by means of such a promise cannot possibly assent to my mode of acting against him and cannot contain the end of this action in himself . . . he who transgresses the rights of men intends to make use of the persons of others merely as a means, without considering that, as rational beings, they must always be esteemed at the same time as ends, i.e., only as beings who must be able to contain in themselves the end of the very same action. (G 430—431/48)

When Kant says that the other "cannot assent to my mode of acting against him," he does not mean merely that if the other knew I was lying he would find my conduct objectionable. This is probable, of course, but he might not. The case might be trivial and he might not care how I act; it must still turn out to be true that he "cannot assent to my mode of acting" if this is a correct analysis of what is wrong. Similarly, when Kant says that the other is "unable to contain the end of this action in himself," he does not merely mean that the other cannot happen to have the same purpose. Obviously he can. In the example, for instance, the deceived lender may be a generous person who was longing for a way to give me the money outright without wounding my pride. Surely this fact would not mitigate

the wrongness of my deception. What Kant means, I believe, is that the other person cannot assent to my action because he is not in a position to. This is because he is deceived. By the nature of the case, he doesn't know how I am acting, and you cannot assent to a transaction you do not know is occurring.[10] In the same way, he cannot "contain in himself" the end of the same action because he is not in a position to. He doesn't know what the real end of the action is, and is therefore not in a position to make it his own — to choose, freely, to contribute to its realization. In the example, the other person believes the end of the action is my temporary possession of his money, when in fact it is my indefinite possession of it. What makes my action wrong is not that he *would* not have chosen to cooperate in promoting that outcome, but simply that he *did* not.[11]

If our way of acting is to be consistent with the unconditional value of free rational choice, others must be able to assent to the transactions in which we engage them, and be in a position to voluntarily choose to further the ends at which those transactions are aimed. There are two things that interfere with these requirements: coercion and deception. The person who is forced to engage in a transaction or further an end, and the person who is fooled into it, cannot be said to assent to the transaction or choose the end. These persons are therefore being treated as mere means to the end of others. As free rational beings, persons must be allowed to choose for themselves what transactions they will engage in and what ends they will promote. This is what it means to treat persons as ends in themselves.

This argument not only shows why lying is wrong, but that it is one of the two most fundamentally wrong things you can do to others. Coercion and deception are the two ways of using others as mere means. Lies are therefore wrong in themselves, regardless of whether they are told with good intentions or bad. Kant condemns benevolent lies for exactly the same reason as mischievous or malicious ones. But Kant's theory of value implies something even stronger than this. A coercive or deceptive character not only renders an action wrong, but robs the end of the action of its value. For the goodness of the end depends upon the rationality of the choice of the end, but an end pursued by way of a bad action has not been rationally chosen. The condition of its goodness — consistency with the value of free rational choice — is not met, and so it is not a good end. In this way, Kant's value theory leads to a sort of unanimity criterion of value.[12] "Good" is a concept of practical reason, and so "what we call good must be, in the judgment of every reasonable man, an object of the faculty of desire, and the evil must be, in everyone's eyes, an object of aversion." (C2 60—61/62—63) For an end to be good, it must be possible for everyone to agree to it.

> . . . reason has no dictatorial authority; its verdict is always simply the agreement of

> free citizens, of whom each one must be permitted to express, without let or hindrance, his objections or even his veto. (C1 A738—739/B766—767/593)

An end achieved through a lie violates this condition, for the deceived person has not had a chance to agree or to cast his veto. Thus, a lie for a *good* end is not only wrong. It appears to be, strictly speaking, *impossible.*

Now this result seems much too strong. For most of us think that there are some cases in which benevolent lies are permissible. The ends in these cases will be ones which, whatever Kant's theory says, we would certainly regard as good if they could be achieved by honorable means. For example, there is Sidgwick's case of the lie told to an invalid to spare him a dangerous shock. Sidgwick points out that we allow the use of coercion on children, idiots, and the insane; as we have seen Kant ranks deception and coercion together. Kant would have to agree that we may sometimes use coercion on such persons, for their own safety and that of others. And if we may use coercion, then we may lie. Presumably there will be cases in which lying is the preferable stategy, since coercion may well involve violence. If there are cases where benevolent deception is allowable, how can Kant's theory account for those cases?

