

Ethical theories and their application

Forcehimes, Andrew T.

2017

Forcehimes, A. T. (2017). Ethical theories and their application. S. M. Cahn & A. T. Forcehimes (Eds.), *Exploring Moral Problems: An Introductory Anthology* (pp. 2-48). Oxford University Press.

<https://hdl.handle.net/10356/155750>

© 2017 Oxford University Press. All rights reserved. This book chapter is made available with permission of Oxford University Press.

Downloaded on 11 Oct 2024 20:26:19 SGT

Ethical Theories and Their Application

Andrew T. Forcehimes

§1. *Introduction*

Your life consists of a series of actions. You do mundane things. You brush your teeth and buy cups of coffee. You do momentous things. You fall in love and have a child. Mundane or momentous, you have no doubt thought about whether what you did is, in point of fact, what you ought to have done. Think, for example, about something you did that you deeply regret. (Take a moment to actually do this.) When thinking about this regrettable action, you are, inevitably, having two very different kinds of thoughts. You are thinking about what happened. You are having *descriptive* thoughts about what was the case. But, insofar as what you did was regrettable, you are also thinking about what should have happened. You are thus also having *normative* thoughts about what ought to have been the case (but wasn't) or what you were required to do (but didn't). Here—in the normative domain—is where ethics resides.

Although you've confronted ethical questions, chances are your answers have not formed a consistent set. Return to your regret. You think that you ought to have acted otherwise. This is a verdict about your action. Now if I asked you why you reached this verdict, you could probably point to certain facts that made the act one you ought to have refrained from performing. Perhaps you harmed someone, broke a promise, were untrue to yourself, or the like. If you are able to say what you ought to do and why you ought to do it, then you have the beginning of an ethical theory. But a mere beginning is insufficient. If you are to stand any chance of thinking about your life in a consistent and coherent fashion, then what you need is a *systematic*

account of what you ought to do and why you ought to do it. You, in other words, need an ethical theory.

This introduction will prepare you for the intense ethical inquiry that takes place in this volume. Ethical inquiry, roughly, proceeds in two broad steps. First, we start with easy ethical cases—ones that we have high confidence that we are getting right—to help us build a theory. Second, we then use this theory to help us work out what to do in the hard cases—ones that we, pre-theoretically, are unsure about. The readings in this volume concern this second step: They deal with the hard cases, like abortion, terrorism, world hunger, and so forth. This introduction concerns the first step. We will first look at what any ethical theory tries to accomplish, then we will survey the dominant positions.

Being familiar with the process of theory construction as well as having the ability to recognize popular extant theories will prove an invaluable resource as you read through this volume. In trying to defend a position on a particular topic, most authors, usually only implicitly, rely on their preferred ethical theory. Your ability to recognize their reasoning as stemming from a particular theory will thus not only make their position more readily intelligible, but also make it easier to locate potential weaknesses. Relying on a particular theory entails that one inherits the problems that come with the theory. So, by being aware of the general objections faced by a theory, you will be able to quickly see the problems that beset specific applications of it.

A final remark before we begin. Ethics, as you shall see, is a trying endeavor. Like all areas of philosophy, the ideas can be complicated and abstract. And like all areas of philosophy, it is often unclear just which theory we should, in the end, believe. Ethics is intellectually taxing. But unlike most other areas of philosophy, we cannot comfortably take the time to sort out our thoughts. Right now, we are living. We have to act. And what we do matters. Ethics is thus taxing in an additional, special way. When we get things wrong, we do not just hold false beliefs. We live disastrous lives. When we are plagued by doubt, we do not just confront the discomfort of uncertainty. Our casting about in the dark might turn out to be deeply morally regrettable. That thought haunts me. And it should, I believe, haunt you too.

Let's thus turn, with some urgency, to figuring out the thing to do.

2. What Is an Ethical Theory?

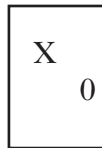
Ethical theories have two main aims. The first is *enumerative*: identify those acts that we ought (or ought not) to perform. The second is *explanatory*: provide an account as to why we ought (or ought not) to perform the acts identified. Let's take each of these aims in turn.

2.1 What You Ought to Do

We can start with the enumerative aim. An ethical theory tries to articulate a general principle that tells us the status of the various actions we could possibly face. There are four main assessments this principle might deliver: *impermissible*, *permissible*, *optional*, and *required*. These assessments are called *deontic verdicts*, because they tell us our various duties. (*Deon* is Greek for “duty.”)

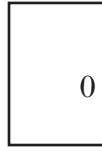
Having a handle on what these four deontic verdicts mean is crucial to understanding what a theory demands. One way to do this visually is by representing each of the actions you could perform at a given time as a door you could proceed through. Say, right now, you have ten possible actions you could perform. Then we could imagine you standing in front of ten doors, each representing a course of action you could take. Now to visualize what a theory says about each of these actions, we could further imagine that some doors are marked with an X. This X indicates that proceeding through the door—performing this action—is, according to the theory, impermissible. With this picture, we can now represent the verdicts a theory might issue as follows:

- An action is *impermissible* if and only if refraining from the action is required. This is a door with an X on it, telling you not to proceed.



The door represents a course of action that you could take, but ethically you shouldn't.

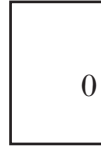
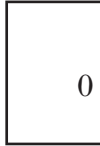
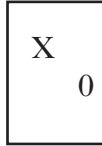
- An action is *permissible* if and only if it is an action that is not impermissible. This is a door without an X on it, telling you it is Okay to proceed.



This is a course of action that you could take, and ethically it is open for you to take it.

- An action is *optional* if and only if it is permissible to perform or not perform the action. Suppose, for simplicity, that you only have three actions available to you. One of these acts is impermissible; the other two are permissible. So the choice set you face looks like this:

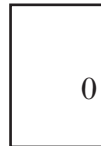
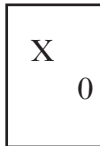
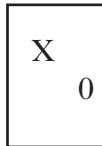
Act 1: Impermissible Act 2: Permissible Act 3: Permissible



Since you can permissibly perform *Act 2* or *Act 3*, both of these acts are *optional*. You need to make sure you refrain from performing *Act 1*, but ethically the rest is up to you.

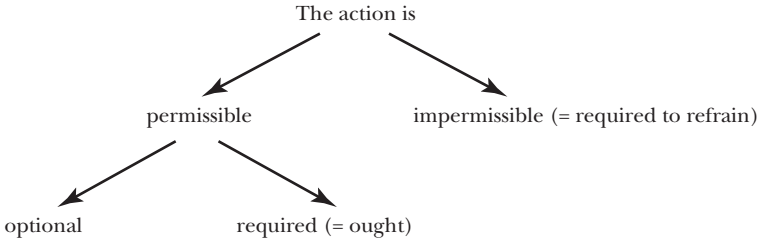
- An action is *required* if and only if it is the uniquely permissible action available—a permissible action that is not optional. Suppose again that you only have three actions available to you. But this time two of these acts are impermissible; the remaining act is permissible. So this new choice set looks like this:

Act 1: Impermissible Act 2: Impermissible Act 3: Permissible



Since you can only permissibly perform *Act 3*, this act is *required*. All of the other acts open to you are ethically blocked.

Notice how these verdicts relate to one another. For example, any action that is required will also be permissible. The below diagram illustrates these connections.



The definitions for our four deontic verdicts have been partly defined in terms of one another. That’s unsatisfying. Nevertheless, we can make some headway toward more satisfying definitions if we introduce the concept of a *reason*: a consideration that counts in favor of, or justifies, our acting in certain ways. As, for instance, when we say, the fact that casting insults harms others gives you a reason not to cast them. The fact—that casting insults harms others—counts in favor of your not casting them.

By deploying the concept of a reason, we can break out of our definitional circle by analyzing “required” in the following, more informative way:

An action is *required* if and only if there is decisive reason to perform the action. The set of reasons in favor of the action outweighs the set of reasons against it.

Defining “required”, and thereby our other deontic verdicts, in terms of reasons fits well with a fairly natural way of thinking about what to do. When thinking about whether we ought to perform some action, we look to all the considerations that count in favor of acting in this way and all of the considerations against it. We then treat these considerations like weights on a scale. If the pros outweigh the cons, then we are required to perform the action. Think back to your regrettable action. There were the pros: that you wouldn’t be embarrassed, that you could help out a friend, and the like. And there were the cons: that it would harm someone, that you would break a promise, and the like. Since regrettable, the cons outweighed the pros. You did what you were required to refrain from doing—what the balance of reasons failed to support.

Defining our deontic verdicts in terms of reasons sits well with what ethical theories are in the business of doing: namely, locating the considerations that count in favor of our actions. Ethics is, at base, justificatory. By citing reasons, we are not trying to explain how the world is; rather, we’re trying to justify how it should be.

2.2 *Why You Ought to Do it*

In citing the reasons that deliver our deontic verdicts, however, we've made a subtle shift. We've moved to giving a specific justification—e.g., the fact that the action will harm others—for *why* an act has the deontic status it does. This brings us to the second, explanatory aim of an ethical theory. Why do acts have the deontic statuses that they do? The answer to this question tells us which facts are genuinely reason-providing. If, for example, you think you ought not have performed the regrettable action *because* it caused harm, then you take harm facts to be one of the factors that determine the deontic statuses of actions. It is, at least in part, because of the harm it would cause that you are required to refrain from acting in this regrettable way.

To get a better handle on this explanatory aim, it will be helpful to start with a broad taxonomy of the types of accounts on offer. And the best way to categorize the various approaches is to think about how you would construct a theory yourself. In constructing an ethical theory, what you would likely do is think very carefully about your life, noting the various features that seem ethically relevant. Next you would try to identify some central idea by which all of these features could be explained.

Here is how we might start such a list:

<u>Deontic Claims</u>	<u>Evaluative Claims</u>	<u>Character Claims</u>
Costlessly saving lives is required.	Pain is bad.	Courage is a virtue.
Favoring your children is permissible.	Injustice is bad.	Modesty is a virtue.
Murdering is impermissible.	Pleasure is good.	Benevolence is a virtue.
Torturing is impermissible.	Equality is good.	Malice is a vice.
Stealing is impermissible.	Beauty is good.	Vanity is a vice.
Lying is impermissible.	Getting one's due is good.	Envy is a vice.

Obviously, this list needs to be greatly expanded. But this gives us a start. The items on this list are symptomatic of the ethical life. We now need a diagnosis that explains all of these symptoms (or explains them away).

As this analogy suggests, constructing an ethical theory is very similar to how a doctor comes up with a diagnosis. If we were doctors, we would first look carefully at our patient, noting the numerous symptoms. We would then offer an account that we believed best explains

why our patient displays these symptoms. Diagnostic medicine is, of course, a descriptive enterprise. Nevertheless, insofar as the process of inquiry is concerned—trying to offer a unified account of the data—the analogy to theory construction in ethics is apt. A diagnosis explains why a patient's body displays the symptoms it does. An ethical theory explains why a life displays the normative features it does.

