CODES OF CONDUCT

Behavioral Research into Business Ethics

DAVID M. MESSICK AND ANN E. TENBRUNSEL
EDITORS

Russell Sage Foundation / New York
Summary

The Psychology of Business Ethics

Russell Hardin

The central problem of ethics is the problem of motivation. Indeed, one of the oldest characterizations of ethics is that it is about the conflict between self-interest and other concerns, such as the interests of others or doing the right thing in some sense. Morality is supposed to help us make the trade-off between our own interests and other concerns. To a large extent, that means that ethics is a subject for psychology as much as for philosophy, and this was evidently the view of one of the greatest moral theorists, David Hume, in his massive Treatise ([1739–40] 1978) the third book on ethics follows that on psychology, which focuses especially on the passions. It is therefore apt that psychologists and philosophers have joined in a conference on behavioral research and business ethics. It will be even more apt if we hereafter contribute extensively to this joint debate, which reached its apogee in the work of one lonely Scottish-Enlightenment philosopher.

In discussions of ethics in business from a psychological perspective, there are several tendencies that mirror tendencies in the modern history of moral debate, some of which are about matters that are almost guaranteed to confuse. I will address three such issues: the problems of facts and values, altruism, and intuitions. I will then turn to an institutional theory of moral behavior—that is to say, a theory of moral behavior in institutional and organizational settings. This theory should be especially congenial to social scientists, and indeed several of the psychologists represented in this book discuss elements of it or parallel to it (for example, Darley, chapter 1; McGill, chapter 12; Wagenaar, chapter 18).

To moral theorists in general, the stance of social scientists would count as remarkably uniformly consequentialist. That is, social scientists (perhaps psychologists and economists more than most) are especially inclined to count as moral that which produces good consequences. Moreover, they are even relatively utilitarian, in that they suppose good consequences are primarily (or even exclusively) related to human welfare. When my account requires specific moral assumptions in order to carry through a particular argument, I will resort to utilitarian principles. It is important to note, however, that the central drive of the account is to fit behavior to organizations and their plausible goals, not specifically to a utilitarian's goals. The overall argument here should therefore be of interest to most moral theorists, including those most ardently opposed to utilitarianism. The rough structure of the argument is that, whatever our moral principle is, we want our organizations to achieve it. If they are to do so, they will have to be properly designed for the purpose, because organizations have effects beyond what the individuals in them intend. A moral theory that is exclusively about individuals and is irrelevant to organizations is irrelevant to business—as well as to most of contemporary life. A moral theory that does not work for institutions cannot work for people.

FACTS AND VALUES

Social science is generally concerned with explanation of what is. Moral philosophy is generally concerned with normative evaluations, with right and wrong or with bad and good. According to a dictum of Hume, one cannot infer an ought from an is (Hume [1739–40] 1978). That is to say, nothing moral follows from the mere facts of any matter. I cannot look at your act of killing your father and judge your parricide, merely from the observed facts, to be morally wrong. If I reach such a judgment, it is because I apply the normative principle that killing one's father is wrong or some other principle from which this conclusion follows, at least for the circumstances of your case. Every substantive normative judgment is, at least in part, an inference from a normative principle.

Then why is it worthwhile to bring social science and philosophy together on such issues? Philosophers bring one set of ideas, and they typically push for systematic treatment; social scientists bring other ideas, and they push for realism and relevance—and, often, common sense. Together, there is likely to be heightened discovery on both sides from the findings and arguments of the other side, as the philosophical commentators (Thomas Donaldson and Patricia Werhane) have noted. Moreover, the moral theory that is forced into engagement with realism and relevance is likely to be much better theory; social science is important for the fundamental arguments of moral theory. And the social science theory or explanation that is forced to be normatively systematic is likely to be a far better prescriptive and policy-relevant theory.

Consider an example of the effects of focusing on real problems for constitutional theory. Arguably the best body of constitutional thought to have come from a brief period is the debates over the U.S. Constitution that took place during its writing and adoption, 1787–1788. This debate was driven by the real and urgent need to find a better system of government, and the debate was carried on by practical men, not theorists. The central figure was
James Madison, especially in the larger part of *The Federalist*, although there were also valuable contributions from dozens of others. John Jay and Alexander Hamilton did some of their best writing for *The Federalist*, although their writing and ideas were inferior to Madison's. Madison was a statesman, but not a philosopher or legal theorist. John Adams, Thomas Jefferson, and perhaps James Wilson were major intellects in their own right, but while Wilson was a leading figure at the constitutional convention and in the politics of ratification in Pennsylvania, Adams and Jefferson were more nearly peripheral; indeed, Jefferson was out of the country and contributed relatively little to the constitutional debates. Though not a major intellect, Madison was very accomplished, very commonsensical, and remarkably open—and he had an unusually good instinctive strategic sense. These men achieved what they did largely because they were forced to work by the facts of the matter.

**ALTRUISM**

Much of the debate in social/psychological motivation is about the prevalence of altruism. Some psychologists and many economists provide clever arguments to show that apparently altruistic actions are really self-interested. Sociobiologists attempt to construct models for how altruism could be genetically selected. Real-world examples and experimental data are elicited to demonstrate the validity of the varied claims. Much of the discussion, however, is confusing in that it often runs together altruism that occurs at very different levels of interaction. Three levels are especially of interest: (1) individual-to-individual interactions over odd events, (2) frequent or large-scale interactions between individuals and other actors; and (3) systematic interactions that are often handled by institutions or by policy. These levels can be motivationally independent; an individual might be altruistic at one of the levels but not at another. They can also be causally distinct; the altruism that would work for problems at one level would not work for problems at another level. There is therefore not a single, simple altruism motivation.

