

## **What's Wrong With Lying?**

**Christine M. Korsgaard**

**Harvard University**

The question I am going to discuss in this paper is: Why is it wrong to tell lies? I am going to discuss three reasons moral philosophers have proposed in answer to this question. I will begin by saying quickly what they are. Intuitionists think that lying is wrong just because it is lying, and for no further reason. Consequentialists think that lying is wrong because it has harmful consequences. And Kantians think lying is wrong because it violates the autonomy of the person to whom you lie.

Before I begin discussing these three views, I want to explain why I think that this question is important. Almost everyone thinks that lying is wrong, that is, that there is a general moral presumption against it. But almost no one thinks that lying is *always* wrong: most of us believe that there are some occasions when a lie is justified or even obligatory. Now what this suggests is that lying is not wrong simply, but that it is wrong for some further reason, and so that we can give an account of what is wrong with it. For if lying is wrong for a reason, then it will be wrong only on those occasions when that reason holds. But there will also be occasions when it is not wrong: occasions when the reason either does not hold, or is overridden. If this is correct, it is of great practical importance, as well as theoretical interest, to know why lying is wrong. Only if you know the reason why lying is wrong will you know when that reason holds, and only if you know when the reason holds will you know when it is wrong to lie, and when it is not. And of course this illustrates a more general point. We need to know what is behind our ordinary moral principles in order

to apply those principles correctly. What I want to discuss, then, is why lying is wrong, with a view to determining when it is wrong, and when it is not.

Proponents of the first view I am going to discuss would be uncomfortable with some of the remarks I have just made. From the late seventeenth century on there has been a school of moral thought according to which moral principles like “lying is wrong” “killing is wrong” “being unfaithful to your friends is wrong” are self-evident truths for which no further reason can be given. These philosophers are called intuitionists, because they believe we know moral truths through direct rational intuition.<sup>i</sup> According to intuitionists, moral principles are simply obvious to any rational person who is thinking clearly. They are obvious in the way that it is obvious that, say,  $2+2=4$ , or in the way that it is obvious that you cannot be in two places at the same time. You can tell that these things are so just by thinking.

Now you will see right away that according to this view, there is not much we can say about what is wrong with lying, any more than there is much we can say about why  $2+2=4$  or why you cannot be in two places at the same time. These are not things for which we normally give reasons, because they are too basic. Some philosophers say that once someone fully understands what these statements mean, he will see immediately that the statements must be true. And the intuitionists say this sort of thing about lying: once you really understand what it is, you will see immediately that it must be wrong. So on this view there is no real answer to the question, “What’s wrong with lying?” The wrongness of lying is basic. The intuitionist says that what is wrong with lying is just that it is *lying*.

You might think that this means that intuitionists are committed to the view that lying is always wrong, or absolutely wrong. But this is not what intuitionists believe. Most

intuitionists distinguish between an action being wrong “*prima facie*” and it being wrong “all things considered.”<sup>ii</sup> To say that something is wrong *prima facie* is to say that it is wrong if everything else is equal. *Prima facie*, lying is always wrong, but all things considered it is not. This is because sometimes telling the truth would require violating one of your other duties. By telling the truth, you may violate a promise to keep a secret; or if the truth hurts, you may violate the duty to be merciful; or if the truth kills, you may violate the duty to preserve life. When two duties conflict, we must violate one of them. If we must violate the duty to tell the truth, lying is not wrong all things considered, although it is still wrong *prima facie*. The intuitionist, then, would be uncomfortable with one of the things I said when I described my project at the beginning. I suggested that lying must be wrong for a reason, and the occasions when it is not wrong must be the occasions when the reason either does not hold or is overridden. Now the intuitionist thinks that lying is not wrong for any reason except the fact that it is *lying*. So the reason against lying is always in force. But the intuitionist would agree that the reason can be overridden, when some more important duty is involved in the case.

Let me point out some advantages and disadvantages of the intuitionist position before I go on. An advantage is that it seems to coincide with common sense views of morality. Many people would agree with the claim that you should not lie unless telling the truth violates some more important duty. And many people would agree with the intuitionist’s model of how we make moral decisions - by balancing and weighing various moral principles. Also, many people would endorse the following claim: even when lying is justified, there is something morally distasteful and regrettable about it. When you tell a justified lie, you usually feel some of the same emotions you feel when you do something

wrong. You feel guilty towards the person to whom you lie and you might even want to apologize when the need for deception is over. The intuitionist tries to explain these feelings by appealing to the idea that although the lie is not wrong all things considered, it is still *prima facie* wrong. He says this is why you still feel bad: because you have done something that is *prima facie* wrong.