By recognizing how theory construction begins, we can classify the various theories by asking: What central idea explains the various features of our ethical lives? Three different answers to this question are prevalent among ethicists. There are those that take the central idea to concern the world we create—*consequentialists*. There are those that take the central idea to concern the actions we perform—*deontologists*. And there are those that take the central idea to be the people we are—*virtue ethicists*. In terms of our list of ethically relevant features, we could classify these theories by saying that consequentialists try to explain the items on the right and left columns on our list of symptoms in terms of the items found in the middle column. Deontologists try to explain the items on the middle and right columns in terms of the items found on the left. And virtue ethicists try to explain the items on the left and middle columns in terms of the items found on the right. The question, of course, is which of these diagnoses is correct.

§3. *Consequentialism*

The normative importance of the connection between acts and outcomes is obvious. Let us thus begin with consequentialism. On this approach, we start by noting that, through our acts, we shape the world we occupy. When you are thinking about whether to do something—say, skip class—the first thought you probably have concerns what will happen if you skip and what will happen if you don't. Those that take the consequentialist approach never move past this thought. They believe that, in figuring out what to do, all that ethically matters concerns what will be brought about. This is not to suggest that consequentialists are narrowly concerned with what causally follows from the act. Rather, the outcome that an act brings about includes everything in the complete world history were the act to be performed. Concern for outcomes, in this broad sense, exhausts the ethical domain, according to consequentialists.

But how precisely do we get from outcomes to deontic verdicts? The answer involves three steps. First, we need to look at the alternatives you have available to you at the time of action. Second, we

evaluate the world that would come about were you to take a given action. We associate an outcome with each action and then ask: How does this outcome compare to the other possible outcomes? Is it better or worse? Evaluating the outcomes of each action gives us a ranking. Third, we use this ranking to assign deontic verdicts to each action. And the natural thought to have concerning ranking is: Why do less than the best? The action with the best outcome is the one you ought to take. Yet, given that there could be ties for best, it is perhaps better to say that you are permitted to do (of the available actions) only what will bring about the best outcome. Doing less than the best is impermissible. Consequentialism, in a word, treats deontic verdicts as a function of the outcomes—in particular the goodness of the outcomes—our actions produce.

We can make these three steps concrete by looking at an example. Suppose you face a life-threatening illness. Your doctor gives you three options: (A) You can have a surgery with anesthetics, which will painlessly cure you. (B) You can have surgery without anesthetics, which will agonizingly cure you. Or (C) you can do nothing, in which case you will die an agonizing death. What ought you to do? If you are thinking like a consequentialist, you are thinking, “It ought to be that I am painlessly cured.” And why is that? Because (A) is better than (B), and (B) is better than (C).

<u>Option</u>		<u>Outcome</u>	<u>Deontic Verdict</u>
A	—————→	Best (cured, no agony)	<i>Required</i>
B	—————→	2nd Best (cured, agony)	<i>Impermissible</i>
C	—————→	Worst (not cured, agony)	<i>Impermissible</i>

This example illustrates the structure of a consequentialist theory. We ought to make the world as good as we can. Facts about the goodness of outcomes are the sole reason-providing facts. They are the only facts that matter for determining the deontic status of an action. But this example also illustrates what is compelling about the view. Without a moment’s reflection, it should be obvious that (A) is the option you ought to choose. Why endure unnecessary badness? The consequentialist globalizes this thought for the rest of the normative domain: It is only permissible to bring about the best outcome.

So far so good. However, in order to get from this consequentialist framework to a specific consequentialist theory, we need to nail down, with precision, what it takes for one outcome to be better than another. How should we evaluate the possible actions available to us?

In the above example, I took it for granted that your agony and death were bad. And these seem like pretty obvious candidates for bad outcomes. How well your life is going (or if it is going at all) appears to be one of the things that makes an outcome better or worse. Let's thus begin with the natural suggestion that facts about well-being are what determines the goodness of an outcome. (A discussion of the major theories of well-being can be found in the *Appendix*.)

3.1 *Egoism*

Suppose we focus exclusively on *your* well-being, holding that it alone is good. That is to say, one outcome is better than another if and only if your life is going better in that outcome. Nobody else matters. Just you. Then, combined with a consequentialist approach, this view would hold that you ought to make the world as good as you can, and the only way to do this is to make your life as good-for-you as possible. This view is known as

Universal Egoism. You are required to perform an action if and only if (and because), of the acts now available to you, this act will uniquely bring about the most good-for-you. That is, you ought to perform the action whose outcome is uniquely best-for-you.

The theory thus tells each of us to do whatever will best promote our own good. This principle serves as the criterion for assessing the deontic statuses of actions. An action is *required* if and only if so acting would uniquely maximize your overall well-being. An action is *impermissible* if and only if so acting would not maximize your overall well-being. And an action is *optional* if and only if so acting would maximize your overall well-being, but there is some other available action that ties in the amount of overall well-being it would bring you.

After hearing these deontic verdicts, universal egoism might not look very promising. Is the impermissibility of an action really located in the narcissistic fact that it wouldn't be best for *you*? Can't the good of others bear noninstrumentally on what you are required to do? The thought driving these questions seems to be this: Universal egoism makes ethics overly and implausibly focused on a specific person—you. Perhaps, however, egoism can be made more plausible by thinking about what nonegoistic theories demand. An altruistic theory maintains that ethics demands, at least sometimes, that you sacrifice your well-being for the well-being of others. Egoism denies this. Thus, if we can show that you are not required to be altruistic, then we are well on our way to universal egoism.

Let's begin by assuming that an altruistic act is required. What you have most reason to do is sacrifice your well-being for the well-being of others. Suppose, for example, that you are outside of a burning building with a group of strangers stuck inside. It is too late for the fire department to help. The only hope for these strangers is you. However, to save these strangers you will be severely burned. These burns will force you to live out the remainder of your existence in agony. Any benefit that you get from saving these strangers will, we can assume, be swamped by this future agony. Now, if the number of these strangers was large enough, most ethical theories will demand that you make this sacrifice. They demand altruism. But I suspect, confronted with this situation, you would pause to ask: Why ought *I* to do what altruism claims I ought to do?

If, as we are assuming, this altruistic act is required, this question would answer itself. It would be like asking: Do I really have most reason to do what I have most reason to do? Your original question, however, does not seem to be trivial in this way. You seem to be really asking for something. What's going on? One interpretation turns to the authority of the deontic verdict altruism delivers. When you try to assume the truth of altruism and then confront a situation where it demands something at odds with what is good-for-you, the deontic verdict seems to lose its authority—it fails to bind you in the right sort of way. To use a political analogy, altruism's demands are like the laws of a country of which you are not a citizen. When a country tries to command a non-citizen it makes sense to say: Why ought *I* to do what this state claims I ought to do? This state has exceeded its jurisdiction. And so too, on this line of argument, have altruistic ethical theories. You cannot have reasons to perform actions that do not ultimately promote your well-being. Such purported reasons do not speak to you; they lack authority.

If this line of argument is correct, then for a reason to have force it must serve your well-being. And by plugging this conclusion into the consequentialist framework—that you always have most reason to bring about the best outcome—we get universal egoism. Of course, the soundness of this argument turns on the claim that only your well-being can grip you. The fact that well-being extends across your whole life should thus give the egoist pause. For this means that egoist must cash out the metaphor of “being gripped” or “spoken to” in such a way that the *present* well-being of others is excluded from consideration but your *future* well-being is included. Yet why think that your future well-being would be any more gripping? That is, why think that reasons are agent-relative (it's all about you) but time-neutral (your present well-being

matters as much as your future well-being)? This asymmetry—placing special importance on *who*, while being indifferent to *when*—is hard to maintain. After all, the prevalence of procrastination, for instance, is best explained by the fact that the good-for-your-future-self does not always resonate with you now. But even if we assume the egoist can justify this asymmetry, problems remain.

Chief among these problems is that, if followed collectively, universal egoism is self-defeating. If we all act as universal egoism prescribes, each of us will bring about an outcome that is worse-for-each than had we not obeyed the theory's prescriptions. To see this, suppose that you and I each have a terrible migraine. I possess a capsule of aspirin, which, given my constitution, will make my migraine slightly less painful. But, given your constitution, the aspirin will make your migraine significantly less painful. You possess a capsule of codeine, which, given your constitution, will make your migraine slightly less painful. But, given my constitution, the codeine will make my migraine significantly less painful. A combination of both aspirin and codeine would completely eliminate the migraine for me and, likewise, for you. What does egoism say about this case? It says that no matter what I do, you are better off by giving yourself the codeine.

		You	
		give me codeine	give yourself codeine
I	give you aspirin	<ul style="list-style-type: none"> • 2nd best-for-me (pain significantly reduced) • 2nd best-for-you (pain significantly reduced) 	<ul style="list-style-type: none"> • Worse-for-me (full pain) • Best-for-you (eliminates pain)
	give myself aspirin	<ul style="list-style-type: none"> • Best-for-me (eliminates pain) • Worse-for-you (full pain) 	<ul style="list-style-type: none"> • 3rd best-for-me (pain slightly reduced) • 3rd best-for-you (pain slightly reduced)

Here's why: If I give you the aspirin, then you can completely eliminate your migraine by taking the codeine. But if I don't give you the aspirin, then the best you can do is make your migraine slightly less painful by taking the codeine. Either way, then, egoism prescribes that you take the codeine. Notice, however, that no matter what you do, I am better off giving myself the aspirin. So, if we both follow what egoism prescribes, we end up getting the next to worse outcome for each of us. But to do what seems like the obvious thing to do in this situation—i.e., swap painkillers—we would both need to disobey universal egoism.

3.2 *Contractarianism*

When confronted with this painkillers case, your initial reaction is probably that this coordination problem has an easy solution. What egoism should demand is that we first make an agreement to do what is mutually beneficial, and then turn to addressing the distribution of the pills. This is precisely what one approach, known as *contractarianism*, holds. This theory aims to retain the driving idea behind egoism—that all of your reasons are reasons to promote your well-being—but avoid the problem posed by the painkillers case by filtering egoism through a contractual agreement. Given the prevalence of coordination problems, egoism demands that we conform to the details of this contract. Very roughly, then, we can characterize contractarianism as endorsing the following:

You are required to perform an action if and only if refraining from so acting is disallowed by principles that everyone would egoistically agree to (under suitable conditions).

This formulation is rough because we would need to know the conditions needed for the contract to qualify as suitable. For instance, we might want to add that the bargainers are idealized in certain ways—e.g., they are informed, rational, and so forth. Without such idealizations, we would in effect guarantee that no consensus could be reached. Misinformed and irrational people are hard to bargain with.

Still, even if we landed on suitable idealizing conditions, we would need to identify just what principles would be agreed to. To identify these principles, it will help to recall what each party wants out of the contract. We all want the benefits afforded by cooperation. We all want this contract to succeed. Yet each of us wants to secure the greatest benefit while taking on the fewest costs. What principles we end up with will largely depend on how we think this tension should be adjudicated. However, let's not pursue this issue further.

Whatever the chosen principles turn out to be, a lingering worry concerns compliance.