One thing that we cannot do is handle such cases by saying *simply* that the person to whom we lie *would* agree to the end for which the lie is told. This would be to adopt the reading of Kant's "possibility of containing the end" criterion which I have rejected. For one thing, this would justify manipulative conduct towards those who happen to share our ends (as in the case of the false promise made to one who would be glad to give you the money outright). Furthermore, for me to make the *judgment* that you would agree to this end is already for me to take the decision out of your hands in the way that the theory forbids. Although Kant agrees that we have a duty to promote the happiness of others, "What they count as belonging to their happiness is left up to them to decide." (MMV 388/46) So benevolence cannot override the essential duty to let others determine their own ends. If I lie with the excuse that you would agree to my end, I am attempting to justify preempting your free use of your reason by appealing to an idea that I have arrived at because I have already preempted it. This clearly will not do.

We could get a more promising argument for permissible lies if we could establish in the pertinent cases that the deceived one would, to use Kant's phrase, "assent to my mode of acting against him." This has some of the same difficulties about preemptive judgment as the rejected suggestion, but at least it will only excuse lies to those who would specifically agree to be lied to, not to everyone who shares the purpose for which the lie was told. But in order to keep these two excuses distinct, we must not allow the inference that someone would agree to be lied to, simply *because* they would agree to the purpose for which the lie was told.

Otherwise we will be back where we started. In any case, that inference seems incorrect. The requirement that each person be allowed the free use of her own reason in determining her actions and ends is a procedural one. It is important not only that the choice be the one that a person would have made, but that she actually gets to make it.

So our argument that the person would agree to be lied to must not appeal to, or appeal only to, the end for which the lie is told. We must justify the lie more directly. It seems clear that in order to annul the procedural requirement that each person decide for herself we must show that the procedure could not in any case be carried out. We must argue that the agent is not in a position to make a free rational decision — that she is not (at present, perhaps) autonomous, and that were she rational she would agree that someone in her (actual) condition may be lied to or coerced. Of course, we do not want to justify telling *any* lies whatever to non-autonomous persons. The hypothetical acceptance of *both* the mode of action *and* the end must be defended. And this is the usual way of justifying our manipulative transactions with children, the mentally handicapped, and the insane. We argue both that they are not autonomous because of their conditions, and that they would if rational agree that *such* persons may be manipulated for *these* ends to be realized.

The ends must be ones that we can fairly presume that non-autonomous persons would hold if they were rational. How can we tell what ends they would choose without being guilty of preemptive judgment? One end that we can attribute to them without preemptive judgement is their humanity itself, their power of free rational choice. Since this is unconditionally valuable, this is an end they would hold if they were rational. Presumably we may lie to someone who lacks autonomy if our end is to restore or preserve her autonomy, or to restore or preserve things which are necessary conditions of it, such as physical well-being. Since all human beings want to be happy, we can also safely attribute this end to them (G 415/33), but we will be on less firm ground here because we cannot without preemptive judgment determine what they would count as part of their happiness. And this must be treated as secondary to the restoration of autonomy, if that is possible.

Because we cannot justify the lie by appeal to its purpose alone, but must also appeal to the lack of autonomy of the person lied to, we will not be able to determine which paternalistic lies are permissible unless we have some criteria for determining who is autonomous. But here we run into a difficulty. Autonomy in the ordinary sense appears to be a matter of degree. Children are thought to acquire it little by little, for instance. So again we get Sidgwick's question about where to draw the line — how we are to decide what counts as autonomous enough to ground a claim against paternalism. Sidgwick thinks that we should refrain from using coercion on those who are "sufficiently intelligent to provide for them-