The disadvantage of the intuitionist view is that it can tell us nothing about when lying is wrong all things considered and when the duty to tell the truth is outweighed by another duty. And this is because we can give no account of why various kinds of actions are right or wrong. If there are reasons for our duties we can weigh the importance of the duties by comparing the reasons for them. But the intuitionist believes that there is no method for resolving conflicts of duty. An intuitionist thinks that if you are well brought up and have life experience you will become a good judge about which of two conflicting duties is the more important. Now this too may coincide with our common sense views about morality. But the need to depend on experience and judgment may be one of the things about common sense morality that we regret. We might have hoped that moral philosophy could help us to resolve such conflicts. More important is the fact that the view makes it mysterious what we are doing when we compare the importance of two duties. If there is no further reason why lying is wrong or killing is wrong, then how can there be a reason why killing is worse than lying? So it is not really clear what it is that the person who is good at weighing the importance of different duties is a good judge *of*. The problem is not only that the intuitionist lacks a method for resolving moral conflicts, but also that it is unclear what the method would tell you if you did have it.

Now I am going to devote the rest of my discussion to two views according to which there is something a little more helpful to say about what is wrong with lying, views according to which we can give an account of why lying is wrong. To keep things simple, I am going to focus on a particular kind of lie. The occasions on which many lies are told are such that those lies are wrong in a variety of ways. Malicious lies, told in order to hurt people, or lies told in the service of concealing crimes or other wrong actions, are wrong in more than one way, and may be more seriously wrong because of the role that they play in evil projects than they are just *as* lies. I want to know what is wrong, if anything, with lies just *as such*, so I want to avoid these complications. There are two kinds of lies that even decent and morally good people are regularly tempted to tell. These are lies told to protect your privacy, and benevolent lies: lies told for good purposes. Because I am not convinced that lies told to protect your privacy are wrong, I am going to focus on the case of benevolent lies. More specifically still, I am going to focus on the two-person case. I am not going to consider the kind of case in which, say, you lie to me in order to protect a third party. That case too may introduce complicating factors. I am going to discuss the case of paternalistic lies: the case where you lie to me for my own good, because you think it would be better for me to believe, or perhaps just to hear, something other than the truth. Two things make this the best case for us to think about. First, as I mentioned, this kind of case is one of great practical concern to most of us, since this is one of the kinds of lie that we are often tempted to tell. Second, this kind of case is clear of all morally complicating factors. Only two people are involved and the lie is told for a good purpose. Whatever is wrong with lies *just insofar as they are lies* should show up in this kind of case.

The two views I am going to describe are often confused, because as you will see there is a wide range of cases in which they will have exactly the same practical results. But in the most important and difficult cases they will diverge, so it is critically important both to distinguish them, and to discover which of them we really believe.

The first view I will discuss is what philosophers call a consequentialist theory.<sup>iii</sup> According to such a theory, morality is entirely concerned with the consequences of actions, with the benefits and harms that result from them. On this view there are reasons why actions are right or wrong: they are right because of the benefits they produce, and wrong because of the harm they do. I want you to notice something important about this view. This view provides us with a way of deciding what is right only if the ideas of good and bad, or benefit and harm, are objective and determinable: that is, if there are objective facts about what is good and bad and we can ascertain what those facts are. The consequentialist view gives us no guidance about what is right unless we suppose that we know, or can establish, what counts as good and what counts as bad. Only then can we determine empirically which actions will best promote the good consequences and avert the bad ones.

According to this view, what is wrong with lying? The answer has to be that lies are harmful, or, to put it more exactly, that lies tend to do more harm than good.

Now offhand it does not seem obvious that lies always or perhaps even usually do more harm than good. It seems to vary from case to case. So you might wonder why proponents of this theory think that there is even a *general* presumption against lying. And of course paternalistic lies are intended to do good, so if there is something wrong with them it must be that, for some reason, they are usually misguided, and fail in their aim. The people who decide to tell them must be making some common mistake in their assessment of the

consequences. If there is a *general* presumption against paternalistic lying, it must rest upon some quite general empirical fact, which paternalistic liars tend to forget.

I think that most people who accept the consequentialist view would make the following factual claim in order to explain the presumption against paternalistic lies. They would claim that people are in fact usually the best judges of what is good or bad for themselves. Suppose someone asks you for information. She thinks it would be good for her to know something you can tell her. If you lie to her for her own good, it is because you think you know better than she does what it is 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 the consequentialist, you will be wrong as a matter of fact, because she is in the best position to judge what is good for her. But according to this view, paternalistic lies are completely justified and right when someone is not in the best position to judge what is best for herself.

Compare this to the intuitionist view we examined first. You will see right away that it compensates for one of the disadvantages of that view, but at the same time that it loses one of its advantages. The disadvantage that it compensates for is the lack of definite guidance provided by the intuitionist position. If we can determine when we are better judges than other people about what is good for them, then we will know when it is all right to tell paternalistic lies. So the consequentialist can give a definite answer about when we may lie and when we may not.