Remember that egoism serves as the foundation for contractarianism. So what happens when adhering to the terms of the contract is not good-for-you? Return to our painkillers case. Suppose, for example, that we both agree to the principle: If it's mutually advantageous to swap pills, we ought to swap. Next consider a twist on the original case. Suppose that you have a fake codeine pill, which initially works just like codeine but kills one week after ingestion. Moreover, this fake pill kills in a way such that the cause of death looks natural. If you give me this pill, I will give you aspirin. And, because you retained your codeine pill, you can completely eliminate your migraine. What ought you to do? Assuming this is a one-off case—you don't want to get a well-being-diminishing reputation—egoism counsels you to secretly defect from our agreement. You ought to give me the fake pill. For recall, the driving idea behind egoism is that you cannot have reasons to perform actions that do not ultimately promote your well-being. And giving me codeine, when you have a substitute that can secure my aspirin without this cost, does not ultimately promote your well-being. The problem for contractarianism should now be obvious. If you can, by force or fraud, secure the benefits of the contract without incurring the cost, then egoism prescribes that you break the contract. You ought to free ride whenever it is best-for-you. Contractarianism's egoistic foundation demands this defection. Why? Because egoism claims that only well-being enhancing reasons can grip you. And, hence, in this case, since adhering to the chosen principles does not enhance your well-being, you have no reason to follow them. The contract loses its grip; it fails to bind.

This brings us back to our original worry with egoism: It heightens your own importance to the point of absurdity. If you think that you ought not kill me in the revised painkillers case or fail to live up to your side of the bargain, then you both reject egoism and think that contractarianism faces a serious problem.

3.3 *Utilitarianism*

The lesson from our discussion of egoism is that ethics should be more impersonal. Egoism tells you that your well-being, and only your well-being, matters. But what is the ethically relevant difference between the good-for-you and the good-for-me? Is there some magic in the use of the possessive—*my* well-being? It seems not. As Sidgwick famously put it, “[T]he good of any one individual is of no more importance, from the

point of view . . . of the Universe, than the good of any other; unless, that is, there are special grounds for believing that more good is likely to be realized in the one case than in the other.” Following Sidgwick, perhaps we should think that our well-being, taken together, is what matters. Your well-being is a goodness-making feature of an outcome, and so is mine. By extending this line of reasoning to everyone, we arrive at our next consequentialist theory—*utilitarianism*. This theory holds that one outcome is better than another if and only if the total amount of well-being is greater. According to this consequentialist approach, you ought to make the world as good as you can, and the only way to do this is to impartially maximize well-being. The following principle thus serves as the criterion for assessing the deontic statuses of actions, according to

Utilitarianism. You are required to perform an action if and only if (and because) of the acts now available to you, this act will uniquely bring about the greatest sum-total of well-being. That is, you ought to perform the . . . action whose outcome maximizes total well-being.

To see what this principle amounts to, we will need to understand precisely what promoting the *sum-total* of well-being means.

It means that the theory is aggregative. In determining the sum-total of well-being attached to the outcome of a given action, we need answers to the following questions:

The Increase Question. For each entity whose well-being would be affected by the performance of the action, how much would the act increase their well-being?

The Decrease Question. For each entity whose well-being would be affected by the performance of the action, how much would the act decrease their well-being?

Once we have an answer to the Increase Question, we take all of the units of well-being that will accrue to each entity and add them up. Then we do the same for the Decrease Question. Finally, we subtract the increases from decreases, giving us the net sum-total.

What utilitarianism demands is that we bring about the outcome with the highest sum-total of well-being. More precisely, an action is *required* if and only if so acting would uniquely bring about the greatest sum-total of well-being. An action is *impermissible* if and only if so acting would not bring about the greatest sum-total of well-being. And an action is *optional* if and only if so acting would bring about the greatest sum-total of well-being, but there is some other act that brings about an equal sum-total of well-being.

Unlike universal egoism, the deontic verdicts supplied by utilitarianism seem initially quite plausible. One area where this is particularly evident concerns our treatment of animals. Just as it appears that there is nothing special about *my* well-being that makes it more important than your well-being, there appears to be nothing special about *human* well-being that makes it more important than animal well-being. (This is why I used “entity” rather than “person” in the Increase and Decrease Questions above.) By utilitarianism’s lights, it is just as bad for me to needlessly break your arm as it is for me to needlessly break the leg of a dog, assuming the decrease in well-being is the same. Intuitively, needlessly causing pain to a dog seems like a bad thing, something we have reason to avoid. Utilitarianism nicely explains this reason. And it locates the explanation in exactly the right place—the fact that the dog would suffer. Contrast this with egoism or contractarianism. Egoism says that if breaking the dog’s leg would best promote my well-being, then I am required to do it. Contractarianism, at best, locates the reason not to break the dog’s leg in an agreement with others. But this agreement has nothing directly to do with the dog. Thus, even if contractarianism can tell me not to harm the dog, it comes from the wrong place—a contract, not the dog’s leg.

Despite its initial appeal, utilitarianism faces serious problems both in terms of its sum-totaling and its exclusive focus on well-being.

In focusing exclusively on the sum-total of well-being, the theory is indifferent to how well-being is distributed. A counterintuitive implication of this exclusive focus is immediately obvious once we remember that there are two ways of increasing the total sum of well-being: We can either increase the well-being of existing entities, or we can bring into existence new well-being—having entities. This latter option is where the trouble lies. Suppose, for instance, that you are in charge of the world’s population. Given the earth’s resources, you have two possible policy options:

Making People Well-Off. Keep population levels static. Earth’s available resources will then permit each member of the limited set of existing people to have exceedingly good lives. In short, have a low population with each person given a very high level of well-being.

Making Well-Off People. Greatly increase the population. Earth’s available resources will then permit each member of the enormous set of existing people to have lives barely worth living. In short, have a high population with each person given a very low level of well-being.

Now for the crucial stipulation: The sum total of well-being produced by each policy is the same. Confronted with these options, what does

utilitarianism tell you to do? It tells you that since, in terms of well-being, each option is equally good, you are permitted to choose either. Yet could Making People Well-Off really be ethically on a par with Making Well-Off People? If you think that treating these two options equally is a mistake, then utilitarianism's sum-totaling must be abandoned. Of course, there are other ways of aggregating well-being. We could aim for the highest average, for example. But let's leave aside these possible revisions, and turn directly to a distribution problem that calls into question utilitarianism's exclusive focus on well-being.

Most of us hold that people getting their due is good. Intuitively, outcomes are better when people get what they deserve. Suppose, for example, your lifelong enemy wants to harm you. With malice in his heart, he tries to push you off of a cliff. However, worked up by the prospect of your destruction, his push carries him over the precipice with you. You both survive the fall, but are gravely injured. As the rescuers arrive, it is clear that, given the severity of the injuries, there is only enough time to save one of you. And now for the crucial stipulation: Whether you or your enemy is saved, the outcome, in terms of the sum-total of well-being, will be the same. What does utilitarianism tell these rescuers to do? Like the population case above, it tells them that since, in terms of well-being, the options are equally good, they are permitted to choose either option. Between saving you and saving your enemy, they could flip a coin. But that seems off. If you think, due to his culpability, the outcome would be worse if they saved your enemy—he caused this mess, after all—then you think that well-being is not the only factor that determines the level of goodness that an outcome can have. Desert also matters. The outcome is better when your enemy receives the level of well-being he deserves.¹

3.4 *Act-Consequentialism*

Utilitarianism, in telling us that the maximization of well-being is all and only what matters ethically, thus faces serious problems. Yet these problems can be alleviated with a simple revision. We can simply drop the idea that well-being is all that matters. As we have seen, people getting their due might also influence the goodness of an outcome. And, if desert can influence the goodness of outcomes, then other goods might as well. Equality and beauty, for instance, are also plausible candidates. Each, it might be thought, is good in itself. Notice we have here moved away from the idea that for something to be good is for it to be good-for-us. Rather desert, equality, and beauty

are good-*simpliciter*. These goods are impartially good, regardless of whether they are good-for-anyone. On this broad consequentialist approach, then, you ought to make the world as good as you can, and the only way to do this is to maximize the total amount of good-*simpliciter*. This broad consequentialist theory is known simply as

Act-Consequentialism. You are required to perform an action if and only if (and because), of the acts now available to you, this act will uniquely bring about the most good-*simpliciter*. That is, you ought to perform the action whose outcome uniquely maximizes the good, period.

Act-consequentialism tells each agent to do whatever will best promote the good overall, impartially assessed.

To be clear, utilitarianism is a species of act-consequentialism, restricting the only thing that is good-*simpliciter* to well-being. Act-consequentialism, as the broader genus, is simply more liberal in the potential goods that may increase the goodness of an outcome. Of course, if we took this broader approach, we would need an account of these other goods, and how they interact, in order to determine the ranking of our available actions. But, once we had this account, we would be well positioned to avoid the above problems associated with the above narrower theories, while retaining consequentialism's compelling structure. Rather than trying to articulate what things are good-*simpliciter*, however, let's conclude with two serious problems for act-consequentialism generally.

Act-consequentialism says that you are permitted to act in a certain way if and only if this act would bring about the best outcome. By breaking apart the necessary and sufficient conditions of this principle,² we can bring out two implications of act-consequentialism that are particularly troubling: namely, its lack of (i) a wide range of permissible actions open to you, some of which will not maximize the good (*options*); and (ii) a set of impermissible actions that you simply ought not to perform even if they maximize the good (*constraints*).

Take first the

Necessary Condition. Only if an action brings about the best outcome, is it permissible to perform the action.

This condition leaves little room for people to live their own lives. For you to act permissibly, you must bring about the most good. Put differently, it is always impermissible to fail to maximize the good. For example, right now you are reading this introduction. Maybe it will do some good. It will help you better understand ethics. But that, act-consequentialism says, is not enough. Your reading this introduction

is not permissible unless, of the alternatives, this is the action that produces the *most* good. That is a daunting demand. Recall the deontic verdict concerning what it would take for an action to be optional—it would have to produce the most good and also be tied with some other available action that produced an equal amount of good. Accordingly, act-consequentialism doesn't hand out a lot of optional deontic verdicts. This dearth of options appears to be a serious problem. Shouldn't you have, at least once in your life, the option to favor or sacrifice yourself, even if it comes at the expense of the overall good? If you think we have such options, then you think that act-consequentialism requires too much.

Unfortunately for act-consequentialism, the problems do not end here. Turning now to the

Sufficient Condition. If an action brings about the best outcome, it is permissible to perform the action.

This condition seems to permit too much. Act-consequentialism tells you that committing *any* act is permissible if it produces the most good. Put differently, it is never impermissible to maximize the good. Indeed, if an action uniquely produced the most good, it would be required. To see why this appears particularly troubling, pick an act that is as abhorrent as you can imagine—e.g., murdering, torturing, stealing, or lying. Next, rig two outcomes such that one outcome, the one where *you* are performing the abhorrent action, is infinitesimally better overall than the outcome where you do not perform the action. Act-consequentialism always comes back with the same verdict. It's not just permissible for you to perform the abhorrent action, it's required. There are no constraints placed on what consequentialism could demand from you. And that, I suspect, is a claim from which you will likely recoil.

We have, however, only scratched the surface in our investigation of consequentialism. Consequentialists have creative ways of responding to the problems raised in this section. But we've covered the basic tenets of the approach, so let's now take stock. The theories that take evaluative claims as primary sit on two extremes. Egoism, with its complete partiality toward self, appears to demand too little of us. Act-consequentialism, with its complete impartiality toward all, seems to demand too much. The seemingly troubling implications of universal egoism and act-consequentialism thus provide a helpful way to orient our discussion. The remaining ethical theories can be seen as attempts to forge a middle path. They try to resist the pull of

act-consequentialism's impartiality without being driven all the way to egoism. How might this be accomplished?

