

Ethics as Rational Choice

John Hooker
Tepper School of Business
Carnegie Mellon University
Revised February 2008

Ethics can be viewed as rational choice. A decision must have a consistent rationale behind it, or else it is not an ethical decision. Rationality may not be a sufficient criterion for ethical choice, but it is necessary. It is useful as well. It can provide an objective guide for decision making in business situations and everyday life.

Although rational choice is popularly identified with rational self-interest, the ethical literature has developed a broader point of view. Neglecting the interests of others is irrational—not because it may eventually damage your own interests—but because it is *logically inconsistent*.

This essay presents three specific conditions that a decision must satisfy in order to be logically consistent. They might be viewed as three Laws of Ethics, analogous to Newton's Laws in physics. They help explain our intuitions as to what is right and wrong. More importantly, they are useful for resolving cases in which our intuitions are unclear.

There are several advantages to viewing ethics as rational choice in this broader sense. It provides a conceptual framework that allows you to analyze complex business decisions that involve multiple stakeholders (as nearly all do). It offers a style of argument that can appeal to all parties, since rational choice, by definition, considers all points of view. It provides a vocabulary with which you can articulate an ethical position and defend yourself from pressure to compromise.

Learning to Make Rational Choices

Making rational choices is a skill, and like any skill, it requires practice. Reading this essay is only the beginning. You should work through “Ethical Analysis of Mini-cases” and make sure you follow the arguments. It is impossible to understand the ideas discussed here until you apply them to real ethical dilemmas. Additional exercises will be provided in class and as homework.

Finally, you should practice analyzing cases in other courses, as well as decisions on the job, from an ethical point of view. Psychological research shows that the key to developing expertise in any endeavor is prolonged, continuous, intelligent practice.¹ This goes for ethical decision making in particular.

¹ See Philip E. Ross, “The Expert Mind,” *Scientific American* (Aug. 2006) 64-71.

Even with practice, intellectual analysis alone won't make your decisions for you. You can't just turn a crank and get the right answer. As in any field, judgment and experience are indispensable, and good decisions come from the heart as well as the mind. Yet wisdom must be built on a foundation of rigorous analysis and clear thinking.

A Case Study

The following case study will be used as an example to illustrate the ideas that follow.

While interviewing for jobs, MBA student Jennifer learned about a very attractive opening at Glamour Finance Inc. in New York City. It matched her interests and abilities perfectly. She interviewed on site, and everyone expressed enthusiasm about her potential for helping clients to realize their business goals. Shortly after the interview, however, Glamour announced a hiring freeze, due to the loss of a major client and subsequent cash flow problems. The freeze dragged on through much of the spring semester, and Jennifer's contacts at Glamour could not predict when it would be lifted.

In the meantime Jennifer received two reasonably good offers from firms with whom she had interviewed before going to New York City. She tried to keep her options open, but graduation was near, and her classmates were talking about the great jobs they had gotten. Her parents were asking about her prospects. Her best friend Heather urged her to get real and accept a job. Finally, when her offers were about to expire, she signed with Midwest Consulting in Cleveland, Ohio.

About a week later Jennifer received a call from Glamour announcing that the firm was hiring again. Her employment contract was ready to sign. Distraught, she told Heather about her rotten luck. Heather's reply was, "What's the problem? Just tell Midwest that an unexpected opportunity came up. Employers understand that these things happen."

1. Be Consistent with Your Goals

An action is a means to an end. There is a goal we want to accomplish, a state of affairs we want to bring about. The first condition for rational choice requires that we (a) make up our mind what our ultimate goals are, and (b) try to achieve them rather than something else.

Utilitarianism

This principle was perhaps first clearly articulated under the name of *utilitarianism*. Jeremy Bentham applied the idea to the criminal justice system of eighteenth-century England. At the time, punishment was based on the ancient idea of retribution, or literally, paying back.

Bentham believed that criminal justice should be rooted in reason rather than an emotional desire to make the criminal suffer. If the goal is to reduce crime, then penalties should aim to deter crime rather than make the criminal suffer. A utilitarian would favor rehabilitation and education for convicted criminals if it were shown to reduce crime.

What is Utility?

Since our actions are intended to achieve an end, rationality requires us to get clear on what those ends are. You may attend class in order to improve your grade, but the grade is itself a means to a degree, which perhaps is a means to a better job, which may be a means to a more comfortable lifestyle, and so forth. What is it all for, ultimately? Some of the classical utilitarians favored the idea that there is a single ultimate end, which they called *utility*. Utility can be pleasure, happiness, or whatever you prefer. High grades, a good job and a big salary are instrumentally good, while pleasure and happiness are inherently good. Rationality requires us to decide what is inherently good, whether it be one thing or many, and aim for it.

Maximizing Utility

The utilitarians go a step further. Suppose I regard happiness as inherently good, my ultimate goal. If I can do something to make one person happier, without reducing the happiness of anyone else, then to be consistent I must do it. Otherwise I don't really believe happiness is inherently good! In other words, I must maximize utility across the whole population, or as the traditional formula goes, achieve the greatest good for the greatest number.

I may protest that *my* happiness is inherently good, not just anyone's. But to be rational in distinguishing my happiness from someone else's, there must be some difference that justifies the distinction. If my happiness is somehow of a different quality than everyone else's, then OK, I am at least internally consistent. But I really don't want to claim this. I just arbitrarily distinguish my happiness from that of others. An arbitrary distinction, however, is an irrational one.

The Glamour Finance Case

In the Glamour Finance case, Jennifer might reason as follows. If she breaks her agreement with Midwest Consulting, then since it is late in the season, her replacement may be less qualified and result in less utility for Midwest and its clients. Yet the job at Midwest doesn't really require a person with her background. Her unique qualifications and enthusiasm for the Glamour job, however, may create significantly more value for its clients than their second-choice employee would. So the gain at Glamour probably outweighs the loss at Midwest.²

² If outcomes are uncertain, as they usually are, one can maximize *expected* utility, which is the sum of each possible outcome's utility multiplied by its probability. What if there is no way to estimate meaningful probabilities? I don't know. Let me know if you have any ideas.