The advantage that it loses is that it fails to capture the common-sense notion that telling a lie, even when it is justified, is morally regrettable. If the *only* reason that it is wrong to tell paternalistic lies is that people are the usually the best judges of what is for their own

good, then when they are *not* the best judges it is not wrong to lie - *not at all*. I suppose we might think it is regrettable that people are not always good judges of what is best for themselves. The world would be arranged more efficiently if everyone were capable of looking after herself. But this is not a *moral* ground for regret. It does not explain the guilt we feel towards those we lie to. So the consequentialist is unable to explain the feelings of moral distaste we have for telling lies even when they are justified. Or rather, he must dismiss these feelings as irrational, perhaps just the result of the fact that we have been trained to feel badly when we lie.

This disparity with common sense morality is not the only thing troubling about the consequentialist view, however. The real trouble emerges when we compare two of the assumptions used in the consequentialist argument against paternalistic lies. One of the assumptions of consequentialism generally is that we have an objective, empirically determinable notion of what is good and bad, beneficial and harmful. Another is that people are, as a matter of fact, usually the best judges of what is good or bad for themselves. These two assumptions do not harmonize well together. If what is good is an objective, empirically determinable matter of fact, then there ought to be a science of the good and people who are experts in that science. These people would be experts about what is good for others as well as for themselves. But if this is so then it does not seem very likely that people in general will be the best judges of what is good and bad for themselves. There are not many matters of fact about which everyone is equally an expert, or especially likely to be expert in his own case.<sup>iv</sup>

The most important philosophical consequentialists, the utilitarians, thought that they could answer this objection. They believed that what makes a life good is some sort of



psychological state, something that can best be measured and assessed by the person himself. Most utilitarians believe that what makes life good is either pleasure, or the satisfaction of desire, and what makes life bad is pain or frustration. Since people know what they want and what pleases them, they know what is best for themselves. But this answer is not really very good. Its appeal depends on the supposition that you have direct conscious access to your own pleasure and satisfaction *when they are present*. But even if this is true, it does not show that the individual knows best what will give her pleasure in the future or will satisfy her future desires. Furthermore, while the existence of present pleasure and satisfaction may be things about which you are an expert in your own case, this does not make you an expert about the means for bringing these states about. And if you do not agree with the utilitarian that the goodness of life consists in mere pleasure or satisfaction, the idea of individual expertise looks even less plausible. Some philosophers believe that a good human life necessarily includes such things as virtue, knowledge, achievement, and friendship, because these things are objectively good. If this is so, the idea that each person is an expert about how to achieve a good life for herself becomes even less plausible, because we know that people often fail to value these things as much as they should.<sup>v</sup>

So there is a second problem with this view. The view does not provide a very coherent explanation of why paternalistic lies are usually wrong. For on this view, the reason not to tell paternalistic lies is that people are the best judges of what constitutes and promotes their own good. But for consequentialism to work, we must have an objective and empirically determinable notion of what is good. And once we have such a notion, it looks as if it is going to be possible for some people to be experts about the good life. The

assumption that people are the best judges of the goodness of their own lives will no longer seem plausible, and the presumption against paternalistic lying will be left unsupported.

If this is correct, the consequentialist view leaves too much scope for telling paternalistic lies. As I have said, it is not even clear that there is a *general* presumption against them. Yet most of us think that there is. When somebody lies to *you* for your own good, and you find out about it, you usually think the liar is a presumptuous busybody, and you resent his action. Paternalism is considered out of line when we are dealing with normal sane and healthy adults.

Before I go on, let me deal with one argument that people commonly make at this point. You might now feel tempted to say that we can argue for the presumption against paternalistic lying this way. I should not tell a paternalistic lie because in order to tell such a lie I must know myself to be an expert about the good life, but in fact, who is to say who is an expert about the good life? Who is to say what is best for a person? There are several ways to reply to this argument. One I have already mentioned. If there is a matter of objective, knowable fact about what makes a life good, as the consequentialist must suppose, then there should be a science of the good life, and we will identify its experts in the usual way we identify the practitioners of a science. Another point is that the claim that we cannot tell who the expert about the good life is will not salvage the consequentialist argument against paternalistic lying, because the consequentialist argument against paternalistic lying appeals to the assumption that each person is the expert in his own case. We cannot let the consequentialist claim *both* that each person is the expert in his own case *and* that there is no way to tell who the expert is. If there is no way to tell who the expert is, then the

paternalistic liar and his victim, are, as far as we know, equally likely to be the expert or to make mistakes.

Now at this point the consequentialist may want to change his tack slightly. He will continue to insist that for some reason there is no reliable way to tell who the expert is, but he will add that, since we cannot tell, it is preferable that people decide for themselves what is best for them. Even if the individual whose good is in question makes a mistake, this is less bad than if the paternalistic liar makes a mistake. The idea behind this claim is that it is less bad to make mistakes on your own behalf than on behalf of another, just as it is less bad to squander your own fortune than one you hold in trust. This, he will claim, is the source of the presumption against paternalistic lying.