§4. *Deontology*

Let's pause to remember what seems to be the force driving egoism and act-consequentialism to their extremes. Surprisingly, the source of these problems seems to be what was so compelling about consequentialist theories in the first place: that it is always permissible, indeed required, to maximize the good. Once pointed in a certain direction, with nothing holding them back, the deontic verdicts of consequentialist theories have no bounds. What we are looking for, then, is a way of directly securing certain verdicts. We want certain parts of your life to go untouched by ethics, leaving you a wealth of options. And we want certain acts to be off-limits, keeping you from getting your hands dirty. With these acts—e.g., murdering, torturing, stealing, or lying—looking to the consequences seems out of place. We don't have to see what these acts will bring about in order for us to know that you ought not to do them. The acts themselves, not their consequences, seem ethically problematic. Deontological theories try to capture this idea. They directly claim that certain actions are optional and certain acts are required. By taking a deontological approach, we could thus hold that ethics includes both options and constraints.³

But what would the structure of a deontological theory look like? The twin goals of securing both options and constraints supply an answer. First, a deontological theory will attempt to enumerate and explain why certain acts are impermissible for you to perform. This, in turn, will tell you what you are required to do. You are required to refrain from such actions. Second, the theory will claim that all actions that fall outside of the set of required acts are optional. If this is right, then ethics plays a limited role in governing over the actions of which your life consists. So long as you have refrained from the impermissible actions, the rest is up to you. By contrast, as we saw, consequentialism takes the jurisdiction of ethics to be all encompassing. Consequentialists have something to say about every act you could perform.

Constraints, as should be clear from this sketch, play a dominant role in deontological theories. Thus, to get a grip on the deontological approach, we need to have a better grip on the nature of constraints. As their name suggests, they are a kind of normative barrier.

They tell you that *you* are required to refrain from performing certain *types* of acts—e.g., murder, torture, stealing, and lying—even if so refraining comes at the cost of the overall good. Constraints thus have two features worth stressing. They concern certain types of acts, and they speak specifically to you. Let's take each of these features in turn.

Think back on our discussion of consequentialism. There we always spoke of specific actions. The explanation for this should be obvious. If consequentialists treat deontic verdicts as a function of evaluative claims, then we need to assess each available action for the amount of good it will bring about to identify which ought to be performed. But part of the worry we had about act-consequentialism stems from the thought that some acts, themselves, are simply abhorrent. And here we've arrived at an important distinction, the distinction between act-types and act-tokens. To help see this distinction, consider the question: How many letters are in the word 'fool'? This question is instructively imprecise. We might plausibly answer: four—counting each specific letter. Or, we might plausibly answer: three—counting each type of letter, and so counting 'o' only once. If we answer four, we took the question to be a question concerning letter-tokens. If we answer three, we took the question to be about letter-types. The same distinction applies to acts. If we take consequences as primary, we are only concerned with *act-tokens*. But, once we move to a deontological theory, we think in terms of *act-types*. Constraints are formulated as barriers, blocking the performance of certain act-types—for example, murder is impermissible. If you then perform the act-token—you commit murder—you've violated the constraint, and thereby done something impermissible.

Even once we've noticed this important point concerning the nature of constraints—that they are typically formulated in terms of act-types—a consequentialist thought might persist. If constraints give us the goal of not performing certain act-types, why not think that we should promote the outcome with the fewest performances of such acts? Consider a schematic case. Suppose that you are in the following position: Unless you commit murder, I will commit two murders. The consequentialist thought here is that if violations of constraints are to be avoided, then we should have the goal of minimizing such violations. If constraints are so important, then shouldn't you commit murder to prevent two additional murders? The answer, which is what drove us to deontology to begin with, is that *you* want to avoid having to be the one performing the abhorrent act. To avoid

consequentialism's pull, we thus need to make a slight retreat toward egoism. We need to formulate constraints in a way that makes essential reference to *you*—the person to whom the constraint applies.

This leads us to the next feature of constraints. The reasons attached to the constraint must be *agent-relative*, not *agent-neutral*. To help see this distinction, suppose you are on a road trip with your family. Your parents tell you and your two siblings to make sure there is as little fighting as possible. Not long after, your two siblings get in a fight. Suppose you know that unless you physically intervene, the fight will continue for the rest of the trip. Your parents' command thus provides you with a reason to fight, because your intervention would bring about less fighting. The reason provided by your parents' constraint against fighting is thus agent-neutral. It gives all the children the same aim: minimize fighting. Now imagine the same scenario, but this time your parents tell you and your siblings, each individually: no fighting! Again not long after, your two siblings get in a fight, which unless you physically intervene will last the rest of the trip. Yet this time you cannot fight to reduce the fighting, for if you did, you would disobey your parents' command. In this second case, the reason provided by your parents' constraint against fighting is agent-relative; it makes essential reference to you, telling *you* not to fight. To capture the idea that it is ethically important for you to keep *your* hands clean (even if it comes at the expense of the overall good), deontology mimics your parents in this latter case. You must be given a weighty (perhaps decisive) reason to avoid violation of the constraint, even if your violation would prevent additional violations of that very constraint. If the constraint doesn't make this essential reference to you, if the reason attached to the constraint is, in other words, agent-neutral, then others' constraint violations would, in determining what *you* ought to do, be equally reason-providing as your own. The constraint would, if agent-neutral, give everyone the same aim: minimize constraint violations. Hence, in the case where unless you commit murder I will commit two murders, you would have stronger reason to murder. A deontological theory thus needs to be articulated in agent-relative terms, meaning that such a theory could potentially give different persons different aims.

Once we recognize these two features of constraints—that they focus on act-types and are backed by agent-relative reasons—we are now in a position to feel the full force of the explanatory burden that the deontologist faces. Not only do constraints wall you off from the good, they wall you off from other deontic verdicts. What you are

required to do is held fixed in the face of what others do (and perhaps what your future self might do). The viability of the deontological approach thus hinges on locating a plausible source of agent-relative reasons not to perform certain act-types. In what follows, we will look at a number of different sources that may supply the grounds for such reasons. We will look at two outside sources of constraints: God and culture. And we will look at one inside source: our own rationality.⁴

4.1 *Divine Command Theory*

Most religious views take God to supply a number of agent-relative constraints. The ethical theory that draws on this theological tradition is known as the

Divine Command Theory. You are required to perform an action if and only if (and because) your performance of such acts is commanded by God.

On its face, the principle divine command theory uses as the criterion for assessing deontic verdicts has some plausibility, and it seems to provide a powerful explanation of the source of our agent-relative constraints. Yet a serious difficulty lurks.

The divine command theory faces an old problem, a dilemma that directly calls into question its ability to explain why actions have the deontic verdicts that they do. Here's the dilemma: Either what is ethically required is required because God commands it, or it isn't. If performing an action is required because God commands it, then the demands of ethics are arbitrary. God could have commanded that we do the most abhorrent acts; and in so commanding, these acts would be required. But if we go for the other horn of the dilemma, maintaining that God commands certain acts because they are required, then God cannot be the source of our agent-relative constraints. God is superfluous to the ethical story. At best, God identifies the constraints we already have. God's commands, therefore, are either ethically arbitrary or unnecessary.

Nevertheless perhaps the divine command theorist is willing to accept the arbitrariness horn of this dilemma. If God commands murder, then murder is what is required. Embracing this counter-intuitive implication, however, does little to help. For the underlying problem, hinted at above, comes from deriving the source of *all* ethical requirements from the commands of God. Why is this a problem? Because it seems perfectly sensible to ask: Why am I required to do

what God commands? This question is not trivial. It seems to demand an answer. But what answer can the divine command theorist give? To be consistent, the answer must be that God has commanded that you ought to do what God commands.

This response, however, will not do. Either we already ought to do what God commands, in which case we are back to the superfluous horn of the dilemma. Or it is not already the case that we ought to do what God commands, in which case the command lacks ethical force, thereby leaving the original question unanswered. In other words, we need a reason, not itself provided by God's commands, to obey his commands. Yet if such a reason exists, then God is not the source of all ethical requirements. And, if no such reason exists, then God's commands lack normative authority, in which case God is not the source of all ethical requirements. Either way, then, God is not the ultimate source of ethical requirements.

4.2 *Cultural Relativism*

There's nonetheless something attractive about looking for the source of the constraints in God. That there is something outside of us handing down the constraints seems to imbue them with the special authority they seem to have. So even if God does not undergird agent-relative constraints, perhaps something outside of us still can. For example, we often hear people say that there are certain things we just don't do around here. The thought seems to be this: Certain constraints apply to people simply in virtue of being a member of a social group with certain norms and customs. Being part of a particular culture, then, provides another possible source of agent-relative reasons to refrain from performing certain act-types. This theory is known as

Cultural Relativism. You are required to perform an action if and only if (and because) your performance of such acts is called for by the norms of your culture.

This theory holds that deontic verdicts are indexed to cultural norms. The idea is that ethical sentences function like sentences that use 'here' or 'now.' For example, when we say, "It's raining here," the truth of the utterance is relativized to a particular location—the location of the speaker. Similarly, according to cultural relativism, the truth of an ethical utterance is relativized to the norms of the culture from which the verdict is issued. Cultural relativism thus asserts a bold thesis. It does not merely make the uncontroversial claim that

deontic verdicts are sensitive to the circumstances; rather it denies the existence of ethical truths that cut across all cultures. Cultural relativism, for instance, tells you that torture is *impermissible* because the norms of our culture include a constraint against engaging in torture. But what would make torture permissible?

Interestingly, cultural relativism gives us two answers. Torture would have been *permissible* had our culture not had norms that include a constraint against engaging in torture. Or, even if the norms of our culture contain a constraint against torture, your engaging in torture would be *permissible* from the perspective of other cultures whose norms lack such a constraint.

Having these two ways of arriving at permissibility is a puzzling feature. Could the norms of a culture really determine the correct deontic verdict concerning torture? If that seems mistaken, then we're led to accept an important desideratum on normative theories: Their deontic verdicts must be *universalizable*, applying to all persons in relevantly similar circumstances. More precisely, if a theory's verdicts are universalizable then the following is true of the theory: For all persons, each person—e.g., you—were you in so-and-so circumstances, you would have a reason to perform such-and-such action. And, accordingly, were I in these circumstances, I too would have this reason. The problem with cultural relativism is not just that it delivers implausible deontic verdicts. The deeper problem is that it violates universalizability. It tells us that the identical circumstances—where all the same facts obtain—can give rise to different deontic verdicts. That's puzzling, to say the least.

Having noted the problems with Divine Command Theory and Cultural Relativism, we now have an even better handle on what a deontological theory is after: locating an authoritative source of agent-relative reasons not to perform certain act-types while retaining universality. So, given this aim, perhaps the source of constraints can be derived from something about all of us. After all, we are the entities to whom the constraints apply. We are *agents*—beings capable of recognizing and responding to reasons.