Measuring Utility

This analysis seems to assume that there is some way to measure the utility of one outcome versus the utility of another. But how? How do you assign a number to someone's happiness or satisfaction? Even if you can do this, does it make sense to add up these numbers across different people?

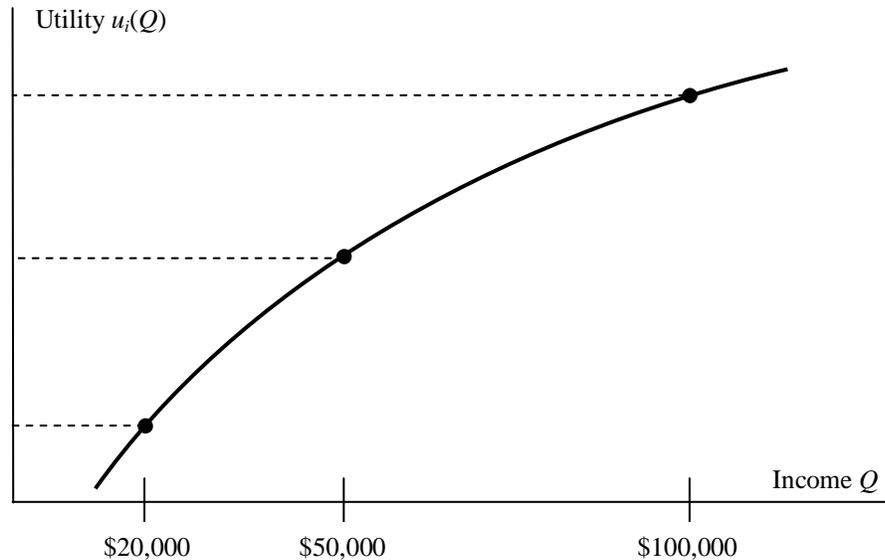


Fig. 1. Hypothetical utility curve for income levels.

Sometimes one action makes everyone better off than any other action. If so, it is the rational choice (this is the *Pareto principle*). But the world is seldom so neat. In 1996 AT&T laid off about 40,000 workers, including some 17,000 managers, typically middle-aged. Every option available to AT&T at the time was harmful to someone. Layoffs would harm the redundant workers, but a bloated payroll would harm the company and therefore everyone that depended on it. The utilitarian test asks whether the layoff would result in gains that outweigh the losses.

The classical approach assumes that each person i has a utility function $u_i(Q)$ that measures the utility of income Q . Although we can't "measure" utility as we might measure sugar for a recipe, we can compare differences in utilities. For example, suppose you make \$100,000 at AT&T, and you are indifferent between two alternatives: (a) taking a lower salary of \$50,000, and (b) taking part in a lottery in which you have an equal chance of keeping your job and losing it (in which case you will find a \$20,000

dead-end job). Then you can say that the utility of \$50,000 income is halfway between that of \$20,000 and \$100,000 (Fig. 1).³

Interpersonal Comparisons

Suppose the AT&T layoff results in income Q_i for each person i , and retaining the workers results in income Q_i' . The utilitarian test says that the layoff is preferable if it results in greater utility; that is, if $\sum_i u_i(Q_i) > \sum_i u_i(Q_i')$.

But how do we compare utilities across different people? If \$100,000 has a certain utility for person i , what utility does it have for person j ? Perhaps less, because person j may have fewer financial obligations or care less about material wealth. How can we put a number on this?

Fortunately, it is not necessary to compare absolute utilities across persons; it is enough to have *unit comparability*. This means that replacing each utility function u_i by a new function $u_i'(Q) = \alpha_i + \beta u_i(Q)$ has no effect on the ranking of alternatives. In other words, the absolute utility level doesn't mean anything across persons, but *relative* utilities mean something. So doesn't mean anything to say that keeping your job at AT&T has utility 163.5 (rather than some other number), but it means something to say that losing your job reduces utility twice as much as getting a 50% pay cut. Utilities determined by the lottery method have this property.⁴

If we assume unit comparability, a classical result of social choice theory states that under certain assumptions about rational choice, it is possible to do a utilitarian calculation. There are multipliers λ_i such that an income Q_i for each person i is better than Q_i' if and only if $\sum_i \lambda_i u_i(Q_i) > \sum_i \lambda_i u_i(Q_i')$. So you can't just add up the utilities to see which decision is better, but you can take a weighted average of the utilities.

³ In general suppose that person i is indifferent between (a) having income Q for sure, and (b) taking a chance of having income Q' with probability p and income Q'' with probability $1 - p$. Then $u_i(Q) = pu_i(Q') + (1 - p)u_i(Q'')$. Values must be arbitrarily assigned to $u_i(Q')$ and $u_i(Q'')$, but once these values are fixed, the utility of other incomes Q can be determined.

⁴ You can check this for yourself. Look at the ratio of the utility loss of a salary cut from Q'' to Q and the loss of a cut from Q'' to Q' :

$$\frac{u_i(Q'') - u_i(Q)}{u_i(Q'') - u_i(Q')}$$

Substituting $u_i(Q) = pu_i(Q') + (1 - p)u_i(Q'')$, this becomes

$$\frac{(1 - p)u_i(Q'') - (1 - p)u_i(Q')}{u_i(Q'') - u_i(Q')} = 1 - p$$

So the ratio is the same regardless of what arbitrary values are assigned to $u_i(Q')$ and $u_i(Q'')$.