But the fact is that the consequentialist theory affords us no grounds for making this kind of claim. Consequentialists do not care who makes the mistakes but only how bad they are. The idea that it is better for people to make their own mistakes really comes from our third view. I will call this third view the Kantian view, after the philosopher from whom I get my account of it, Immanuel Kant.<sup>vi</sup> The Kantian view is that each person *has a right to decide* what is good for herself. Each person has a right to decide what she will count as a good human life, and to live in accordance with her own conception of the good. The consequentialist view gains some illicit plausibility from its similarity to this view, so let me underscore the difference: according to the Kantian view, it is not that there is some determinate notion of what is good, and that each person is the expert about this in her own case. It is that each person *has a right to decide* what counts as good in her own case because it is her own case. It is not a question of an empirical fact about who has expertise, but rather of a right you hold simply on moral grounds – because you are the person whose life it is.

People tend to confuse these two accounts of why each of us has authority over his own life. That is, we tend to confuse the factual claim that people are experts in their own case with the moral claim that people have a right to decide about their own case. I think that this makes the consequentialist view look more compelling than it really is.

I have claimed that the consequentialist view and the Kantian view are often confused with each other. This is because adherents of both views claim that the basis of the presumption against paternalistic lying is the importance of people running their own lives. When we lie to someone for her own good, we take her life out of her hands, and into our own. Both the consequentialist and the Kantian think this is a bad thing, although both also think that there are some cases in which paternalistic lying is justified. But the Kantian view turns out to be much stricter, in the sense of permitting fewer such lies, than the consequentialist view. For despite the similarity in the accounts these two views give of why it is wrong to tell paternalistic lies, the difference is critically important. The consequentialist thinks people should run their own lives because of a factual condition that often fails to hold: they are experts about their own good. The Kantian appeals not to a theoretical fact, but to a pure moral claim: a person has a right to run his own life because he is the person whose life it is. Possession of this moral right does not depend on whether people are especially good at running their own lives or not. For a Kantian, the reason why paternalistic lies are wrong is always in force, even if it is sometimes overridden, whereas the consequentialist reason may often simply fail to hold.

Some comparisons will help here. Consider the way we might argue for the universal franchise, for everyone's right to vote. Sometimes the argument is based on empirical claims about the effectiveness of the universal franchise in selecting good rulers. For instance, one

may argue that when everyone pools their wisdom about government, individual biases, distortions, and limitations are corrected for: the group makes a more intelligent choice than any individual would. Or one might argue that if each person votes in a way that represents her own interests, the candidate selected will represent majority interests. Supposing that this is a good thing, the argument is that the group makes a better choice than any individual would. Both of these arguments rest in part on empirical or factual claims: they suggest that the voting public is a good judge of, or an expert on, the quality of its rulers. These arguments are like the consequentialist argument against paternalistic lies. They say that the public should be in charge because of the public's collective political expertise. A different kind of argument appeals to the fact that since everyone lives in this country and is affected by its policies and its fate, everyone has the *right* to vote for its rulers. It does not matter whether we are especially good at choosing competent rulers or not. Since it is our society, the choice of its rulers rightly rests with us. This is like the Kantian view of why we should not tell paternalistic lies. The Kantian view says you should run your life because it is *your* life, just as this view says you should have a say about your country's rulers because it is *your* country. The Kantian view is based on a pure moral right, not a doubtful empirical fact. The idea that the public has collective political expertise could be wrong. Some of us think that there is evidence that it is wrong. But the claim that everyone has a right to a voice in the political decisions that affect his own life is a moral claim, not a factual one.

Here is another case that has the same structure: the defense of the jury system. One might try to argue that the twelve people together have some kind of collective expertise on justice. When the twelve people pool their wisdom, the collective effect overcomes the distortions, biases, and limitations of individuals. This is a claim of empirical

fact. On the other hand, we know that people who actually participate in juries sometimes come out feeling that it did not work this way at all. Sometimes they feel that the collective effort produced a good judgment, but at other times they come out feeling that making a good collective decision is uphill work against obtuseness and prejudice. There is a quite different argument available for the jury system, which is that when someone is accused of a crime against the community, it is morally right that representatives of the community should decide the case. The jury makes the decision not because it is especially good at making such decisions, but because the jury represents the community and the community has the right to judge the offender. This kind of argument is a moral one and does not depend on the empirical claim that a jury has some sort of collective judicial expertise.

The jury system and the universal franchise are institutions of democracy. They make a useful introduction to Kantian ideas because Kantian ethics is based on a view of how human beings should treat one another that is deeply democratic.