4.3 *Kantianism*

That we are agents is a thought worth exploring. Right now you are reading this essay. If I asked you why you are doing this, you might say, "In order to learn about ethics." Here you supply me with a reason, justifying your action. Contrast yourself in this regard with a nonrational

entity, say, a rose. If I asked you why does this rose bloom, you might say, "In order to attract pollinating insects for reproduction." Though this answer sounds similar to your previous one, it is not. With the rose, your cited reason is not justificatory. The rose, unlike you, does not have a rational nature. You, not the rose, are sensitive to reasons. External forces, like the sun, do not drive you. You are free in a way that the rose is not. And it is only in virtue of this freedom that ethical questions become salient. Insofar as you are responsive to reasons, you can act. The rose cannot act. It's blooming is akin to your having an involuntary spasm. The rose does not ask: Ought I to bloom? Only persons, beings with rational natures, confront normative questions.

That you have a rational nature, of course, does not mean that you always act rationally. Depressingly, most of us act irrationally all too often. But now we've arrived at an intriguing idea. Perhaps agent-relative constraints fall directly out of our own rational nature, such that violating the constraint can only be done on pain of our own rationality. That would be a neat result. Unlike the divine command theory or cultural relativism, this would locate the source of ethical requirements in you—in the mere fact that you are rational. Surely your own rationality binds you. So how could we establish this tight connection between rationality and agent-relative constraints?

To answer this question, return to universalizability—the thought that our deontic verdicts need to apply to all persons in relevantly similar circumstances. Previously, we concluded that universalizability seemed like a plausible desideratum on ethical theories. Given our discussion of rationality, we can now defend this desideratum. The connection between rationality and universalizability is easiest to see when it comes to forming beliefs. Insofar as we are rational, we cite reasons. You do not simply believe, for example, that the gas in your car is low. Rather, if you are to rationally form a belief concerning how much fuel your car has, you look for reasons. For instance, you see that your car's reliable gas light is on, and this justifies your belief that the gas is low. Notice an interesting feature concerning the evidence that your gas light provides: It is not person specific. If it's rational for you to believe that the gas is low because of the gas light, then, based on this same reason, were I in your situation, it would also be rational for me to believe the gas is low. You don't have a monopoly on gas-light reasons. Universalizability thus appears to be constitutive of rational belief formation. If some fact serves as a genuine reason for you to believe, it must serve as a genuine reason for me, assuming we are in the same circumstances. And this idea is not

restricted to the formation of beliefs. Presumably, universalizability is also constitutive of rational action. To act rationally is to act on reasons I can coherently see others, in relevantly similar circumstances, acting on as well. In short, since rationality demands consistency, we are driven to universalizability.

This link between universalizability and rationality leads us to a *universalizability test*. If a purported reason for acting can't apply to all rational beings in relevantly similar circumstances, then it is not a reason. With this test in place, we are in a position to see how constraints arise from our own rationality. We apply the universalizability test broadly, identifying acts whose features are such that any purported reason in their favor fails the test. If permission-granting reasons for certain acts systematically fail, we can conclude that you have an agent-relative constraint against performing them.

To see how this test works, imagine that you are in dire need of money. Imagine further that you are over at a wealthy friend's house, where cash is carelessly left around. Suppose you are contemplating taking some of this money. It would mean little to your friend, who will not even realize it is gone. But taking some would greatly increase your well-being. Here you might be thinking

Steal. The fact that taking what's not mine will increase my well-being gives me decisive reason to take it.

Now, in order to act rationally, everyone must be able to act on this reason. Our universalizability test thus begins by trying to imagine a world where everyone acts on the reason you claim to have in *Steal*. That is, we imagine a

Full Compliance World. In this world, everyone complies with *Steal*. Whenever taking what's not one's own increases one's well-being, one takes it.

We next ask: Is this Full Compliance World possible? Here we are asking could all rational agents adhere to *Steal* as if a law of nature (like gravity). The answer, it seems, is No. If everyone followed *Steal*, it would destroy the very institution (property) that serves as a means to achieving the end in question (your owning the well-being-enhancing goods). In such a world where everyone complied with *Steal*, the very institution of ownership would cease. It would be pointless to steal, since no one would own anything. Accordingly, *Steal* fails the universalizability test. If not everyone could rationally act on the reason captured by *Steal*, no one could rationally act on it. Failing universalizability is the hallmark of irrationality. And since you don't have a monopoly on

well-being-based reasons to take what's not yours, you are irrational if you *Steal*. But perhaps we can say more. Perhaps no reasons in favor of stealing can pass the universalizability test. If so, there would be an agent-relative constraint against stealing—stealing, the act-type, would be impermissible. That would be a substantial result; it would show the possibility of generating constraints through rationality alone.

The deontological theory we've been exploring brings us close to the theory espoused by Immanuel Kant. Kant's theory is complex, but one of his central ideas is that the reasons that serve as the determining grounds of our deontic verdicts need to be sharable by all rational beings. In his words, "Act only in accordance with that maxim [i.e., the statement describing what you are going to do, and the reason you are going to do it] through which you at the same time can will that it become a universal law." We can state this principle as follows:

You are required to refrain from acting in a certain way if and only if (and because) you couldn't rationally will as a universal law the maxim associated with this action.

This is an elegant way of generating constraints. Rationality is what makes normativity possible; we must have agents to have reasons. But rationality, fully appreciated, is also the source of our agent-relative constraints. What's more, there is an intuitive thought—how would you feel if everyone did that?—that Kantianism nicely captures. (Think of how this thought can be deployed, with great effect, to deal with our painkillers cases from the discussion of Egoism [§3.1] and Contractarianism [§3.2]. Purported reasons to defect on a mutually advantageous agreement for personal gain are not universalizable.)

Yet there are problems for Kantianism too. The universalizability test appears to tell us that we are required to refrain from certain acts when intuitively performing them is permissible. Suppose that, because you need cash on hand, tomorrow you will pull your remaining funds out of the bank. Could all rational agents adhere to this maxim as if a law of nature? It seems not. The universalizability test thus tells us that you acted impermissibly. That looks like the wrong verdict. This action is obviously permissible. And these examples are easy to multiply: Taking the last chair in an auditorium, donating the most to charity, watching live TV shows (the content of which is not others watching live TV), and so on.

And the problems for Kantianism do not end here. The universalizability test seems to tell us that we are required to perform certain acts when intuitively refraining from them is permissible. Suppose

that, out of romantic honor, were you to fight in a duel, you would aim to miss. Is a world where everyone adhered to this maxim possible? It seems not. For then, like the institution of property in *Steal*, the institution of dueling would break down. According to the universalizability test, then, it would be impermissible for you to aim to miss. Again, that seems like the wrong verdict. The universalizability test thus either needs to be refined or abandoned.

Fortunately Kantians have another, highly influential, derivation of agent-relative constraints from rational agency. Recall that rationality demands consistency. And, to state the obvious, much of our ethical lives consist of interactions with other people—i.e., other rational agents. When interacting with a person could you, without inconsistency, use your rational agency to undercut theirs?

To answer this question, return to what using your rational agency involves. Recall that rational agency is the capacity to recognize and respond to reasons. Think about how you use this capacity on a daily basis. Take any action. You see, for example, the beautiful rose, and think that having this beautiful rose would be nice. The rose then becomes your end, and you set off for some scissors. Though done a hundred times a day, this ability to set ends is a striking phenomenon. By setting ends, you seem to be able to generate reasons for yourself. But how could you have this reason-generating ability? One way—perhaps the only way—is if you yourself were an end. Yet does your being an end depend on anything else? Presumably not. You—as a rational agent—are an end in itself. Notice, however, that this same progression of thought applies to anyone with this capacity for rational agency. Accordingly, we arrive at the idea that all rational beings are ends in themselves simply by reflection on what it takes to act as a rational agent.

We now have an answer to the question whether you could, without inconsistency, use your rational agency to undercut the rational agency of another. When you act in this way, you are using a rational agent as a mere means to one of your ends. Yet, as we just saw, to act at all you are committed to treating rational agents as ends in themselves. If you undercut the rational agency of another, you are not treating a rational agent as an end in itself. So your action is inconsistent. Thus, to act in this way is to act irrationally, and hence unethically.

This line of reasoning led Kant to put forward another principle, which states: “Act in such a way that you always treat humanity [rational nature], whether in your own person or in the person of any

other, never as a mere means but always as an end.” We can state this principle as follows:

You are required to refrain from acting in a certain way if and only if (and because) this act uses another rational agent merely as a means and not as an end.

When you use people as a *mere means* you use them in ways to which they could not rationally consent, thereby blocking them from sharing in your end. That is to say, you use them in a way that they cannot share your reason; you undermine their rational agency.

Given the two abilities that come with rational agency—the ability to recognize reasons and the ability to respond to them—the two most obvious ways of using someone as a mere means are deception and coercion. In cases of deception, one’s ability to recognize the actual reasons in play is distorted by another. In cases of coercion, one’s ability to respond to the reasons is thwarted by another.

In order to see how this principle works, consider in detail a case of deception. Suppose again that you are in dire need of money. And again, you are over at your wealthy friend’s house. As it nears supper-time, you tell your friend that, if he pays, you will pick up takeout. He agrees. Your statement, however, is a lie. You plan to take the money, never to see your friend again. This deception uses your friend as a mere means. You do not allow him to share your goal; the possibility of rational consent is ruled out by the lie. You treat your friend like a tool to achieve your ends. You show your friend the same respect you might show to an ATM, not the respect due to a rational being. In short, your lie is inconsistent with seeing humanity as an end in itself. And seeing humanity as an end in itself is something, simply in virtue of acting at all, you implicitly accept. Your lie is therefore irrational. As before, if all reasons to lie were inconsistent with treating people as ends in themselves, there would be an agent-relative constraint against lying.

Still, problems remain. We can best see these problems if, as we did with act-consequentialism, we break apart the necessary and sufficient conditions of Kant’s principle. Let’s begin with the

Necessary Condition: Only if an action uses another rational agent merely as a means and not as an end are you required to refrain from performing the action.

This condition appears to condemn too little. Although much of our ethical lives consist of interactions with other rational beings, this does not exhaust the possibilities. As we saw earlier—in the case involving the dog’s leg—animals can have their well-being reduced,

and such reductions seem ethically relevant. But what could a Kantian say if we substituted a dog's leg for the rose in the example above. You see the dog's leg, and think that having this leg would be nice. The leg then becomes your end, and you set off for some scissors. (Let's just assume that dogs are not rational agents. If you think dogs are rational, then we can change the example to a lower animal that can still feel pain.) Clearly, you ought to refrain from cutting off the dog's leg. But this act does not use other rational agents merely as a means. It uses a nonrational dog. And hence, the action is one you are not required to refrain from. That is hard to believe.

We can next turn to the

Sufficient Condition: If an action uses another rational agent merely as a means and not as an end, you are required to refrain from performing the action.

This condition appears to condemn too much. Intuitively, it is permissible to use a person as a mere means, if the end to which they are being put is sufficiently important. Return to our schematic case where unless you murder I will commit two murders. Now let's make two revisions. First, the act in question for you is using someone as a means. You can achieve the performance of this act in a number of different ways—e.g., deception or coercion. The second revision is that the number of murders I will commit is high enough that the permissibility of you using someone as a mere means is ethically irresistible. For example, suppose that unless you use someone as a means I will commit one thousand murders. Even if we raise the number to a million, the sufficient condition of Kant's principle will always come back with the same verdict: You are required to refrain. Again, that is hard to believe.