Lack of Information

Another potential problem with the AT&T case is that it may be hard to predict the utility outcome of a given decision, even if we know how to measure utility. This is not unusual. We frequently don't have enough information to know which alternative maximizes utility.

Fortunately, the utilitarian test doesn't require omniscience. It requires only that our actions be rationally aligned with our goals. If I have evidence that action A creates more utility than action B, then consistency requires me to choose action A. However, if the evidence is scarce or ambiguous, and I have no particular reason to believe that one action is has better consequences than another, then the utilitarian criterion imposes no particular obligation. Either choice satisfies the utilitarian test.⁵

This doesn't give me a license to be lazy, however. I can't just say that I am not going to worry about which action maximizes utility. I can't simply say that since I have no evidence either way, I can do whatever I want. If happiness is my goal, for example, it is irrational for me to act without trying to predict which action will result in happiness. It is though I want to drive to a certain shop but don't bother to find out which route to take. If I drive around aimlessly, I am simply irrational. It makes no sense to begin the trip without making some effort to look at a map or ask a friend about the route.

How much effort must I exert to be rational? It depends on my goal structure. If I organize my life around creating certain kinds of utility, then it only makes sense to invest some effort into determining how to achieve this. Naturally, if I spend too much time researching the issue, then this will cut into my efforts to achieve my goals. If I spend too much time studying maps, then I will have too little time to shop when I arrive. To be rational, I must find a balance between information gathering and action that I can reasonably believe maximizes the utility that results.⁶

In the AT&T case, the CEO must find a practical tradeoff. Time is short, and a decision is required. As it happens, the effects of layoffs have been studied, and the firm's economists should know the literature. The managerial staff has probably already

⁵ This differs from classical utilitarianism, which requires me to choose that action that actually maximizes utility. The classical theory is consequentialist, while the theory presented here is a deontological reinterpretation of utilitarianism.

⁶ Yes, the issue of what is a proper balance is itself an issue that I should research to some extent. This is related to the concept of *bounded rationality* introduced by my former colleague Herb Simon, but it differs in an important way. For Simon, bounded rationality is a corrective for rational agent models in economics. It takes into account the fact that human beings have limited knowledge and ability to choose rational acts, even acts that are rational in the narrow, self-interested sense used by economists. In the sense relevant here, however, humans can be completely rational when acting without full knowledge, because the evidence at hand may indicate that further investment in fact-finding will consume more effort than it is potentially worth (or to use a term from decision theory, the *net expected value of perfect information* is negative). Incidentally, the problem of balancing knowledge acquisition with action has been formalized in the study of partially observable Markov chains.

projected the effects of a layoff on stock price and other business indicators. Typical individual utility curves are known, and there are probably data on the wealth distribution of the affected groups. AT&T analysts should be able to make a ball-park estimate of total utility impact within a few days. Wouldn't be interesting if large companies routinely employed people to run this sort of analysis?⁷

When There are Several Goals

A goal structure that recognizes a single overriding goal, to which everything else is a means, is probably inadequate. In fact, we will find that it is frequently inconsistent with the other two conditions for rational choice. A rational person recognizes multiple ends that are desirable for their own sake. Perhaps knowledge, beauty, and service to others are valuable irrespective of whether they are associated with happiness or some other good.

Multiple goals can conflict with each other, however. This raises the problem of how to trade them off. Value pluralism is a deep problem and has been the focus of much research in ethics. Ideally, our goal structure should include some understanding of how to balance the goals. Otherwise, it is like wanting a car that has high quality and a low price. It doesn't really tell you anything.

When Utilitarianism Is Not Enough

Passing the utilitarian test is not enough to ensure that an action is ethical. The action must conform to all three conditions for rational choice.

If I fail to vote in the next national election, there is an infinitesimal chance, if any chance at all, that this will affect the outcome. (Even if my vote would make a difference, it is far from clear that I vote wisely!) Furthermore, it is inconvenient for me to travel to the polls, and I may even suffer an accident on the way. Unless I take pleasure in the act of voting, which we may suppose I do not, the utilitarian choice is to stay home. But this does not seem ethical, particularly since most people could use the same excuse for not voting.

Students sometimes make a utilitarian calculation when tempted to cheat on exams. Suppose for the sake of argument that grades don't depend on the distribution of scores, so that one person's dishonesty has no effect on the welfare of others and boosts his own. The utilitarian act is therefore to cheat. The calculus may change if cheating propels cheaters into a career with inadequate knowledge. But it seems wrong to cheat even if they learn the material later.

⁷ Some businesses have already adopted Triple Bottom Line accounting as part of their corporate charters or by-laws. This concept is related to maximizing total net utility due to its concern for People, Planet and Profits (the three bottom lines). The United Nations International Council for Local Environmental Initiatives has established standards for TBI accounting.

We will see that cheating and failing to vote violate the second condition. The utilitarian test can therefore be misleading if it alone is applied, particularly to decision making at the level of the individual.

Utilitarianism for Policy Decisions

Utilitarianism tends to be a more reliable indicator of the right choice when applied to *policy* decisions that govern a large number of people, as when a government passes a law or a large corporation adopts a regulation.

If a government wishes to enact a mandatory voting law (as is done in Australia, Belgium, Singapore, Uruguay, and elsewhere), the utilitarian test is helpful. A rule that says, “Don’t vote if it is inconvenient,” clearly doesn’t maximize utility, because it leads to the collapse of democracy. Even the opposite rule, “Vote if you are eligible,” seems suboptimal, because it asks the seriously ill to drag themselves to the polls. Democracy will not suffer if a few sick persons stay home. Perhaps an optimal voting rule would arrange eligible voters on a scale according to the inconvenience of voting and make exceptions for those at the high end. The cut point can be located so as to maximize overall utility.