Kant argued that you must never treat a human being, either yourself or another, as mere means to an end, but always, as he put it, also at the same time as an end in himself.<sup>vii</sup> The idea initially sounds puzzling. Obviously, Kant is not saying that we should never use a human being as a means at all, if a means is something that furthers our enterprises. We use each other as means all the time. The clerk who checks you out in the grocery store, the teacher who explains what you need to know, and the friend who gives you a ride to the airport all further your enterprises and so are means to your ends. But there is nothing pernicious about these relationships. And indeed, Kant insists on this himself when he says you must not treat a person as a *mere* means. What makes the difference between treating someone as a means in a permissible way and treating her as a *mere* means is whether you

also treat her at the same time as an end in herself. But what does this mean? Normally, we think of an end as a purpose, an aim, something we try to bring about. Of course it makes no sense to talk about treating a human being like that. But an end is also something that we do not violate, something that we do not act against. And this is more what Kant has in mind.

In fact, the distinction Kant has in mind is that between respecting someone's autonomy, or her capacity for self-government, and treating her as if she were merely a tool you may use to promote your own ends.<sup>viii</sup> To respect someone's autonomy, not to violate it, is to treat her as someone whose beliefs and actions are, and should be, controlled by her own reason. It is to allow her to decide for herself what to do and what to think and what to be. Kant thinks that respect for autonomy requires not only a certain way of acting towards other people but even a certain way of thinking about them, a way that you will recognize as being in the ordinary sense "respectful." For example, everyone knows how enraging it is when someone who disagrees with you simply assumes that your view must be the result of prejudice or passion, or even that since you disagree with him you must be stupid.<sup>ix</sup> These attitudes are disrespectful. We expect people to talk to us, and even to think of us, in a certain way: we expect people to deal with us on the assumption that we have reasons for our views, even if they disagree with those reasons. And we expect them to acknowledge our right to act on those reasons. If someone disagrees with you and you want her to agree, you may try to change her reasons by arguing with her, but you must not try either to bully her or to trick her into thinking or acting as you think best. This is true even if it is very important what she thinks. Suppose you are arguing about candidates and she is going to vote. If she among others votes wrongly, a bad candidate will be elected, and if it is to a

powerful office, the results can be serious indeed. But this does not make it permissible to bully or trick her into voting the way you think best, any more than it makes it permissible to steal her registration card and vote in her place. She has a right to a vote: this means she *has a right to decide* for whom she will vote. In a democracy, poor judgment does not disqualify a person for citizenship. In Kant's theory, it does not disqualify her for the respect due to every rational being.

The requirement of respect for autonomy affects your treatment of others most directly when you are involved in a cooperative project, when you must act together. In our dealings with others, we often enlist their assistance towards the achievement of our own ends. We are to this extent asking them to contribute to our ends. So the requirement of respect for autonomy says that others have a right to decide whether or not to contribute to our ends. And it follows that we must treat them in a way that *enables* them to make that decision.

Let me use one of Kant's own examples to illustrate the point.<sup>x</sup> This is an example of a lie, but not of a paternalistic one. Suppose I am in need of some ready money, and I think perhaps I may get it from you. Since you have cash flow problems of your own, however, I think it is unlikely that you will give me the money unless you believe that I will pay you back next week. Actually, however, my financial situation is so bad that I will not be able to pay you back in the foreseeable future at all. But I don't think you will give me the money if I tell you the truth about that. So I go to you and say, "lend me \$50, would you? I promise to pay you back next week." I make a lying promise.

What exactly is wrong with this? On Kant's analysis, it is that you do not get to decide whether you will contribute to my end. You cannot decide whether to contribute to



my end because you do not know what it really is. You think that the result of our transaction, the result to which you are being asked to contribute, is my temporary possession of your money. But actually the result of our transaction will be my indefinite possession of your money. You have been tricked into contributing to an end different from the one to which you thought you were contributing. I am treating you as a tool rather than as a rational being, in this sense: what I say to you is just whatever I think will *work* to produce the result I want. I do not consider the fact that since you are contributing to this end you also, so to speak, have a right to vote on whether this end will be brought about or not. I treat you as a tool because, when I decide what to say to you, I think only about what levers to pull to get what I think is a good result. The Kantian view is deeply democratic, as I said earlier, in the sense that it implies that every rational being has a right to cast a vote about any end to which her own actions are to contribute.

But now notice an important thing: this analysis of what is wrong with the lie holds even if, contrary to what I thought, you *would* have lent me the money if you had known the truth. Indeed it holds even if you specifically want me to have the money, and you were planning to offer it to me. Whether you can share in my end depends not on whether you *would* have contributed to it if you had known what it really was, but on whether you *actually* had the chance to decide whether to contribute to it or not. Consider a comparison. Suppose I steal your voter registration card and vote in your place. When you remonstrate with me, I say, “Oh, it’s all right, because I voted for the candidate you would have voted for anyway.” Now, maybe this will make you feel somewhat less upset. But it does not make what I did “all right.” You will still feel that I have wronged you, even though the consequences are the same as if I had not. In the same way, what is wrong with my lie does

not rest in its consequences, but in the fact that I have taken the decision out of your hands. That is just as wrong whether I make the decision the way you would have made it or not. And of course this means that the analysis applies to paternalistic lies as well. The lie is no less wrong if the end is one that you think is good for me. In fact since it is my own good that is involved and I have a special right to decide what is good for myself, paternalistic lies are in a way worse than others.