We can now take stock. We have investigated three attempts—God, culture, and rational nature—to explain the existence of agent-relative constraints. Each faces problems. Of course, deontologists have forwarded ways of responding to the problems raised in this section. Instead of digging deeper, however, let's end by returning to options.

Our investigation into deontology was motivated by the quest to make room for options and constraints. We have been focusing on constraints because we assumed that, if constraints could be defended, the deontologist could claim that so long as you live within these barriers, the rest of your actions are optional. This claim, however, is suspect.

To see this, return to utilitarianism's plausible idea that, between two outcomes, if one outcome has a higher level of well-being than

another, you have a reason to favor that outcome. Now let's suppose that among your available actions you face two outcomes, both of which involve no constraint violations but one produces a higher level of well-being. Are both of these actions optional? If well-being is reason providing, then, assuming neither action violates a constraint, the set of reasons in favor of the outcome with the higher level of well-being does not match the reasons in favor of the outcome with the lower level of well-being. You have more reason to bring about the outcome with more well-being. And you are required to perform the act that the balance of reasons favors. Thus the deontologist cannot simply say that actions that lie within the barriers set by the theory's constraints are optional. So, even if agent-relative constraints can be secured, work for the deontologist remains, if they hope to also secure options.

At this point, we might be suspicious that supplying an explanation of agent-relative constraints and options is possible. Deontological theories leave too many normative mysteries. Act-consequentialism might, at this point, appear rather alluring. However, before we succumb to consequentialism's pull, we should explore a final approach.

§5. *Virtue Ethics*

Perhaps our focus on acts and actions—on conduct—pointed us in the wrong direction. Think about eulogies or obituaries. Most do not list the actions the deceased performed. Instead, they talk about the kind of person the deceased was. In other words, they talk about the deceased's character. And this talk is not limited to funeral homes. Often, in everyday discourse, character is primary to conduct. When I was a kid, a ubiquitous Gatorade commercial demanded that young basketball players "Be like Mike." Perhaps we should follow this linguistic evidence, and treat character as primary to conduct. On this approach, it is only by first figuring out who we ought to be that we can figure out how we are required to act. We treat our deontic verdicts as a function of our character traits. We first ask what attributes you ought to possess, and then use those traits to figure out what to do.

To understand this approach, we first need to understand what it is to have a certain character. Consider Gatorade's claim that young basketball players should be like Michael Jordan. The idea seems to be that Jordan has traits or dispositions to respond in certain ways that make him a good basketball player. So, the thought continues, if you want to be a good basketball player too, then you should also try to cultivate Jordan-like dispositions. Of course, dispositions can also be bad. My dispositions, when it comes to basketball, are dispositions

young basketball players should not cultivate. We might distinguish the good from the bad dispositions by calling the good dispositions *virtues*, and the bad ones *vices*. This is why those who take character claims as primary are known as virtue ethicists.

Here it is important to notice a subtle difference between evaluative claims like “pain is bad” and “Michael Jordan is a good basketball player.” To differentiate this type of evaluative claim from good-for and good-*simpliciter*, we call this a claim of *attributive goodness*. Attributive uses of good pick out those things that are good-of-a-kind. For example, an eye that can see accurately at a range of 40 meters is better than an eye that can only see accurately at 20 meters. Accordingly, we call an eye that has certain traits that make it better than other eyes, within a certain comparison class, a good eye. This thought is easy to apply to Michael Jordan. He is better than most, if not all, basketball players.

Once we understand this appeal to attributive goodness, it is not hard to predict how the rest of the story unfolds. We generalize the point we’ve been making about Michael Jordan to human beings. The answer to the question—who should we be?—is then determined by identifying those dispositions that good humans have. But what makes for a good, distinctively human, life? One attempt to answer this question, made famous by Aristotle, leans heavily on the idea that human beings have a certain function. On this suggestion, just as what it takes to be a good eye is determined by what it takes for an eye to perform its function well, what it takes to be a good human is determined by what it takes for a human to perform its function well.

But what precisely is the function of human beings? Unlike eyes, this question is not easily answered. We might look to the activities that are unique to human beings (as opposed to plants and animals). But, needless to say, this strategy is not terribly promising. Too many distinctively human activities are either ethically irrelevant (e.g., whistling for fun) or deeply unethical (e.g., murdering for fun). However, even if this way of identifying our function fails, perhaps, by turning to modern biology, ethology, and psychology, we could identify the characteristic function of human beings. Instead of pursuing this inquiry further, let’s turn to the question of how we get from character to conduct.

How do we get from the virtues to deontic verdicts? To answer this question, consider how one cultivates certain dispositions. If one wants to be a good basketball player, one does not play like me; rather, one plays like Michael Jordan. Similarly, if one wants to be a good person, one emulates what good people do. Virtuous persons serve as exemplars. They inform us how we ought to act. Looking to

them tells us what to do. If you understand this point, you understand why I said, when introducing virtue ethics, that deontic verdicts are a function of our character. And here we've arrived at the principle that stands at the heart of

Virtue Ethics. You are required to perform an action if and only if (and because) this action is what a fully virtuous person (acting in character) would do in the circumstances. That is, you ought to do whatever the completely virtuous agent would characteristically do.

That you ought to act only as the fully virtuous act is an attractive idea. Nevertheless, on closer inspection, the idea raises concerns. We want a theory to be capable of advising less than perfect ethical agents, but virtue ethics appears unable to do so.

Virtue ethics supplies an analysis that uses a counterfactual—were the fully virtuous person in my situation, she would (if acting in character) do such-and-such. This creates a problem. Suppose, for example, that you are a coward. Fortunately, your university is offering a class called “Cultivating Courage 101.” It seems like you ought to attend. However, you are required to do only what the completely virtuous person would do. And, obviously enough, the perfectly courageous person is not going to attend. The class would be pointless for this person. Why waste time perfecting perfection? Thus if you ought to do what the perfectly virtuous person would do, you ought not attend either. But not attending would be a mistake. You are a coward and would benefit from instruction. The problem with the criterion supplied by virtue ethics for assigning deontic verdicts should now be clear. The counterfactual used to get us from character to conduct seems to rule out the requirement to do certain actions that would make us more like the virtuous. For what the *fully* virtuous would never need to do is perform actions to make themselves *more* virtuous. And that looks like a devastating problem for a view whose main focus is character development.

My disclaimer should now be familiar: Virtue ethicists have offered responses to this problem. As before, instead of pursuing these responses, now that we have an overview of the three main approaches, let's turn to how we might assess the various theories.

§6. *Theory Assessment*

After reading this introduction, you probably have a preferred theory. All the ethical theories we have looked at thus far, however, fail to get everything we want. So chances are, in your honest moments, you

take your preferred theory to be a failure; it's simply that you think rival theories fail worse. But to competently make this assessment—that the others fail worse—we need to have some handle on what makes one theory better than another.

When devising a theory, the place to begin is with some paradigmatic deontic verdicts that we think any theory needs to capture. So we might start, as we did in (§2.2), with a list of intuitive ethical judgments. When constructing our theory, we want it to fit, at least initially, these judgments. If it fits, then we say it has *intuitive appeal*: The theory does not have counterintuitive implications.

Intuitive appeal, needless to say, can only take us so far. To see why, assume that all we wanted was to have a maximum fit with our ethical judgments. We could achieve this aim by simply claiming that our list of ethical judgments just is our theory. But surely this will not do, for two reasons.

First, sometimes we just don't know what to think about certain cases. I suspect, for example, you do not have a considered judgment about how many people I would need to murder before your murdering to prevent my murdering became permissible. We thus need to have a theory that can tell us, when brute intuition isn't up for the task, what to think. If a theory can tell us what to think in all ethical cases—hard and easy alike—we say it is *complete*. Theories that can only address what we owe to other human beings, for instance, appear to be incomplete. They are mute on what we owe to animals—a topic where ethical questions arise.

The second reason a theory consisting of a list of our ethical judgments will not do is this. Sometimes one's intuitions are not consistent. Perhaps what you think about one case flatly contradicts what you think about another (without you noticing it). Or perhaps you waffle about a case depending on your mood. A theory that makes inconsistent claims, however, cannot be true. So here too intuitive appeal is not enough. We also want a theory that makes consistent claims. If a theory is consistent, we call it *coherent*.

Intuitive appeal, completeness, and coherence are criteria for assessing how well a theory achieves its enumerative aim. Yet, even if a theory achieves this aim, it may turn out to fail to achieve its explanatory aim. We saw above, for example, that contractarianism can supply reason not to hurt animals, but the explanation supplied comes from a contract (not the hurt inflicted on the animal). That explanation appears misplaced, even though this seems like the correct deontic verdict. To take another example, utilitarians say that, if

it fails to bring about the greatest sum total of well-being, slavery is impermissible. That, again, seems like the correct verdict. But does it seem like the correct explanation? Wouldn't the correct explanation turn to the lack of respect shown to the slave, not just the loss in well-being? As these examples help us see, we not only want our theory to come down on the right verdict, we also want it to come back with the correct explanation of that verdict. We call such a theory that can supply this proper explanation *explanatorily efficacious*.

Putting these criteria together, we want an intuitively appealing theory that can answer any ethical question we throw at it, we want to ensure that these answers do not conflict with one another, and we want a plausible explanation for the answers supplied.

A final note. When trying to achieve completeness, coherence, and explanatory efficaciousness, our initial intuitions are likely to be heavily revised. Indeed, ensuring completeness and coherence alone will force serious revision to our initial ethical judgments. Thus we should not treat our untutored intuitions as the final word. A theory that demands serious revision of our initial list of judgments might, for all that, turn out to be the best (perhaps the only) one capable of being complete and coherent. Accordingly, my practice above of introducing cases that conflict with the necessary and sufficient conditions of a theory was all too hasty. Perhaps, for example, upon reflection we are willing to live without constraints and options if act-consequentialism turns out to be the only complete, coherent, and explanatory efficacious theory. But before reaching such a conclusion much more work would need to be done.

§7. *Applied Ethics*

Having learned about the major ethical theories, we can now start thinking carefully about how they might be applied. This volume covers a wide swath of particularly vexing, and extremely important, ethical problems. Their importance should be obvious with only a cursory glance at the table of contents—most of the topics are matters of life and death. As mentioned at the outset, this importance contributes to why these problems are so vexing. They are not problems we can safely ignore. Take abortion. It could be that each of us is a bystander to the murder of millions. Take euthanasia. It could be that each of us is a bystander to the continuation of much needless suffering. These are problems where both action and inaction are

both ethically weighty. So we need to be sure to land on the correct deontic verdict. The trouble is that often we face serious uncertainty.

The uncertainty we face is twofold. First, most of us are unsure which ethical theory is correct. If we knew the correct ethical theory, then we could solve most problems in applied ethics with the following argument form:

1. [Ethical Theory] is true
2. If [Ethical Theory] is true, then actions that fit [Description of Circumstances] are [Deontic Verdict].
3. Description of Circumstances.
4. Deontic Verdict

Sadly, not only do we not know (1), but we also do not know (3). That is, second, we often do not know all of the ethically relevant descriptive facts. Is the fetus a person? What will be the implications of permitting euthanasia for the society at large? Would, for example, it increase the rates of teenage suicide?