selves." Presumably Sidgwick thinks that "intelligence" is a matter of degree, and a certain degree of it is sufficient for expedient self-government. Whether you have this degree of intelligence will show up in various ways, among them the choices you actually make when left to your own devices. The one deciding whether to resort to paternalism judges whether the other is making her choices well. Now the Kantian cannot solve the problem about ascertaining and measuring autonomy this way, for several reasons. First, because Kant does not have an empirical theory of value, we are not in a position to judge the quality of the choices of others. Of course we can see in the ordinary way if someone is poor at the choice of means, and we can ascertain whether his choices are immoral. But because each person determines for himself what he will count as part of his happiness, a person's choice of the ends he will pursue is not otherwise open to assessment. We cannot decide someone is insufficiently autonomous merely because he makes some choices we would not have made in his place. But further: we cannot decide someone is not autonomous because he is regularly making decisions we know are bad in the sense of immoral. It is the capacity for autonomy, not its actual exercise, which gives a person the right to self-government. Kant's theory may allow us to use manipulative tactics on children and the insane, on grounds of insufficient autonomy. It does not allow us to use them on persons who are immoral. We must have a way to distinguish persons who should be regarded as capable of autonomous conduct, however they may actually behave, who therefore are held responsible when they go wrong, and to whom we have no excuse for lying, from persons who should be regarded as incompletely developed or ill, who therefore are not held responsible when they go wrong, and to whom we may lie under the conditions described above.[13]

What makes this especially difficult is that the capacity for free rational choice, as Kant understands it, is not really a natural attribute at all. Freedom of the will is an idea of pure practical reason, and, according to Kant, a property that we ascribe to ourselves when we act. We ascribe it to ourselves because the moral law commands categorically, and we recognize that we can do what we ought. (C2 30/30; 42/43) We do not ascribe it to ourselves because we have theoretical evidence of its presence — that would be impossible, for in Kant's philosophy freedom is beyond the limits of theoretical knowledge. Similarly, we ascribe it to other persons not because we have theoretical evidence that they have it but because it is a duty to do so. Respect for humanity in the persons of others demands that we attribute to them the capacity for free choice and action, regardless of how they are acting:

> Thus it is also with the reproach of vice, which must never burst out in complete contempt or deny the wrongdoer all moral worth, because on that hypothesis he could

> never be improved either — and this latter is incompatible with the idea of man, who as such (as a moral being) can never lose all predisposition to good. (MMV 463–464/129)

Actual conduct, then, does not provide evidence for or against freedom. Furthermore, since every human being is imperfectly rational, we all fall short of acting completely autonomously. But then how are we to distinguish the psychotic, whose conduct should not be attributed to freedom, from the evil person, whose conduct should? We cannot appeal directly to the kind of explanation their conduct admits of. Kant's view is that *all* conduct, viewed theoretically, admits of a deterministic causal explanation. An evil person's conduct is attributed to but not *explained* by free choice. The idea of freedom has no explanatory role at all. Freedom is something we attribute to ourselves and others for moral purposes only. (C2 51–58/52–59; 134–137/139–142)

In Kant's view, moral concepts generally are concepts of pure practical reason. They do not describe or explain anything about the world. They get their hook into the world via the practical reason of each one of us, as individuals who make choices and decisions, who are therefore bound by the moral law, and who in light of that must attribute freedom to ourselves and others. Practical concepts are ideal concepts, and the things in the world to which we apply them, under the direction of the moral law, fall short of them. No one exhibits perfect rationality, and so no one exhibits the perfect freedom which the moral law teaches us that we have. But this means that facts about the world will not decisively settle the question which natural objects these concepts are to be applied to and which not. What we find in the world is a continuum of more or less imperfectly rational beings. As a rule we are obliged to treat them as capable of perfect rationality. If we decide that somewhere along that continuum we will draw a line, beyond which we are excused from that obligation, the decision will be to some extent arbitrary. Or in any case, it will not be made by seeing how close various persons come to actual freedom. We cannot say: "children are not yet *in fact* fully free." Theoretical statements about freedom have no standing. The most we could say is that children should not be treated as fully free. But the grounds for this practical statement are unclear. The pressure of the moral law is towards treating every human being as a free rational being, regardless of actual facts.

§4

Both Sidgwick and Kant uncover problems about applying moral concepts to the world. Sidgwick thinks they are usually imprecise, and that ultimately we should be guided instead by the precise concept, utility. Kant thinks that they are precise in themselves but don't exactly fit the things we find

in the world, because they are ideals. He thinks we should apply them anyway, without regard for the world's imperfections. In this section I discuss another example of a moral concept about which it is particularly clear that the lines between Sidgwick and Kant are drawn in the way I have suggested.