More generally, there are two conditions under which your autonomy is violated. One is when force or coercion is used to make you contribute to an end. The other is when lies are used to trick you into contributing to an end. In both cases what is wrong is that you do not get to decide whether to contribute to the end or not. The conditions under which you are able to decide for yourself are that you have power over your own actions and knowledge of what is going on. Force and coercion, on the one hand, and lies, on the other, undercut these conditions. And so force and coercion and lies are, according to this view, the most fundamental forms of wrongdoing - the roots of all evil. Morality demands that we resist the ever-present temptation to *manage* things ourselves, and instead share our decisions - and so our knowledge and our power - with all who are concerned.

When I was discussing the consequentialist view, I suggested that questions like “who’s to say what’s really best for someone?” really get their rhetorical force from the Kantian view. Kant did not believe that we have an empirically determinable notion of what is good in a human life. But he did not deny that some things are better for people than others. He believed something more like this: you may think that you know what makes a human life good, but you have nothing to go on but the authority of your own reason. There are, after all, no independent tests for the correctness of views about what makes a

human life good. If whatever authority your conception of the good has comes from your own reason, you must concede the same authority to anyone else's reason. And this is why you cannot lie to another in order to control the results of her actions, even, or rather especially, when it is for her own good.

On the Kantian view, the presumption against paternalistic lying is very strong. If someone asks you for information pertinent to her own life, it is because she has decided that she should have that information, that it is good for her to know. If you lie, you take it upon yourself to decide that it is not good for her to know. But she is the one who has the right to decide what it is good for her to know, where the information is relevant to her own life, and you are not. Perhaps the information she requests is painful or tragic. She may or may not have realized that, but you must assume that because she asks she has decided it is worth it to her to know the truth. It is not for you to say that a deluded pleasure is better for her than an honest grief. You may make this choice for yourself, although I think that when it comes down to it hardly anyone ever does. Most people want to know the truth about their own situation, and tend to reject paternalism when they imagine themselves as the objects of it. But even if you would make this choice for yourself, you cannot make it for someone else.

Although the presumption against paternalistic lies is strong, they are not excluded altogether. I said earlier that Kant's view makes lying and force and coercion, as the main ways of violating people's autonomy, the most fundamental forms of wrongdoing. But we do think the paternalistic use of force is sometimes justified, so we will also think that there is such a thing as a justified paternalistic lie. But the view is still very strict, for lies will be justified only in the same kinds of cases as force would be. Yet many people in fact tell

paternalistic lies in circumstances in which they would not dream of resorting to force. They tell lies to people who are merely upset or vulnerable, for instance, when they would only use force on someone who was out of his mind. According to Kant's view, they are wrong. Lies and force are wrong in exactly the same way.<sup>xi</sup> They violate people's autonomy. You cannot tell lies just to make people feel good, or to divert them from what you think are poor decisions, any more than you could use force for those reasons. And just as it is nearly always morally regrettable to use force on a person, so it is nearly always morally regrettable to lie, because it is regrettable to use a human being as a mere tool, even as a tool for his own good.

When exactly is paternalism justified? The answer is that telling paternalistic lies, like using paternalistic force, is justified only when we are dealing with people who are incapable of exercising their own reason, or severely crippled in the use of it by some condition like insanity or drunkenness. And the best justification for telling paternalistic lies, apart from protecting the incompetent person from physical harm, is developing or restoring his autonomy. For example, the encouraging lies we tell to children about the quality of their performance meet these two criteria. Children are rationally undeveloped, and encouraging lies help them develop the confidence they will need for autonomous adult life. But notice that we also use force on children to an extent that would be outrageous in dealing with adults. We ought to feel the same way about telling lies to adults. Using paternalistic force on adult human beings is justified only in cases involving immediate threats to life and limb or severe mental illness. So the same should be true of paternalistic lies. Only when people are severely incapacitated for making their own decisions may other people interfere. And of course it is important to add that it is cheating, or at least losing your nerve, if you treat

the fact that someone makes a decision that you disagree with as *evidence* that she is severely incapacitated for making her own decisions.