In applied ethics, we thus confront unavoidable and momentous problems compounded by normative and descriptive uncertainty. Contained in these pages are a number of attempts to puzzle through what ethics demands in these hard cases. As you shall see, much of the argumentation is tempered by the humbling fact that we have not landed on such a theory; and, even if we had, we might not have an accurate description of the circumstances when going to apply it.

As should now be clear, applying ethics is no easy task. But you should not be put off by its arduousness. Ethics is in its infancy. Much work remains. It's our goal, as we peel through the layers of the debate, to see which view comes out on top. After reading this collection, I expect you'll be left with the sense that ethicists make progress. This makes applied ethics an exciting enterprise. For we do not know what the next stage in the development of ethics will bring.

Appendix: Well-Being

Well-Being

A theory of well-being tries to answer the question: What makes a life noninstrumentally good for the subject living it? A theory of well-being has two main aims: Identify what things are good-for-you, and explain why this is so.

To help think through the various theories of well-being, it is helpful to begin with a list of what is symptomatic of a life that is good-for-you, and what is symptomatic of a life that is bad-for-you. Here are some things that might spring to mind:

Goods

Pleasure

Health

Love

Friendship

Freedom

Knowledge

Accomplishment

Beauty

Bads

Pain

Illness

Betrayal

Manipulation

Unfreedom

Deception

Failure

Ugliness

Both of these lists could be longer, but this will give us a start. Is there some feature about all of the items on this list that might explain why a life with the goods is noninstrumentally better for you than the life filled with the bads?

In trying to locate this unifying feature, it will be helpful to keep in mind why I keep adding the “noninstrumental” qualifier. This qualification is needed because the presence of many things in your life, like ice cream, may increase your well-being. But no one has an ice cream theory of well-being, because, at best, ice cream is only instrumental—it’s a means—to increasing your well-being. And, in devising a theory of well-being, we are trying to capture what noninstrumentally—i.e., what ultimately—makes your life go better. Ice cream does not by itself make your life better. We might think, for example, that ice cream causes your well-being to increase because you take pleasure in experiencing ice cream, and pleasure is noninstrumentally good-for-you. If we thought that the benefits of ice cream are ultimately traceable to pleasure, then we would hold that pleasure facts are the kind of facts that are non-instrumentally good-for-you.

In searching for a unifying feature for our lists above, the importance of the instrumental/noninstrumental distinction should now be clear. Perhaps only one item is noninstrumentally good-for-you, while the rest are simply a means to bringing about the presence of that one item. If so, then just as one disease might explain many symptoms, one of the items on our list of goods might explain all of the others. If, for instance, we could deploy the same strategy we deployed for ice cream—looking to pleasure as the ultimate explanation—to

all of the items on our list of goods, then we could plausibly hold that pleasure is the only thing that noninstrumentally benefits you.

Hedonism

The idea just suggested—that pleasure is what ultimately explains the presence of the other items on our list of goods—has much appeal. Take accomplishments. We can test whether your accomplishments are noninstrumentally good-for-you by asking: What makes a life full of accomplishments better-for-you? If your accomplishments ultimately make you better off, then the answer is simply the accomplishments themselves. Does that seem correct? Imagine all of the famous artists whose accomplishments came after their deaths. Was Vincent van Gogh made better off by his posthumous success? Clearly, his great works made others better off. But our question concerns only van Gogh: Did success make *him* better off? If you think his accomplishments did nothing for him—after all, he was dead when success came—then you will need to give up on the idea that we are made better off by our accomplishments themselves. Suppose that's right. We would then need an account of why accomplishments are regularly thought to be part of a life that is good-for-you. Here we might, like the case of ice cream above, point to the fact that we take pleasure in success.

If we could repeat this strategy for each of the other items on our list of goods, and use pain for our list of bads, then we could hold that your well-being is entirely constituted by facts about pleasure and pain. In support of this idea, it is worth noting that this same strategy is not easily deployed against pleasure (or pain). Pleasure does not seem instrumental to some other good. If I asked you why you are eating ice cream, and you said because it is pleasant, and I then asked why do you want to experience pleasure, you probably would not know how to respond. Pleasure is where the explanation gives out. The theory that holds that pleasure and pain exhaust the things that can be noninstrumentally good-for-you is known as

Hedonism. Your well-being consists exclusively in facts about pleasure and pain. To the extent that (and because) your life contains experiences of pleasure, you are benefited. To the extent that (and because) your life contains experiences of pain, you are harmed.

We have already looked at one attractive rationale for accepting hedonism: its ability to explain many of the things contained in a life that is good for the subject living it. But, by returning to the thought

that what happens after you die cannot benefit or harm you, we can proffer an additional argument in favor of hedonism. When thinking about the case of van Gogh, we noted that it seemed odd to think that he could have been made better off by his posthumous accomplishments. But why think this odd? Because van Gogh, since dead, was not around to *experience* his success. By lacking the possibility of experience, van Gogh's accomplishments never got the chance to benefit him. They never made it into his head, so to speak. If that strikes you as plausible, then you hold the

Experience Requirement. Changes in your well-being must involve your experience. In order to be benefited or harmed, your experience must be affected.

The experience requirement makes hedonism, which holds that well-being essentially involves your experiences of pleasure and pain, hard to resist.

Yet, even if we accept hedonism, questions remain. In order to figure out how well your life is going, we will need to figure out how to calculate how good a particular pleasant (or painful) experience is for you. On a flat-footed proposal, we multiply intensity times duration. This view is called *quantitative hedonism*. Suppose you face a choice between an agonizing surgery and a dull headache. Quantitative hedonism plausibly holds that, if the surgery was extremely short but the headache lasted for the rest of your life, you should go with the surgery.

So far so good. Unfortunately, this invites a worry. Suppose that your life could be made much, much longer—say, a million years. Suppose further that year after year your life contained only the most mild of pleasures, such as the pleasure you get from a good stretch. Each year you get the pleasure equal to a good stretch, but since your life is so long by the time it's over you have experienced the pleasure equal to a million good stretches. That's a lot of pleasure. Compare that life to the life you have now. And let's just assume that, given your much shorter existence, you will end up with less pleasure than that resulting from a million good stretches. Would you want to swap lives?

If you wouldn't, then you reject quantitative hedonism. Presumably, what you think is missing are higher, intellectual pleasures. You think what the life of good stretches lacks is the pleasures that come from relationships, the arts, and so on. These higher pleasures are arguably different in kind from the animal, bodily

pleasures associated with licking ice cream or having a good stretch. What is missing, then, from the strict quantitative calculation is the *quality* of the pleasures. This view is known as *qualitative hedonism*.⁵ To accurately capture the magnitude of the pleasure, we need to multiply the intensity and duration times a quality rating. Of course, with qualitative hedonism, we will need a list of the kinds of pleasures that qualify as of higher quality. (And perhaps a similar list for pains.) But let's leave the specification of higher pleasures aside, and instead turn to a problem that plagues hedonism generally.

Hedonism, either quantitative or qualitative, says that only *experiences* of pleasure and pain matter for your well-being. This implies that your well-being consists wholly of your subjective mental states. Given the attractiveness of the experience requirement, this seems a virtue. Nonetheless, holding that your mental life fully determines your well-being has disquieting implications. For it suggests that your mental life could detach completely from reality, while leaving your well-being unaffected. Imagine, for example, that you have struggled for years to earn your university degree. Today is graduation. As you walk across the stage to shake the provost's hand, your heart is filled with joy. Would our assessment of your well-being change if your graduation were the product of an experience machine—a machine that could, without you knowing the difference, simulate reality? It is tempting to answer Yes. All other things equal, between actual graduation, and the mere experience of it, better for you that your accomplishment is actual. Yet according to hedonism, you-in-the-world and you-in-the-machine, because both enjoy the same mental states, enjoy the same level of well-being. In deciding whether to live a real life or float in a tank with electrodes attached to your brain, assuming the two were experientially identical, you could flip a coin.

Desire-Satisfactionism

If you take the connection to reality to matter essentially to how well your life is going, then hedonism must be abandoned. Perhaps it was a mistake to require that all benefits (or harms) must be experienced. Still, we do not want to venture too far away from your psychology. After all, whatever determines your well-being must benefit *you*. Perhaps, then, we should think of well-being as a fit between your mind and the world. And we have a name for the mental state that has this world-to-mind direction of fit: a desire. If we hold

that the satisfaction of your desires exhausts the things that can be noninstrumentally good-for-you, then we have arrived at a theory of well-being known as

Desire-Satisfactionism. Your well-being consists in having your desires satisfied. To the extent that (and because) your desires are satisfied, you are benefited. To the extent that (and because) your desires are not satisfied, you are harmed.

We can head off a potential misunderstanding by stressing that the sense of “satisfaction” in play is not a feeling. You need not feel satisfied in order to be made better off. Rather, a desire is satisfied when the object of the desire, in fact, obtains. For instance, if you desire that your mother is healthy, this desire is satisfied if and only if your mother is actually healthy. So long as you have this desire, and the world conforms to it, you are made better off. Indeed, you are made better off regardless of whether you know about it or not.

But can desire-satisfactionism supply a unified explanation of our lists of goods and bads? Arguably, it can. Note that all of the items on the list of goods are things that people generally want to have in their lives. You probably desire health, love, friendship, and so on. And you are probably averse to illness, betrayal, manipulation, and the like. So, although we did not initially include the satisfaction of desires on our list, we could hold that the reason all of these things are usually good-for-you is because, if you are like most people, these are things that you want in your life. Desire-satisfactionism can thus explain why the list of goods contains the items that it does by holding that most people desire these things, and getting the things you desire is what ultimately makes you better off. Desire-satisfactionism is thus appealing because it too, like hedonism, offers a unified explanation of what is symptomatic of a good life.

Moreover, we can mount a powerful defense of this explanation by thinking about what it would take to harm you—i.e., what it would take to make you worse off. Suppose, for instance, that I cannot take the responsibility that comes with freedom. What I most want is someone to control my life. That is, my deepest desire is to be enslaved. Now suppose that my enemy comes along and wants to harm me. It would be a very poor strategy for my enemy to make me unfree. To be sure, if my enemy enslaved me, controlling every aspect of my life, the goal of harming me would, it seems, be thwarted. We can imagine my rejoicing every day at my enslavement. The upshot of this example is that by desiring one of

the items on the list of bads, it appears to cease to be bad-for-me. Desire-satisfactionism nicely explains why this is so.

Suppose this leads us to accept desire-satisfactionism. In order to figure out how well your life is going, we still need to figure out how to calculate how much the satisfaction (or frustration) of a given desire influences your well-being. The standard answer holds that the stronger the desire the greater the increase (or decrease). Return to the idea that you want the world to fit your mind. When you have desires, it is as if you are constructing a model of your dream world—what you most want the world to be like. But, clearly, parts of this dream world are more important to you than others. You'd probably be willing to give up on the part of the dream world where you are licking ice cream in order to secure the part where your mother is healthy. You prefer the latter to the former. According to the standard proportional view, having your desire that your mother is healthy satisfied increases your well-being to a proportionally greater extent than having your ice cream desire satisfied.