A utilitarian analysis is not a reliable way to resolve Jennifer’s job dilemma, since it is an individual decision rather than a policy decision. Yet it may help her MBA program design its job placement policy. Some business schools (including the Tepper School) deny their career services to students who renege on an employment contract. The justification is that a strict policy creates a good reputation for the school, which attracts more recruiters, which in turn makes everyone better off on the average—even if some people like Jennifer pay a price. This is a classic utilitarian justification.

Utilitarianism and Justice

Even when utilitarianism is restricted to policy making, it may not always provide a just solution. It is true that there is already a strong principle of justice in utilitarianism simply because everyone’s utility is given equal weight in the calculation. One cannot (arbitrarily) give greater weight to members of the upper class or of a certain race, for example. Furthermore, utilitarian solutions show at least some preference for equality due to the principle of decreasing marginal utility. As one acquires more resources, utility rises at a decreasing rate. A fixed amount of resources may therefore bring more utility when they are distributed widely rather than concentrated in a few persons. This introduces a bias in favor of more equal distributions.⁸

⁸ The utilitarian bias toward equality is limited, as shown by a mathematical analysis. Let the utility that results from giving Q units of some resource to person i be $c_i Q^p$. The exponent p is less than 1 when there are decreasing marginal returns. The coefficient c_i indicates the person’s ability to use the resources; c_i is presumably larger for persons who are intelligent, well positioned in society, or advantaged in other ways. The goal is to maximize $\sum_i c_i Q_i^p$ subject to $\sum_i Q_i = R$, where Q_i is the amount of resource allocated person i and R is the total amount of resource available. If $p = 1$ (i.e., marginal utility is constant), then the most advantaged person gets all of the resources. Otherwise the problem can be solved by associating Lagrange

But utilitarian calculations nonetheless endorse highly unequal and apparently unjust distributions if they happen to maximize overall utility. They may require us to pay CEOs exorbitant salaries while eliminating the minimum wage, deny health care to the poor, refuse to hire the handicapped, and so forth, if these policies raise the average utility despite hurting those at the low end. To determine whether these policies are ethical, we must apply the other conditions for rational choice.

2. Have a Consistent Rationale

The condition for rational choice tells us to decide what we want and then consistently aim for it. This may be an inadequate guide for individual choices, however, and it may not ensure distributive justice. We need a second condition that addresses individual duty and fairness.

Acting for Reasons

The second condition is based on a premise even simpler than the first condition: we always act for a reason. The reason may be to achieve some ultimate goal, as assumed by the first condition, or it may be some other sort of reason. Whatever the case, there should be something we take to justify the action. For example, if I choose not to vote, there must be some reason I so choose. Perhaps it is because voting is inconvenient.

The argument from here out is reminiscent of the utilitarian argument. If a reason justifies an action for me, then it justifies the action for anyone. When I choose an action for myself, I choose it for anyone who has the same reason. Either the reason justifies the action or it doesn't. If it does, then it justifies the action for anyone to whom it applies.

For example, I must regard my reason for not voting as a reason for anyone's not voting. I might protest that my reason does not work for people who enjoy voting. Then I really have two reasons for not voting: it is inconvenient, and I don't enjoy it. If these are really my reasons, then I am committed to saying that they justify nonvoting for anyone to whom they apply.

multiplier λ with the constraint. The optimal solution satisfies the Lagrangean equations $pc_iQ_i^{p-1} = \lambda$ for each i and $\sum_i Q_i = R$. It is therefore

$$Q_i = R \frac{c_i^{1/(1-p)}}{\sum_j c_j^{1/(1-p)}}$$

This gives more resources to the more gifted persons but no longer gives everything to the most gifted. As the exponent p drops to 0, the allocation becomes proportional to c_i . So the most nearly equal distribution that a utilitarian can endorse is to give each person resources in proportion to that person's ability to use them.

Generalization Test

So far there is nothing wrong with my decision not to vote. But one of the reasons I don't vote is almost certainly that others will vote even if I don't, and democracy will be preserved. If it were otherwise, I would be first in line at the polls. So my rationale presupposes that most other people will vote. But it also presupposes that most other people won't vote. When I choose not to vote due to inconvenience, and so forth, I choose the same for everyone to whom these reasons apply, and they apply to most people. So my rationale presupposes that most other people will vote and most other people won't vote. This is irrational and inconsistent and therefore unethical.

This is sometimes called the *generalization test*, although it is really a rationality test. It is historically associated with the philosopher Immanuel Kant, who called it a "categorical imperative." It can be phrased: *the reason for your action should be consistent with the assumption that everyone who has the same reason will act the same way*. Kantian ethics, like utilitarianism, is at root a call to rationality.

The generalization test is different from the utilitarian test because it doesn't look at the consequences of the act in question. To satisfy the first condition, an action must result in as much utility as any other available action. The second condition doesn't care about how much utility the action creates, as long as the rationale is consistent. Of course, maximizing utility could part of be the rationale for an action, but this is incidental as far as the second condition is concerned.

Let's apply the generalization test to cheating on an exam. The student mentioned earlier cheats presumably because it will improve his grade and career prospects. But it will improve his career prospects only if most people are honest enough for grades to be meaningful, despite the fact that they have the same reasons to cheat. So part of the student's reasons for cheating is the assumption that other students will not cheat even though they have the same reasons to cheat. So the student's rationale is inconsistent with the assumption that others who have the same rationale will cheat. It fails the generalization test and is therefore irrational and unethical.

A special case of the generalization test is the *free rider* principle: other things being equal, one shouldn't be a free rider on the efforts of others. In some European countries people sometimes ride the city bus without paying the fare, because the driver doesn't check whether they paid. Free riding is possible only because most people pay the fare even though they have equally good reasons to ride for free.