The concept in question is that of "legitimate government." Both Kant and Sidgwick consider the question when a government established by revolutionary overthrow of a previous government becomes legitimate. Sidgwick says:

> All are agreed that usurpation ought to be resisted; but as to the right behaviour towards an established government which has sprung from a successful usurpation, there is a great difference of opinion. Some think that it should be regarded as legitimate as soon as it is firmly established: others that it ought to be obeyed at once, but under protest, with the purpose of renewing the conflict on a favourable opportunity: others think that this latter is the right attitude at first, but that a usurping government, when firmly established, loses its illegitimacy gradually, and that it becomes, after a while, as criminal to rebel against it as it was originally to establish it. And this last seems, on the whole, the view of Common Sense; but the point at which the metamorphosis is thought to take place can hardly be determined otherwise than by considerations of expediency. (ME 300)[14]

Kant, on the other hand, places himself firmly in the first camp Sidgwick mentions. As is well known, Kant believes that any revolution is wrong, since the existing government must be taken to represent the general will of the people. For a revolt to be legitimate, it would have to be in *accordance* with the general will of the people, since coercion that does not meet this condition is mere illegitimate violence. But since the existing government represents the general will of the people, the revolutionary party of course does not. (MMJ 318—321/84—87) However, Kant also holds that if a revolution succeeds, the new government is immediately legitimate and owed the same allegiance as the old was. (MMJ 323/89) Taxed by a reviewer[15] with the oddity of this position, Kant replies:

> Every matter of fact is an object that is an appearance (of sense); on the other hand, that which can be represented only through pure reason and which must be included among the Ideas — that is the thing in itself. No object in experience can be given that adequately corresponds to an Idea. A perfect juridical [just] constitution among men would be an example of such an Idea.
>
> When a people are united through laws under a suzerain, then the people are given as an object of experience conforming to the Idea in general of the unity of the people under a supreme powerful Will. Admittedly, this is only an appearance; that is, a juridical constitution in the most general sense of the term is present. Although the [actual] constitution may contain grave defects and gross errors and may need to be gradually improved in important respects, still, as such, it is absolutely unpermitted and culpable to oppose it. If the people were to hold that they were justified in using violence against a constitution, however defective it might be, and against the supreme authority, they would be supposing that they had a right to put violence as the supreme prescriptive act of legislation in the place of every right and Law. (MMJ 371—372/139—140)

The legitimate state, like the autonomous moral person, is an idea of pure practical reason. Kant is clear that the form of government that corresponds to this idea is a republic: a constitutional government in which legislation is carried on by representatives of the citizens (MMJ 340—341/112—113; PP 349—353/93—97) and in which every adult is or can become a free, equal, and independent citizen (MMJ 313—315/78—80). But it is also clear that that no actual government is adequate to this idea. Still, we don't feel like saying that no government has ever really been legitimate. Once again we get the problem of saying what comes close enough. Perhaps it seems as if in this case we can easily say what comes close enough: the existence of a constitution and a legislative body of elected representatives. But even this would mean that there had been no legitimate governments until very recent history. And if we suppose a legitimate government must meet another criterion which seems minimal and obvious — universal adult suffrage — then there have been no legitimate governments until this past century (our own, not Kant's). Thus, rather than deciding how close a government has to be to the ideal before we will call it legitimate, Kant thinks we should simply treat *any* existing government as legitimate. Like Sidgwick, although for a different reason, he does not see where to draw the line.

§5

The two theories that I have been examining take extreme positions on paternalistic lies. Sidgwick allows too many, Kant apparently allows none. I have traced the extremism of both views to a common source: a problem about applying ethical concepts to the world. Both reject a popular model of applied ethics, according to which we work like this: we find out which objects in the world fit the ontological categories with which our principles operate (for instance, which things are in fact mere fetuses and which are persons), and then marshall in the principles to see what they say about those kinds of objects. On this popular model the real work — the hard work — of applied ethics is not ethical thought at all, but ontology or metaphysics. Once the ontological issues are settled, the ethical principles are easy to apply. Sidgwick and Kant both reject this model, because of a mismatch each perceives between the world and the concepts we employ in ethical thought. The ontological issues cannot be settled in the way proposed.