But should we, as our initial formulation has it, hold that *any* satisfied desire increases your well-being? Remember, a theory of well-being hopes to specify the facts that benefit *you*. And this creates a problem. For in the unrestricted form of the theory given above, even if the object of your desire has nothing to do with you, if satisfied, you are benefited. Yet it seems implausible to hold that desires that have nothing to do with you, whose satisfaction you never find out about, make you better off. The satisfaction of such desires does not redound to your well-being. To illustrate this problem, consider this: At some point, you've likely read a news story about some disaster in a far-off part of the globe. And, while reading, you probably desired that the people facing this disaster turn out all right. Then, after reading, you probably went back to living your life never to think about the disaster or the people facing it again (until now). Suppose that, in fact, unknown to you, the people facing this disaster did turn out all right. Of course, that would be good-for-them. But would it be good-for-you?

If you think not, then we need to add a restriction to desire-satisfactionism: namely, the desires need to be about you. On this restricted form of the theory, only self-regarding desires, if satisfied, increase your well-being. This restriction removes the problem posed by the case of the disaster survivors. Yet how precisely to specify this restriction proves challenging. For example, many people got their heart's desire when Armstrong took his first steps on the moon. Surely this was a good thing for these people, even though the desire

involved—that Armstrong walk on the moon—was not self-regarding. But let's assume that we can properly specify this restriction, and turn to another problem.

We've assumed thus far that satisfying your actual (self-regarding) desires increases your well-being. But surely many of your actual desires are based on false beliefs or irrationality. Think back to what your dream world was like five years ago. What did your five-years-younger self most desire? With this in mind, next imagine that, at that time, your dream world was fully realized. The people that you wanted to be friends with became your friends, the stuff you wanted to own became yours, you got the career you wanted, and so on. Thinking about it now, do you think your five-years-younger self, in getting everything, would be made maximally better off? You probably think that some of the friends, the stuff, and the career would not have been good-for-you-then. And you probably think this because you have in the last five years grown older and wiser. Older and wiser, you realize that many of your desires back then were based on false information and poor reasoning. These ill-informed and irrational desires would not, if satisfied, benefit you back then. To be sure, their satisfaction may have positively harmed your five-year-younger self. If that's right, then we will need to make another revision to desire-satisfactionism. Rather than holding that the satisfaction of your *actual* desires makes you better off, we hold that the satisfaction of your *ideal* desires—the desires you would have if you were rational and fully informed of the relevant facts—make you better off.

However, like the self-regarding restriction, properly specifying this idealization proves a challenge. One immediate concern is how to specify the relevant facts needed to qualify as informed. As a first pass, we might hold that any fact that would influence our desires counts as relevant. But this will not do. Some facts would influence our desires simply by thinking about them, not because they corrected our false beliefs. If, for example, you learned what your ice cream would look like a few hours after consuming it, you might lose your desire to eat it. This implication is unwelcome. Again let's leave the proper way of specifying the idealization aside, and turn to a more general worry for desire-satisfactionism.

We noted that what makes desire-satisfactionism attractive is that it explains why someone cannot harm you by giving you what you want. In making what can benefit or harm us depend on our wants in this way, desire-satisfactionism is a subjective theory of well-being. A theory is *subjective* insofar as it says that something is noninstrumentally

good-for-you just in case you take a positive attitude toward that thing—you give that thing a mental thumbs-up. Desiring (and perhaps taking pleasure) are ways of giving this thumbs-up. But note, subjectivism makes well-being hostage to people’s psychology, and that can be troubling. For people’s psychology can be, and this is a gross understatement, weird. Suppose, for example, that my deepest desire, even if fully informed and rational, is to pinch as many people as possible. Suppose further that everyone indulges my desire. I get what I most want. Over the course of my existence I gleefully pinch thousands of people. Desire-satisfactionism, as a subjective theory of well-being, tells us that I lived a life that was very good-for-me. But a life spent pinching people is both mildly annoying and pointless. Could such a life really be best?

Objective List Theory

If you think that my life of pinching is not best, then you will have to abandon subjectivism, appealing directly to certain objective goods. This theory is known as the

Objective List Theory. Your well-being consists in having your life contain certain objective goods and lack certain objective bads. To the extent that (and because) your life contains these goods, you are benefited. To the extent that (and because) your life contains these bads, you are harmed.

The objective list theory, as its name implies, is both pluralist and objective. It is pluralistic because, as typically conceived, the approach gives up on the hope of finding a single, unifying good that makes your life noninstrumentally better-for-you. The theory instead opts for a collection of goods—like those on our list at the start of the Appendix—each of which is capable, by itself, of increasing your well-being. However, though usually pluralistic, the objective list theory need not be. For the theory may have a very short list, consisting of only one item. We would, nevertheless, want to classify this monistic account as a version of the objective list theory. The constitutive feature of the theory is thus its objectivity. The objective list theory says that the item(s) on the list of goods increase your well-being regardless of whether you take a positive attitude toward them or not. Put differently, benefits and harms can accrue to you independently of your psychology.

The appeal of the objective list theory stems primarily from the conviction most of us have that a pointless or base life cannot be good-for-you even if you endorse it. But the view also gains plausibility

from reflection on how we consider the goods and bads in our lives. When thinking about love or friendship, for example, do you think these things are good-for-you because you take a positive attitude toward them, or do you take a positive attitude toward them because they are good-for-you? In answering this question, imagine trying to explain to your beloved that the reason his love makes you better off is ultimately because you take pleasure or desire it. This, I suspect, will be met with resistance. The reason for this resistance is that most of us intuitively hold that love is *desirable*; it is fitting to desire it for its own sake because it is good (and not the other way around). And the same, it seems, goes for friendship and many other goods.

Suppose, then, that we go in for an objective list theory. In order to figure out how well your life is going, we will still need to be told what items are on the list of goods, and what items are on the list of bads. Additionally, we will need to figure out just how much the presence of each good adds to your well-being, and how much the presence of each bad diminishes it.

One promising way to address both of these issues at once appeals to what it would take to live a perfect human life. On this *perfectionist* approach, our list of goods is determined by what is essential to our natures as human. If friendship and love are demanded by the development and exercise of your essential human capacities, then friendship and love will be included on the list of objective goods. By looking to what perfects you, as a human being with a certain nature, we can locate what is good-for-you. Perfectionism also provides a way of calculating the magnitude of the benefit the presence of each good brings. The goodness of each good is set by how much closer their possession brings you to the ideal. If friendships make you approximate the perfect human life to a greater extent than health, then the presence of friends increases your well-being to a greater extent than health. Of course, the plausibility of perfectionism depends on the plausibility of what is essential to being human lining up with what is good-for-you. But why think these two are connected?

We can bring out the force of this question by remembering that the objective list theory holds that the things that are good for you float free from your attitudes toward them. This attitude-independence implies that the best life for you according to the objective list theory could be one that leaves you miserable. You may be entirely averse to everything that makes for an ideal human life. You might give everything on the list of goods a mental thumbs-down. For example, suppose you are a misanthrope. You hate people. You do not want friends

or lovers. They leave you cold, disgusted. Quiet solitude brings you the greatest pleasure, and is what you most desire. If human relationships are essential to living the ideal human life, then, even though you hate them, friends and lovers would be, according to the objective list theory, good-for-you. If you think that seems mistaken, then you accept the

Resonance Requirement. In order for something to be good-for-you, it must resonate with you. You cannot be benefited by something that you do not endorse or that alienates you.

Accepting this requirement demands that we reject the objective list theory. For the theory's constitutive feature is its objectivity—the independence of the goods from our attitudes. The resonance requirement thus brings us back to subjectivism. The appeal of a subjectivist theory of well-being should now be obvious: It is sensitive to *you*. To maintain an objectivist theory, this sensitivity must be sacrificed.

Where does this leave us? What theory of well-being should be accepted? The answer, in the end, depends on ~~how you come down on~~ the experience and resonance requirements. Rejecting either seems a cost, but accepting either invites serious worries. Perhaps there is a way to adjudicate this conflict. Perhaps a hybrid theory, which holds that what makes someone's life go better is (i) the presence of one of the objective goods, and (ii) desiring or taking pleasure in that good, could get us all that we wanted. But let's leave such exotic possibilities unexplored. Surveying the above theories should be sufficient to give you a flavor of what a theory of well-being needs to accommodate. And, more importantly, it should now be clear why all plausible ethical theories take well-being to be one of the factors that determines the deontic statuses of actions.

Notes

1. Of course, if we also hold that what makes someone deserving is failing to do what is ethically required, then circularity looms.
2. p is a *necessary condition* for some q when the falsity of p guarantees the falsity of q . That is, q cannot be true unless p is true. p is a *sufficient condition* for q when the truth of p guarantees q . That is, p 's being true is enough for q 's being true. ~~Being a mother is a sufficient condition for being female.~~ "If and only if" is a way of capturing both of these conditions. The "if" captures the sufficient condition. The "only if" captures the necessary condition.

3. A terminological point. Much contemporary deontological theorizing favors the language of *rights*—e.g., the right to life, the right not to be harmed, the right to be informed, property rights, and so on. Rights are, in this sense, entitlements—claims to be treated in a certain way. Though popular, talk of constraints and options is more illuminating than talk of rights. Talk of rights often masks whether we are talking about an option or a constraint. For instance, if I say you have a property right to this book, I might be saying that, since it's yours, you have the option to use this book as you see fit—you could read it or burn it. Yet I might also be saying that, since the book is yours, I have a constraint against taking your book or interfering with your reading or burning it. Or I might mean that, if a friend is in possession of the book, she has a constraint against keeping it and you have the option to (perhaps forcibly) take it. Most of the time context makes clear which sense of "right" we are talking about. But sometimes it doesn't. Thus, to avoid ambiguity, I will speak directly in terms of options and constraints. In any case, our discussion below could, with merely cosmetic alteration, be rephrased in terms of rights.
4. A conceptual possibility is worth noting. It may turn out that these sources—God, culture, and rational nature—end up delivering a theory whose deontic verdicts align perfectly with the deontic verdicts supplied by, say, utilitarianism. That is, the views we examine below may, in the end, not turn out to enumerate deontic verdicts that include constraints. Of course, were this surprising result to occur, although both theories would give the same answer to the enumerative question of an ethical theory, they would differ in their answers to the explanatory question. Utilitarians will explain why actions have the deontic verdicts that they do by pointing to the good; whereas theorists who subscribe to one of the theories we will discuss below will point to some other explanation—e.g., God's commands. Having flagged this possibility, we can safely set it aside. As traditionally conceived, all of the theories discussed below are thought to depart significantly in terms of the deontic verdicts they deliver from the consequentialist theories examined above.
5. Is qualitative hedonism a form of hedonism? The worry can be brought out by the following dilemma: Either the quality rating is experienced as pleasant, or it isn't. If the quality rating is experienced as pleasant, then a difference in rating must be merely a difference in terms of intensity or duration, in which case we are back to quantitative hedonism. If the quality rating is not experienced as pleasant, then qualitative hedonism is not a form of hedonism, since it claims that something other than pleasure can increase well-being.