Why Acts Must Have Reasons

This whole affair is based on the premise that acts are based on reasons that are taken to justify the action. Why should this be so? Because it is how Western culture distinguishes free action from mere behavior: free agents are *rational* agents and therefore act for a reason. If a mosquito bites me, this is mere behavior. I don't judge the mosquito morally, because it didn't "freely choose" to bite. The bite was merely the

result of chemical reactions and whatnot in the mosquito's body. Human actions can also be given a causal explanation of this kind, but they can be explained in a second way: by talking about the agent's reasons. It makes no sense to say that the mosquito bit because she thought to herself, "I'm going to bite that human because I'm hungry and I think I can get away before I get swatted." However, it is often very reasonable to explain human actions by attributing reasons to the agent.

In principle the behavior of computers and robots could someday be more easily explained as the result of the machine's own deliberation than as the outcome of an algorithm. This would make machines moral agents with duties and rights.

When a rationale is inconsistent, as when it fails the generalization test, then it is not really a rationale. Inconsistent reasons cannot form a rationale for anything and therefore cannot explain behavior. This means that action in violation of the second condition (or any condition) for rational choice is not really action at all, but mere behavior. It is like the mosquito's bite, which can be given only a causal explanation.

What Is the Real Reason?

Mary Smith has discovered a clever way to justify her failure to vote. She is staying home from the polls because voting is inconvenient, others will vote anyway, and she is Mary Smith. She points out that this is generalizable. It is perfectly consistent with this rationale for every Mary Smith who finds voting inconvenient to stay home. So her refusal to vote satisfies both the first and second conditions for rational choice.

This rationale is generalizable, all right, but there are two problems: it's not Mary's rationale, and it's not a rationale at all. If it were Mary's rationale, she would have to vote if she learned that she is really Mary Jones. In reality she would still refuse to vote, because her name has nothing to do with it. In fact her "rationale" is not a rationale at all, for the same reason. There is no discernible connection between one's name and whether one should vote.⁹ A rationale must give some sort of reasonable explanation for the action.

What's to prevent Mary from nonetheless *saying* that her name is part of the reason, and "justifying" her action on that basis? Nothing. She can say anything she wants. The conditions for rational choice don't force anyone to be rational. They only tell you what it means to be rational.

When Conditions for Rational Choice Seem to Conflict

The utilitarian test says that failure to vote is OK because it maximizes utility, while the generalization test says it is wrong. Isn't this a conflict? No. The utilitarian test doesn't

⁹ ...unless of course she is registered to vote under one name and not another. But in this case the reason she fails to vote as Mary Smith is that she is not registered under that name, not the simple fact that she is Mary Smith. Failing to vote because you are not registered to vote is perfectly generalizable, although failure to register if you are eligible, simply because it is inconvenient, is probably not generalizable.

say that staying home is OK, only that staying home passes this particular rationality test. An act must pass all three tests to be ethical.

But there may be a more serious problem. A failure to vote violates the generalizability test, but voting also violates the utilitarian test because it doesn't maximize utility. So it seems we are damned if we do and damned if we don't.

Not really. Both tests evaluate whether there is a consistent rationale for the action in question. If I value happiness, or whatever I take to be a measure of utility, and I have an opportunity to increase overall utility, then I am irrational not to do so. However, in this case, I can't maximize overall utility without falling into another kind of inconsistency (violation of the generalization test). If I maximized overall utility in this way, I would not really be acting, because an action must have a consistent rationale to be an action. So I have no opportunity to *act* so as to increase overall utility.

Another way to put this is that a single-minded goal of maximizing utility, measured by wealth or happiness or whatever, is irrational in the world as we know it. My value structure must be more complex than this. For example, I might see participation in a democratic process as valuable for its own sake.

The Glamour Finance Case

Jennifer is actually faced with two decisions. She decided to accept the Midwest Consulting offer, and she now must decide whether to sign with Glamour Finance.

Jennifer's acceptance of the Midwest offer may not be an act at all. She was under pressure from her peers and parents to get a job. It might be hard to construct a rationale for Jennifer's eventual acceptance of the offer. Perhaps only a psychological reason, a *cause*, could explain it: she gave in to pressure without thinking matters through clearly. If so, her choice was not an ethical act because it was not an act at all.

As for the Glamour offer, let us suppose that Jennifer decides to take it. The key question is, what are the reasons? The job is better suited for her, and she wants it much more than the Midwest Consulting job. Are these reasons generalizable? Clearly not, because Jennifer's decision to break her commitment implies the possibility of making a commitment in the first place. If students abandoned their commitments whenever they got a better offer, commitments would be meaningless and companies would disregard them. Jennifer's rationale for breaking the contract presupposes that most other students don't regularly break employment contracts even when they have the same reasons to do so. This makes it ungeneralizable and unethical.

Jennifer may protest that her reasons are more complex than I have made out. She acted in good faith all along. The offer from Glamour was unexpected, and she didn't interview after signing with Midwest. These are all part of her reasons. Making exceptions in such cases may be consistent with the practices that make commitments possible. If so, breaking the contract passes the generalizability test.

Yet on reflection this, too, seems unlikely to be generalizable. Imagine what it would be like. You could interview for a number of jobs. Then if one of the companies offers you a job, you say, "Sure, I'll take the job. I won't look for another job, but if some company offers me a better one, I'll take it, of course." Since companies receive no commitment, they would have no incentive to offer one. So when you interview with a company, the interviewer would say, "We are happy to hire you. We won't look for someone else, but if someone sends us a better resume, we'll hire that person and forget about you, of course." Obviously in this regime there are no meaningful commitments.

There may be another way for Jennifer to escape from the commitment. Students in this predicament often point out that employment contracts generally have some kind of escape clause. Suppose for the sake of argument that the contract allows the employee to resign after giving notice, perhaps two weeks. Perhaps Jennifer could simply tell Midwest she is giving them her two-week notice. This is perfectly legal.