For Sidgwick, the problem arises because the concepts of Common Sense morality are imprecise. A clear and precise concept would be one that picks out its objects in a way that is non-arbitrary and decisive.[16] Only a precise concept can function in a real self-evident moral principle. Since the concepts of "adult," "sane" and "veracious" are all imprecise, "never lie to a sane adult" could not possibly be a self-evident moral principle. The

concept of utility is precise. Pleasure and pain seem to have clear moral weight, yet they are also matters of theoretical fact. So utility allows us to map practical concepts onto theoretical ones perfectly. It makes the concepts of "right" and "wrong" fit the contours that the world provides — or more precisely, that the set of concepts we use in theoretically explaining the world provides.

For Kant the problem arises not because moral concepts are imprecise, but because they are ideals of pure practical reason. Nothing we find in the world is fully adequate to them, yet we must apply them to things that we find there anyway, for moral reasons. Thus, each person ought to be treated like a free rational being, despite our imperfect rationality; just as each government ought to be treated as the embodiment of the general will of the people, though most real governments are corrupt despotisms born of illegitimate violence. One must not lie to a free rational being, or revolt against the general will of the people. If lying for a good end, or revolting for a better government, were *justified* simply because actual persons and states fall short of the ideal, a slippery slope argument would quickly lead to extreme conclusions, since all persons fall short of perfect autonomy and all governments fail to embody the general will of the people.

So perhaps we should say that what we are looking for is not good *reasons* for paternalism or revolution, since there are not good reasons for what the moral law disallows, but *excuses* for these kinds of action. We think we may be excused from treating certain persons as autonomous and certain governments as legitimate. But in order to decide that there are some human beings to which we will not apply the concepts of full autonomy, or some governments too despotic to count as legitimate, we apparently will have to draw some lines that are not firmly grounded either in theoretical facts about those persons and governments or in the moral law. We must *decide* who to count as a free rational being, and what makes a state legitimate. These decisions cannot be based on finding out which objects in fact are free rational beings or really embody the general will, because there are no such facts. If applied ethics is done this way, its work is not metaphysical, but ethical and practical all the way down. Kant's own solution is to treat every government as legitimate and (almost) every human being as free and rational. This is already a pragmatic decision, and it may seem to us to be — not the incorrect one, because this is not a matter of correct or incorrect — but a poor one. But now we must ask what the criteria for making these decisions should be. If we don't pick those who must be treated as free rational beings by determining who is in fact a free rational being, how do we do it? We need some alternative method, or our decisions will be arbitrary.

The answer is that of course such decisions don't have to be *completely* arbitrary. In deciding what objects to apply our ethical concepts to, we use

available theoretical concepts that approximate them.[17] Although we cannot find a perfect match between ethical concepts and natural facts, theoretical considerations do provide guidance. After all, it is theoretical facts which teach us which things are even potentially rational beings. The ethical concept of a moral person is mapped on to the naturalistic concept of a human being. And theoretical facts do teach us that there is a sense in which children and the insane are not as "autonomous" — that is, not as much in control of themselves — as ordinary adult persons are. The sense of "in control of oneself" here will not be very determinate. It will be a matter of degree, and hardly adequate to Kant's notion of transcendental freedom. But we must use if for casuistical purposes, because it is all we have.[18]

Although Sidgwick wants to use the principle of utility directly for decision-making rather than for concept-formation, one can imagine a Sidwickean proposal that we make the notion of an autonomous rational agent precise by tying it to considerations of utility.[19] If what is at issue is not that the concept be applied correctly, but that it be applied well, the utilitarian will read "well" as "expediently." So the Sidgwickean might urge us to count persons as autonomous when it is expedient — in utilitarian terms — to do so. This would give us a clear way of deciding — in terms of theoretical facts — who was to be counted as autonomous and who was not. Applied ethics would then recover the intellectual cleanliness of the popular model — the real work would be theoretical. But the trouble with the proposed way of achieving precision is that it would make the notion of autonomy *subordinate* to that of utility. Freedom would not function as an independent value at all. To only be truthful when it is useful is to care only for utility, not for freedom. If we adopt the Sidgwickean strategy, we can seek empirical goals, but we cannot live up to rational ideals. Besides, if we are going to use values in deciding how to apply our ethical concepts, why should we only use utility? To say that our concepts can be applied well or badly needn't mean well or badly in the utilitarian sense. Why shouldn't we apply them in a way that maximizes fairness or personal freedom?