Let me illustrate this last point with a real life story that a physician told me about another physician. This physician told a patient that he had a terrible illness and was probably going to die. The patient went home and committed suicide. The physician felt that she had been “burned” and resolved not to tell the truth about such matters again in the future. Let us ask ourselves what this physician was thinking. One way to understand the physician’s decision is this. Perhaps she is a consequentialist. She thinks that she’s an expert on the goodness of life, and she thinks that she knows that committing suicide is bad, even when it is done to avoid living through a terrible and probably hopeless illness. So she takes the incident as evidence that telling people the truth about their illnesses tends to produce a bad result, and therefore is wrong. Another way to understand the physician’s attitude is this. The physician is a Kantian, but one who has lost her nerve. She believes that people ought to make their own decisions and so that they ought to know the truth about their own circumstances. But, just because she disagrees with this particular decision, she concludes that the patient who made it was incompetent to make decisions, and perhaps that all seriously ill people are incompetent to make decisions. Kantianism suggests another attitude is possible here and may even be required. The physician should not have decided that she was “burned.” She should have allowed that the patient had the right to choose a quick suicide over a lingering illness. The patient might have felt that he had a good reason to commit suicide, that it was better for him than living through the illness. And if so, his decision, and his right to make it, must be respected.

Of course this does not mean that you can never withhold the truth, nor am I saying that my description of this actual patient's mental state is obviously correct. Perhaps he was rendered mentally incompetent by the news of his illness and killed himself in a moment of insane despair. But my description is not obviously wrong either: I do not know whether the physician in question had any independent reason for believing the patient was incompetent to make the decision, apart from the fact that she disagreed with the decision he made. The point of the example is not to say that you must always tell the truth. The point is twofold: first, you have the right to withhold the truth about someone's life when he requests it only when you have some evidence of mental incompetence or irrationality. You cannot withhold the truth just because it will hurt, or because you do not approve of what you think the person will do once he knows it. Second, you should not treat the bare fact that someone disagrees with you as evidence of mental incompetence or irrationality.

It is a hard question which of these three views about what is wrong with lying is the correct one. I am going to close with some speculations that are not directly about which view is correct, but about what I think most of us really believe. I think that elements of all three views appear in ordinary moral discourse about the subject, but that in fact not many people believe that lies are just wrong, and that no reason why they are wrong can be given. So I do not think that many people share the intuitionist's views. In the case of paternalistic lies, especially, we find that people often jumble together the empirical claim that people are experts about their own good with the moral claim that people have a right to decide what is for their own good. The same phenomenon occurs with respect to the two cases I used as comparisons: the defense of the universal franchise and the defense of the jury system. I will offer a psychological speculation about why we tend to jumble these arguments

together. I do not think it is just a product of intellectual confusion. I think that most of us in fact believe something more like the Kantian account, but find it, for two reasons, intimidating to defend our decisions and opinions on the basis of it. One is that people find it intimidating in general to defend their decisions and opinions on the basis of their moral theories, and many of us are relieved when we can make our case by appealing to “hard” empirical facts instead. Although nowadays people tend to exaggerate the indefensibility of moral views, correct moral positions are certainly harder to establish than empirical facts. So the view that individuals, or voting publics, or juries should make certain kinds of decisions because of their expertise is a more comfortable one to defend than the view that these agencies just have the right to make these decisions, regardless of expertise. And since the argument that individuals, and voting publics, and juries are experts usually leads to the same conclusion as the moral position most of us really hold, we prefer to argue from these safer grounds. If you doubt that, imagine this: a political scientist proves to us, on the basis of extensive research, that a certain panel of experts would choose better rulers than the voting public does. He proposes that we should give up our votes and turn the selection of our rulers over to these experts. If this scenario distresses you – if you find yourself hoping that his results can be disproved – then your commitment to the universal franchise is not based on the idea that the public is a collective expert. If it were, and you learned that the voting public is not the most qualified expert, you would simply give up your support of the universal franchise, without any regret. The fact that most of us would find the political scientist’s conclusions distressing shows that our commitment to the universal franchise comes from elsewhere. I suggest it comes from the moral view that the public *has a right to decide* who its rulers shall be. The idea that the public is good at choosing rulers does not

really explain our position, but it does prop it up, and make it a less alarming one to hold. The same is true in the case of paternalistic lies. If you feel *relieved* when someone argues that people should make their own decisions because people usually *are* experts about their own good, it is because that consideration props your commitment to letting people run their own lives, not because it explains that commitment.

There is a second, although related, reason why the Kantian position can be a frightening one to hold. It is frightening to stand by and watch a person make a mess of his life, and so it is frightening to think that morality might require us to do that. On Kant's view you may argue with such people, but if you cannot convince them you may not resort to tricks or force. But it is also frightening to stand by and watch the public elect a candidate with a poor grasp of economics or of nuclear science, and it is frightening to stand by and watch the question of an accused person's fate be settled by twelve ordinary, uninformed, biased, and sometimes unintelligent citizens. We do willingly stand by and let voting publics and juries make these decisions, however, and it is because there are cases in which what we really believe is that it is just as important who has the right to decide as it is whether the decision is a good one.