Actually we have assumed along that any option Jennifer chooses is legal. Breaking the law is ungeneralizable, because if people routinely broke the law for mere convenience, there would be no laws to break. Although breach of contract is not a crime, it is a tort, which is a violation of civil law. So let's grant that there is some way to renege on the commitment to Midwest legally. The question is whether it can be done ethically.

Jennifer might argue that what is legal is ethical in this case. According to the contract, she only promised to work until giving a two-week notice. But she clearly promised more than this, whatever the contract may say. Everyone recognizes that a change in life circumstances may require a change of job, such as marriage, children, or unexpected financial problems. Or after working at a company for a while, one may be ready to move on to another position. In such cases, giving notice is reasonable and expected. But barring unforeseen circumstances, Jennifer promised to work for Midwest *for the time being*. Similarly, Midwest promised to employ here *for the time being*, unless there is an unexpected change of circumstances, such as a serious financial setback, or Jennifer's failure to do the work.

Jennifer is still unconvinced. She insists that there is still something special about her case. She not only waited a long time before signing with Midwest, but the offer from Glamour was delayed due to factors beyond the company's control. Glamour wanted to hire Jennifer all along, but she got caught in a timing trap. Allowing everyone in these special circumstances to break employment contracts would not undermine the practice of making commitments.

But there are any number of "special" circumstances that could have put Jennifer in her present position. The HR chief at Glamour is ill and gets behind on offers. An internal audit delays hiring authorization. The media reports on Glamour's financial situation are unduly pessimistic. Every hiring situation is special in some way. If people break contracts whenever the situation is "special," then contracts become impossible. To satisfy the second condition for rational choice, Jennifer must think about the whole

range of situations in which she would break the contract—and why—and only then apply the generalization test.

Back to Consequences

I said earlier that the second condition for rational choice doesn't care about the consequences of the act in question. Yet applying the generalization test seems to require knowledge of consequences. One must know that if all students broke employment contracts under certain conditions, the practice of making employment contracts would be undermined.

This is a consequence, but it is a hypothetical consequence of generalized behavior rather than a consequence of one individual's action. In fact, there is no need to predict the actual consequences. Since we are only testing for consistency, the relevant factor is what the agent reasonably believes are the consequences.¹⁰

When I fail to vote, my understanding of the world tells me that if everyone with my reasons didn't vote, democracy would collapse. Whether I am right about this doesn't matter. What matters is that I believe it, and this generates the inconsistency. What if I fail or refuse to acknowledge that nonvoting would defeat democracy? Then I am already irrational and therefore unethical.

Things would be different if I lived in a country where most people love to vote and don't find it inconvenient at all, and if they are politically similar to the few who hate voting. Then I could reasonably believe that democracy would get along just fine if everyone who found voting unpleasant and inconvenient stayed home. In this case my nonvoting would be generalizable.

So the generalization test might be rephrased: *the reason for your action should be consistent, based on your understanding of how the world works, with the assumption that everyone who has the same reason will act the same way.*

The Veil of Ignorance

The philosopher John Rawls proposed a vivid way of understanding the generalizability of reasons. On his view, you must make a decision *without knowing who you are*. When Jennifer decides to sign with Glamour, for example, she could be Jennifer, a manager at Midwest, or another student. As Rawls put it, she must decide behind a "veil of ignorance" as to her station in life. She will find out who she is only after she makes the decision. Her reasons must be sufficient for her choice no matter who she turns out to be, which is the heart and soul of the second condition for rational choice.

It is easy to misinterpret this criterion. It does not mean that one should figure the probabilities and maximize the expected outcome. An AT&T executive, for example,

¹⁰ The same is true of the utilitarian test as interpreted here, although the historical utilitarian test is generally interpreted to be concerned with the actual consequences.

might decide it is a good bet to lay off some middle-aged employees to make a company more profitable. This would be a disaster for her if she were one of those terminated, but there is a much greater chance she would be someone who benefits from the layoff. She is willing to take her chances. Rawls says this is not enough. She must construct a justification for the layoff that she would find equally convincing if she were transported into the body of one of the redundant employees.

Distributive Justice

Rawls used his idea of the veil of ignorance to analyze distributive justice. He arrived at two principles:

Liberty Principle: A policy must result in the greatest basic liberty for everyone.

Difference Principle: A policy must not result in inequality unless the inequality makes everyone better off.

The Liberty Principle might be defended on the grounds that agency itself requires a certain amount of freedom to achieve one's purposes. There can be no consistent rationale for restricting basic liberty, since having a rationale for any action presupposes the ability to choose one's actions.

The Difference Principle says, for example, that it would be unjust for the government to raise taxes on the poor and reduce taxes on the rich, unless this would make both poor and rich better off. This leads to a "lexicographic" comparison of utilities (maximize the utility of the worst off, then the utility of the second worst off, and so on).¹¹

Rawls clarifies the Difference Principle by saying that it applies only to how society provides people opportunities, such as job openings and access to health care (as well as burdens like taxes). So it is unjust to create some jobs that pay less than others unless this allows all jobs to pay more. Furthermore, everyone should have a chance to qualify for these opportunities.

It is not hard to justify equality of opportunity, since otherwise we must decide who gets the opportunities. We must decide who will have an affluent family and access to education, and who will be come from poor circumstances with no such access. Since the choice must be arbitrary, it must be irrational, and therefore no such choice should be made.

It is harder to justify the remainder of the Difference Principle. Perhaps it fails the veil of ignorance test: no one could rationally agree with creating jobs that pay less than necessary after learning that one holds such a job. Rawls uses a social contract argument.

¹¹ Let $u = (u_1, \dots, u_n)$ and $u' = (u'_1, \dots, u'_n)$ be two distributions of utilities to n persons, ranging from the worst off to the best off in each case ($u_1 \leq \dots \leq u_n$ and $u'_1 \leq \dots \leq u'_n$). Then u is preferable to u' if for some k , $u_i = u'_i$ for $i = 1, \dots, k - 1$, and $u_k > u'_k$. This lex max criterion can also be the basis for a utilitarian test under certain conditions; see the Appendix.