A natural objection is that if we work this way there is a danger that we will simply adjust our ethical categories and principles reciprocally until we get whatever result we want. We could just decide, for example, that all of the categories of persons to whom we are tempted to tell benevolent lies should count as incompletely autonomous. Or that benevolent lies in certain cases aren't really lies at all. Of course there are dangers of this kind, but it is not clear that we have any option but to face them, and to try to be intellectually honest. If the concepts we use in ethics cannot be tied down to the world by firm theoretical criteria, the precision of the popular model of applied ethics is just not available.

It is true that the principle of utility gives us a way of recovering that

precision. But as we have seen, the price of defining autonomy in terms of utility is to give it no real weight at all. So now we must ask whether Sidgwick's demand for precision in ethical concepts is worth the cost of giving up rational ideals. It is clear enough why in scientific discourse we should want concepts that fit as neatly as possible onto the world. It is not at all clear why that should be appropriate for normative discourse. If there are ideals of practical reason, which outstrip the things we find in the world, we cannot give them up merely because they make applying ethical principles a difficult and uncertain enterprise. If we value the Kantian ideal of free and non-manipulative relations among rational beings, for instance, and we want to approximate that ideal in the empirical human community, we must learn to be truthful and straightforward with one another, regardless of our imperfect autonomy, and the bad results to which it may sometimes lead. In ethics, we cannot always trim our concepts so that they will fit neatly onto the world. Sometimes what we must do instead is try to reshape the world so that it will be more adequate to our concepts.

NOTES

[1] For a discussion of the way various moral theories handle this issue, and of the tendency of utilitarian and Kantian approaches to go to extremes, see Igor Primoratz, "Lying and the 'Methods of Ethics'," *International Studies in Philosophy* Volume XVI (1984) 35—57.

[2] The works of Henry Sidgwick and Immanuel Kant are cited parenthetically in the text. The following abbreviations are used:

ME	Sidgwick, *The Methods of Ethics*
C1	Kant, *Critique of Pure Reason*
G	Kant, *Foundations of the Metaphysics of Morals.*
C2	Kant, *Critique of Practical Reason.*
MMV	Kant, *The Metaphysical Principles of Virtue.*
MMJ	Kant, *The Metaphysical Elements of Justice.*
PP	Kant, *Perpetual Peace.*
A	Kant, *Anthropology from a Pragmatic Point of View.*

For further information and an explanation of pagination see the bibliography at the end of the paper.

[3] See, for example, *Apology* 19e—20c and 24d—25c; *Laches* 186c; *Protagoras* 319bff; and *Meno* 89eff.

[4] In the sense of "right" developed in Chapter V of Mill's *Utilitarianism*: we have whatever rights it is useful that we should have.

[5] Sidgwick rejects Mill's view that some pleasures are better qualitatively than others which are similar or greater in quantity. (ME 94—95, 127—128) If one does accept Mill's view, the problems about paternalism might seem to be worse, for the better pleasures can only be identified by those who have training in and experience of them. It will be better for the rest of us to have this training foisted upon us, but we will only be able to see this after the training. In Chapter III of *On Liberty*, however, Mill argues that these pleasures are best cultivated in the course of choosing a life for oneself.

[6] Interestingly, one of the reasons for concealing the utilitarian news from the general public is that ". . . the concealment [of those actions which are only right when concealed] would in most cases have importantly injurious effects on the agent's habits of veracity."

(ME 490) The enlightened few will also have to take into account the danger to their habits of veracity, as Sidgwick has made clear earlier. (ME 482)

[7] Kant also worries about polite salutations. He asks, "Can an untruth from mere politeness (e.g. "your obedient servant" at the end of a letter) be taken as lying? Nobody is deceived by it." (MPV 431/92–93) Probably both Kant and Sidgwick would have appreciated the irony of the fact that the formula that occupies this place in modern official letters is "sincerely".