And the fact is that most of us do not believe that anyone is an expert about what constitutes a good life for a human being, and therefore our view that people ought to be in charge of their own lives *cannot* be based on the consequentialist theory that people are experts about their own good. It is, rather, based on the idea that the problem of what is worth doing, worth having, and worth knowing in a human life is one that every rational human being has the right to solve for herself. Interference with that right, no matter how



well-intentioned, is in the deepest way disrespectful, and is almost always unwarranted. This is why is telling paternalistic lies is almost always wrong.

---

<sup>i</sup> There have been many intuitionists, especially in the tradition of English-speaking philosophers. Some of the major ones include Samuel Clarke, *A Discourse Concerning the Unchangeable Obligations of Natural Religion* (London, 1706); Richard Price, *A Review of the Principal Questions and Difficulties in Morals* (London, 1758); William Whewell, *The Elements of Morality, Including Polity* (Cambridge: 1845); W. D. Ross, *The Right and the Good* (Oxford: Clarendon Press, 1930); and H. A. Prichard, *Moral Obligation and Duty and Interest. Essays and Lectures by H. A. Prichard* (edited by W. D. Ross and J. O. Urmson. Oxford: Oxford University Press, 1968.) Selections from Clarke and Price may be found in D. D. Raphael's two-volume *British Moralists 1650-1800*. Indianapolis: Hackett Publishing Company, 1991.

<sup>ii</sup> I get the term “*prima facie*” from W. D. Ross.

<sup>iii</sup> Consequentialism was introduced into philosophy independently by William Paley in *The Principles of Moral and Political Philosophy* (London, 1785) and Jeremy Bentham in *An Introduction to the Principles of Morals and Legislation* (London, 1789). Other influential presentations of consequentialism include those of John Stuart Mill, in *Utilitarianism* (1861) (Indianapolis: Hackett Publishing Company, 1979), Henry Sidgwick in *The Methods of Ethics* (1<sup>st</sup> edition, 1874, 7<sup>th</sup> edition 1907. Indianapolis: Hackett Publishing Company, 1981); and G. E. Moore in *Principia Ethica* (Cambridge: Cambridge University Press, 1903). Selections from Paley and Bentham may be found in volume two of D. D. Raphael's *British Moralists 1650-1800*.

---

<sup>iv</sup> The Greek philosopher Socrates, the first moral philosopher in the Western tradition, frequently noted this difference between ethics and technical subjects: that there do not appear to be any experts in ethics. He sometimes suggests this is evidence that ethics is not a technical subject, which puts him at odds with the consequentialists' way of thinking. See for instance Plato's dialogue *Protagoras* (trans. Stanley Lombardo and Karen Bell. Indianapolis: Hackett Publishing Company, 1992), pp. 13-14, 319b-320a in the "Stephanus numbers" commonly used in scholarly references to the works of Plato, which indicate the page and section of the relevant volume of the Greek text of Plato as edited by the French scholar Henri Estienne (in Latin, Stephanus).

<sup>v</sup> Of the consequentialists I mentioned in note 3, all were utilitarians except G. E. Moore, who believed in the objective goodness of certain complex states of affairs involving love, friendship, the appreciation of beauty, and so on. John Stuart Mill, however, believed that some pleasures are better in quality (not just quantity) than others, and that we do have to consult experts to determine which ones are better.

<sup>vi</sup> Two of Kant's main ethical writings, *Grounding for the Metaphysics of Morals* (1785) and the Part II of *The Metaphysics of Morals, The Metaphysical Principles of Virtue* (1797), may be found in Immanuel Kant: *Ethical Philosophy*, trans. James Ellington. (Indianapolis: Hackett Publishing Company, 1983). His third major work on the subject is the *Critique of Practical Reason* (1788) trans. Werner Pluhar (Indianapolis: Hackett Publishing, 2002).

<sup>vii</sup> See *Grounding for the Metaphysics of Morals*, *op. cit.*, p. 36; p. 429 in the marginal numbers found in most translations, which refer to the pages in the relevant volumes of the standard German edition of Kant's complete works, *Kant's Gesammelte Schriften*, edited by the Royal

---

Prussian (later German) Academy of Sciences (Berlin: Georg Reimer, later Walter DeGruyter & Co: 1900-).

<sup>viii</sup> Or perhaps merely as an obstacle you can run roughshod over. That kind of treatment also violates the Kantian requirement.

<sup>ix</sup> See *The Metaphysical Principles of Virtue*, op. cit., pp. 128-129; pp. 463-464 in the Prussian Academy edition.

<sup>x</sup> *Grounding for the Metaphysics of Morals*, op. cit., p. 37, pp. 429-430 in the pages of the Prussian Academy edition.

<sup>xi</sup> In fact it is arguable that lies are a *deeper* violation of another's autonomy, since when you lie you do not merely ignore the fact that your victim has his own reason, you make use of it: if he were not a rational being, lying wouldn't be a way of getting him to do something. On the other hand, the use of force may be wrong in an *additional* way, since if it is not used carefully the person who is forced may be injured, and using force means risking that possibility.