He says basically that we could all agree in advance (from an “original position” behind a veil of ignorance) on inequalities that make us better off no matter who we are, but we could not agree on an arrangement that makes some of us worse off than necessary.

3. Be Consistent with Who You Are

The first and second conditions for rational choice impose formal conditions on our decisions. The third condition begins to examine the content of those decisions, while still using rationality as a guide. It says that rational decision making must start by making sense of our role in the world. We can't decide what to do until we decide why we're here.

Naturally, cultural tradition, religion, and philosophy have much to say about this. The third test for rational choice, however, tries to isolate minimal conditions for any rational conception of one's purpose—conditions that everyone can agree with, insofar as they are rational (as defined in the Western tradition). The idea goes back at least to Aristotle.

Teleological Explanation

Teleological explanation is a way of understanding the purpose or function of things. (*Telos*, from the Greek, means purpose.) It is not hard to identify the purpose of artifacts like cars or computers, since their designers give them a purpose. But a teleological approach can help us understand other things as well, such as the human body. It is true that science emphasizes causal explanation. It explains respiration, for example, by talking about chemical reactions that transform nutrients to acetyl coenzyme A, which initiates the Krebs cycle, which produces adenosine triphosphate (ATP), and so forth. Yet the complexity of the human body would be unintelligible if we did not give it a teleological or functional explanation as well. The function of the heart is to pump blood, the function of the lungs is to provide oxygen to the blood, and so on. The molecular biologists who tell us about chemical reactions acquired their first understanding of the body when their kindergarten teachers told them about the heart and the lungs.

Virtues

There is an old tradition that finds a purpose for human life in the scheme of things much as we find a purpose for the organs of the body. We say that the heart's function is to pump blood because it does so and is uniquely suited to do so. Similarly, human beings are uniquely suited to certain kinds of activity. We are rational beings. We can apprehend beauty. We are capable of trust, loyalty, friendship, honor and courage in a self-conscious way that apparently characterizes no other creature. One might conclude that our purpose here is at least in part to bring these qualities to the world. They are traditionally called *virtues*. No one can prove that this is why we are here, but no one can prove that the heart exists to pump blood. Nonetheless it is a hypothesis that helps us make sense of things.

Humans are also uniquely capable of monstrous cruelty, and one may wonder why this would not also be a virtue. Yet no organ of the body can kill like the heart, since a slight electrical disturbance will do the trick. The heart's pumping behavior, not its ability to kill, helps us to understand how the body works. Similarly, cruelty doesn't help us explain human existence, but makes that existence even harder to explain. Rather, it is by regarding human beings as the world's source of rationality, aesthetic sensibility, trust, loyalty, honor, friendship and courage that we are able to make some sense of our predicament.

Another way to put this is that the virtues are part of our *essence*; they help define who we are. A heart can have two chambers or four, but an organ that doesn't pump isn't a heart. A human being can be tall or short or male or female, but a human being without rationality, without any of sense of beauty, with no understanding of friendship, is not fully human.

Being Consistent with Who We Are

The first condition for rational choice says that we must act with some ultimate goal or goals in mind. The second condition states that there must be a consistent rationale behind our choice of action. The third condition says that any decision about what to do must follow rationally from some understanding of why we are here. This calls for teleological explanation, and as just argued, a teleological analysis concludes that a purpose of human life in general is to exhibit the virtues. It is therefore irrational, which is to say unethical, to stray from them.

Integrity (wholeness) is the result of being consistent with who we are. The third condition for rational choice is really an ethic of integrity. We lose integrity when we compromise our honor, abandon a friend, or do a shoddy job that is beneath our intellectual ability. There are various psychological expressions for this: we can't live with ourselves, can't look at ourselves in the mirror, can't sleep at night, and so forth. We are no longer whole persons because we have walked away from who we are as human beings.

The pressures of business life can make it difficult to develop some of the virtues, such as loyalty and friendship. A white-collar or managerial employee tends to be a free agent who moves from one firm to another in pursuit of better salary offers and advancement opportunities. The employee feels free to depart for greener pastures in the middle of a company project, and the firm feels free to terminate the employee in mid-career. Workers and managers who share a commitment to each other arguably live fuller lives than those linked only by transitory economic incentives, if only because they can develop their capacity for loyalty and friendship. On the other hand, a competitive business environment encourages one to develop intellectual competence, which is an equally important virtue and equally part of one's essence as a human being.

Culturally Specific Purposes

People frequently adopt ultimate purposes that are rooted in a cultural or religious tradition. One may define one's essence as being a member of the family, part of the community, or a servant of God. In Western cultures it is not uncommon for people to define themselves at least in part by their career. Choosing a career is a key rite of passage for many young people, and they don't fully define themselves until they settle on one. The career therefore helps to define the ultimate purpose on the basis of which rational decisions are made.

One can argue that the three conditions for rational choice are themselves culturally specific because they give primacy to rationality, reflecting the Western view that human beings are autonomous, rational individuals. Yet insofar as one aspires to rationality, as conceived in the Western tradition, one must observe the necessary conditions for rational choice.

The Glamour Finance Case

A virtue ethics perspective would argue for Jennifer's accepting a job in which she could develop her unique abilities, such as the Glamour Finance job. Walking away from Midwest Consulting is perhaps not a breach of loyalty, either, because no relationship has really been established. On the other hand, a breach of agreement would compromise Jennifer's honor and integrity. Honor is part of integrity, because it is part of who we are.