[8] Actually, Sidgwick is not even sure of this. The concept of utility is only precise if pleasures and pains are in fact commensurable among themselves and with each other, and this could only be proved if we could have two different sensations at the same time, which is impossible. So "the belief that every pleasure or pain has a definite intensive quantity or degree must remain an *a priori* assumption, incapable of positive empirical verification." The only reason he gives for its acceptance is that "the belief in its general validity is irresistibly suggested in reflection on experience, and remains at any rate uncontradicted by experience." (ME 146)

[9] This reading of the argument for the Formula of Humanity is explained and defended in more detail in my "Kant's Formula of Humanity" in *Kant-Studien*, Band 77, Heft 2 (April, 1986) 183–202.

[10] Even if you happen to know that the person is lying to you, there is a sense in which you cannot assent to the transaction. If you say to her: "I know you don't mean to pay me back, but here's the money," the nature of the transaction is changed. You are not assenting to a false promise, but giving a handout.

[11] This argument is explained more fully in my "The Right to Lie: Kant on Dealing with Evil" in *Philosophy and Public Affairs* Volume 15, Number 4 (Fall, 1986), 325–349.

[12] It follows from this theory of value that some states of affairs cannot be assessed as good or bad directly. The cases I have in mind are those involving competition. If two persons both want the same particular, irreplaceable object (a job, say, or a house), practical reason does not say which outcome is good. It can only say that whatever outcome results from procedures which (according to the categorical imperative) are fair is good.

[13] I should say that it is not clear to me whether Kant always saw that this theory has the troublesome consequences about making these distinctions that I am about to describe. In the *Anthropology*, for instance, he says confidently: "Simple, misguided, stupid, foppish, foolish or offensively silly people differ from the mentally deranged not merely in the degree but also in the kind of their mental disorder . . ." (A 202/74) One wonders what makes him so sure. When discussing the question when someone may be exempt from punishment on grounds of insanity, he claims that philosophers and not the medical faculty must settle the question because it is "purely psychological." But then he says that this is only because of lack of scientific knowledge; there could be a physical cause of the action, but "physicians and physiologists in general have not yet reached a deep enough understanding of the mechanical element in man." (A 213–214/83–84) So I cannot claim that the aspect of Kantianism which I am expounding was explicitly acknowledged by Kant, at least as far as judgments about the autonomy of actual persons is concerned. In the next section, however, I will show that he acknowledged it explicitly about another moral concept, that of legitimate government.

[14] In general, Sidgwick cannot see how the consideration of how much time has passed can possibly make a difference *by itself*. Time is a continuum, and so as in the other cases we have been looking at, dividing points seem to be arbitrary if we cannot determine them by expediency. Another case where this worries Sidgwick is that of sexual purity. He says:

> . . . where divorce by mutual consent, with subsequent marriage, is legalised, we do not call this an offence against Purity: and yet if the principle of free change be once admitted, it seems paradoxical to distinguish purity from impurity merely by less rapidity of transition . . . (ME 358)

If I have a new lover every ten years surely I am not promiscuous. If I have a new lover every ten days perhaps I am. But where, in between, shall we draw the line? Every five years? Every two years? Every month? In his footnote Sidgwick says:

> It should be observed that I am not asking for an exact quantitative decision, but whether we can really think that the decision depends upon considerations of this kind.(ME 358n)

It is clear that he thinks that the temporal considerations must work in connection with "expediency" (utility) before they can make a difference.

[15] Professor Friedrich Bouterwek of Göttingen, in the *Göttingen Journal* (Number 28, February 18, 1979). Kant appended his replies to Bouterwek to the second edition of *The Metaphysical Elements of Justice.*

[16] In a slightly different way, Derek Parfit is concerned about the same problem. He is not concerned about precision as a criterion for self-evidence. But after showing us (stunningly) that our concept of a person is not as metaphysically precise as we might have thought, he concludes that personal identity is "not what matters." See Derek Parfit, *Reasons and Persons.* Oxford: the Clarendon Press, 1984.