It is never consistent to act contrary to a virtue, unless it is for the sake of another virtue. In this case there is in fact a conflict of virtues, which must be resolved by striking a balance. The ancient Greeks viewed the ability to find a balance as itself a virtue, which they called *sophrosyne* (a word with no English equivalent). Jennifer must find a course of action that is most consistent with who she is.

Appendix

Some Results from Social Choice Theory

We assume that every individual i has a real-valued utility function $u_i(x)$ that measures the utility of any state of affairs $x \in X$, where X is the set of all possible states of affairs. Let $u = (u_1, \dots, u_n)$ be a vector of utility functions, one for each individual. Also let R_u be a relation that indicates which states are preferable to others, taking into account the utility of all individuals as represented by the vector u . $xR_u y$ means that x is at least as good as y , and $xP_u y$ means that x is strictly preferable to y ($xR_u y$ but not $yR_u x$).

The goal is to find a *social welfare function* w that indicates which states are preferable to others. That is, $w(u_i(x)) \geq w(u_i(y))$ when $xR_u y$ and $w(u_i(x)) > w(u_i(y))$ when $xP_u y$.

It is assumed that any rational set of preferences satisfies certain properties:

Weak Pareto condition. A state of affairs that is better for everyone is preferable. That is, for any $x, y \in X$, if $u_i(x) > u_i(y)$ for all i , then $xP_u y$.

Anonymity. Renaming individuals has no effect on preferences. That is, if for permutation π of $1, \dots, n$ we have $u_i(x) = u_{\pi(i)}(x)$ for all i and all $x \in X$, then $R_u = R_{u'}$.

Independence of irrelevant alternatives. Preferences among a subset of states depend only on the utilities of those states. That is, for any u, u' and any subset $Z \subset X$, if $u(x) = u'(x)$ for all $x \in Z$, then $xR_u y$ if and only if $xR_{u'} y$ for all $x, y \in Z$.

Interpersonal comparability is indicated by the extent to which individual utilities can be transformed without affecting preferences. A vector of real-valued functions $\phi = (\phi_1, \dots, \phi_n)$ is an *invariance transformation* if it has no effect on preferences; that is, if $u'_i(x) = \phi_i(u_i(x))$ for all i and all $x \in X$, then $R_u = R_{u'}$. The *invariance class* (class of invariance transformations) indicates the degree of interpersonal comparability. The larger the class, the less the comparability.

Unit Comparability

Unit comparability occurs when the invariance class contains any transformation ϕ for which $\phi_i(u_i(x)) = \alpha_i + \beta u_i(x)$, where $\beta > 0$. In other words, preferences are unchanged when we offset each individual's utilities a different amount and rescale each by the same positive multiplier.

It can be shown that given unit comparability, states can be ranked by taking a weighted average of the individual utilities. That is, there is a social welfare function of the form $w(u(x)) = \sum_i \lambda_i u_i(x)$ where each $\lambda_i > 0$. This result is related to linear programming duality and is based on the same mathematical theorem (the Farkas Lemma).

In fact, linear programming can be used to find the multipliers λ_i . Just write the linear constraint $\sum_i \lambda_i u_i(x) \geq \sum_i \lambda_i u_i(y) + \epsilon$ for every $x, y \in X$ for which $xP_u y$ and find a feasible solution of this constraint set along with $\lambda_i \geq \epsilon$ for all i (where ϵ is a small positive number).

Noncomparability

Utilities are not comparable at all when the invariance class contains all monotone increasing transformations ϕ . In this case it can be shown that there is a *dictator*: a single individual whose utilities determine which states are preferable. That is, there is an individual i such that $xP_u y$ whenever $u_i(x) > u_i(y)$.

Full Comparability

Full comparability occurs when the invariance class contains any transformation ϕ for which $\phi_i(u_i(x)) = \alpha + \beta u_i(x)$, where $\beta > 0$. In other words, preferences are unchanged when we offset each individual's utilities by the same amount and rescale each by the same positive multiplier.

It can be shown that given full comparability, states can be ranked by a welfare function that considers the average utility and the distribution of utilities, where the distribution is described by a positively homogeneous function of deviations from the mean. (A function g is positively homogeneous if $g(\lambda x) = \lambda g(x)$ for all x and all $\lambda > 0$.) So there is a homogeneous welfare function of the form

$$w(u(x)) = \bar{u}(x) + g(u(x) - \bar{u}(x)e)$$

where g is homogeneous, e is a column vector of ones, and $\bar{u}(x) = \frac{1}{n} \sum_i u_i(x)$.

Some examples of welfare functions having this form are:

- $w(u(x)) = \bar{u}(x) + \gamma \min_i \{u_i(x) - \bar{u}(x)\}$, which yields a utilitarian welfare function when $\gamma = 0$ and a Rawlsian function when $\gamma = 1$.
- $w(u(x)) = \bar{u}(x) - \frac{1}{n} \left(\sum_i (u_i(x) - \bar{u}(x))^2 \right)^{\frac{1}{2}}$, which measures inequality by the standard deviation (divided by $n^{1/2}$).
- $w(u(x)) = \frac{1}{n^2} \sum_i (2i - 1)u_i(x)$, with individuals indexed so that $u_1(x) \geq \dots \geq u_n(x)$. This measures inequality by the Gini coefficient.

Ratio Scale Comparability

Ratio scale comparability occurs when the invariance class contains any transformation ϕ for which $\phi_i(u_i(x)) = \beta u_i(x)$, where $\beta > 0$. In other words, preferences are unchanged when we rescale each individual's utilities by the same positive multiplier.

Given ratio scale comparability, states can be ranked by a homothetic welfare function; that is, $w(x) = g(h(x))$, where g is monotone increasing and h is homogeneous. If w is assumed to be additively separable and $u(x) > 0$, then for some real number r , w has the form

$$w(u(x)) = \sum_i \frac{u_i(x)^{1-r}}{1-r}$$