[17] But we do not have to be slavishly attached to the use of theoretical concepts. For example: In Kant's theory one of the ideas of practical reason is property. To say something is my property is to say that I am injured, or wronged, if someone uses it without my permission. (MMJ 245/51) It is clear why I am injured if someone wrests something from my hand, so we have an empirical notion of possession as physical detention. But this does not exhaust our idea of property, since if something is my property I am injured if another uses or removes it even if I am not present. This shows that the notion of property is rational and not empirical. (MMJ 250/57) Kant traces the notion to practical reason by an argument that shows that if we could not own things we could not successfully use them, and if we could not use them our freedom of action would be restricted for no reason. (MMJ 246/52—53) But the important thing for my purposes here is the flexibility of the way this concept is imposed on the world. In many cases, the objects we own are ordinary physical objects, and there are good reasons of simple convenience for this. But we don't always own ordinary physical objects. Your right to a library book you have checked out would be a piece of property in Kant's theory (all "acquired rights" are property); but in that case what you own is not the physical object (you may not use it for kindling or tear out the pages for scratch paper), but only the privilege of reading from it for a certain length of time. Or think of the way you own the furniture in your office, or shares in a corporation's stock. The point is that we may divide the world into owned objects quite differently than we divide it into physical objects. And as the case of library books shows, our reasons for dividing it differently may be moral and evaluative. The library loan system makes the "intelligible" book available to more people, and so may be defended either in terms of utility or fairness.

[18] I do not mean to imply that we should just use the *intuitive* notion of say, "being in control of oneself". We should map our practical concepts onto theoretical ones that are as philosophically well-defined and respectable as we can make them. Here the philosopher will have much to learn from the scientist. Those working on applied ethics must learn as much as possible about the conditions of the persons about whom these decisions must be made from, say, psychiatrists, child psychologists, and doctors. The notion of self-control must be analyzed carefully. But it will still never turn out to be adequate to the ideal of transcendental freedom, and so decisions about cutoff points will still just have to be made.

[19] Although Sidgwick does not make this proposal about deliberate concept-formation, there is some suggestion in the *Methods* that he thinks that historically, considerations of utility have influenced the formation of our ethical concepts. (ME 423—459) If this were true it would give one sense in which Common Sense morality would be "unconsciously utilitarian."

BIBLIOGRAPHY

Citations to the works of Kant and Sidgwick are incorporated into the text, using the abbreviations that appear below in boldface type.

Kant, Immanuel.: **C1** (1781/1787), *Critique of Pure Reason.* The first pair of page numbers refer to the first (A) and second (B) editions; the second to the translation by Norman Kemp Smith (1965). Macmillan, St. Martin's Press, New York.

— **G** (1785) *Foundations of the Metaphysics of Morals.* The first page number is that of the Prussian Academy Edition Volume IV: the second is that of the translation by Lewis White Beck (1959). Bobbs-Merrill Library of Liberal Arts, Indianapolis.

— **C2** (1788) *Critique of Practical Reason.* Prussian Academy Volume V; translated by Lewis White Beck (1956). Bobbs-Merrill Library of Liberal Arts, Indianapolis.

— **PP** (1795) *Perpetual Peace.* Prussian Academy Volume VIII; translated by Lewis White Beck in *On History*, edited by Lewis White Beck (1963). Bobbs-Merrill Library of Liberal Arts, Indianapolis.

— **MPV** (1797) *The Metaphysical Principles of Virtue.* Prussian Academy Volume VI; translated by James Ellington in *Immanuel Kant: Ethical Philosophy* (1983). Hackett Publishing Company, Indianapolis.

— **MEJ** (1797) *The Metaphysical Elements of Justice.* Prussian Academy Edition Volume VI; translated (in part) by John Ladd (1965). Bobbs-Merrill Library of Liberal Arts, Indianapolis.

— **A** (1798), *Anthropology from a Pragmatic Point of View.* Prussian Academy Edition Volume VII; translation by Mary Gregor (1974). Martinus Nijhoff, The Hague.

Sidgwick, Henry.: **ME** (1907), *The Methods of Ethics*, 7th edition. Hackett Publishing Company, Indianapolis, 1981.