Most of the discussion of altruism in philosophy and in sociobiology is focused on individual-to-individual problems. The traditional philosophers' example is the altruism of stepping into a pond to rescue someone—typically a small child—foundering there, at the mere cost of muddy or shrunk clothing that might ruin one's day. Such cases fit Hume's characterization of natural duty: We know that our own action produces enough good to justify our taking it (Hume [1739–40] 1978). Many one-on-one actions that might once have been called altruistic are now seen to be at least congruent with self-interest. For example, I behave generously toward you because I know you well and will benefit from continuing to know you well. If you were a complete stranger in a one-time encounter, I might be far less generous. As Hume noted, it is typically self-interested to keep one's promises; similarly it is congruent with self-interest to behave generously toward many people. A striking fact of general altruism, however, is that the weight of individual-to-individual problems in odd circumstances is relatively small in the balance of all problems. The problems with which individuals—or at least those who live in contemporary cities—most often have to contend are on the second level of interactions. And, arguably, the problems that have the greatest weight for accomplishing genuine good in the world are on the third level, institutional and policy problems.

Much of the discussion of altruism in institutional settings such as business seems to focus on problems on the second level and to treat them as matters for individuals to resolve. In my position in my organization, I can do good for a large number of people only at some cost to myself. For example, as an independent auditor of corporate practices, I can act as an individual to report management misbehavior. But if I do, I might not be rehired by the corporation the next time it is due for an audit, nor will I be fully trusted by other firms that might hire me. This is a standard example that seems likely to be a very commonplace problem. Insofar as it is commonplace, we should seek institutional devices to give incentives for generally good behavior. At one time, the norms of professionalism were thought adequate to control such conflicts of interest. Those norms were also thought to be the only available device. In many professional contexts the past norms are still pushed, but there are also institutional devices to change the self-interest calculation. In many contexts, self-interest may be the more reliable device for trumping self-interest.

Many of the people—both at professional meetings and in ordinary life—with whom I have discussed the middle category of frequent one-on-one interactions seem to be unsure how to deal with it. Some are unsure that, for example, giving to a street beggar improves the beggar's life. Hence, they are not sure that charity really is altruistic. Others are apparently unsure primarily because they think the larger problem is to improve many lives, which might be better if the system of street begging were not as remunerative as it is. Some of these people seem to feel that they often suffer from a moral variant of the logic of collective action. A gift to the beggar of the moment would make life better for that beggar but would make life for the whole class of potential beggars worse by giving them perverse incentives (R. Hardin 1990). Those incentives have led to the rise of what is perhaps a large class of very accomplished professional beggars who have mastered deceptive techniques. This moral variant of the logic of collective action is a peculiar but increasingly important social dilemma of a kind that should interest social psychologists in the social-justice perspective (Kramer and Messick, chapter 3).

As I have noted earlier, the three categories of altruism are motivationally independent. It was said of Bertrand Russell that he loved humanity but did not much care for individual humans. Someone with such views might hap-
pily vote for altruistic policies but balk at face-to-face altruism. Hume, to the contrary, said that he felt far greater sympathy for one individual's suffering that he could see than for the vastly greater suffering of the millions destroyed in an earthquake in China. He also thought that state-managed distributive justice would be pernicious in its effects (R. Hardin 1988). Someone with his views might vote against redistributive policies but act altruistically toward an individual in distress. In the case of the intermediate category of frequent individual interaction with a particular class of demands for altruism, people might react roughly as they would to the other class to which they tend to assimilate it, either the strictly occasional one-on-one encounter or the systematic interaction. Or, given the difficulty of understanding it causally, they might simply be unsure how to treat it.

The categories are also causally distinct. The chief reason many people support some variant of President Clinton's health-care reform is that they altruistically wish to have health care provided to the substantial fraction of Americans who have no coverage. Presumably, almost no one would suppose that individual-to-individual altruism could handle this problem. One's altruistic impulse here could be successfully put to work only through an institutional arrangement. John Darley (chapter 1) says, aptly, that harmdoing is a team sport; so too must altruism be if it is to be effective at all in many contexts.

Given the causal structures of various demands for altruism, it follows that one or another level of altruism will be more or less necessary if the relevant good is to be done. Without individual-to-individual altruism, individuals in idiosyncratic dire distress in acute moments will not be helped in time. Without enough altruism toward systematic problems to create institutions and establish government policies, various welfare goals cannot be met. These two categories of altruism are often necessary if their respective goods are to be achieved. That people are prepared to act altruistically in one arena, however, need not imply that they will act altruistically in the other arena. Mean-spirited citizens might be wonderful face-to-face, and wonderful citizens might be mean-spirited face-to-face.

Economists are sometimes thought to believe that there is no significant incidence of altruism. Moreover, they are thought to hold that if there were much altruism, it would be harmful; self-interest would produce better results on the whole than altruism would. This second claim is not as silly as it superficially seems to many critics. If most people most of the time were not being partial to their own and their families' interests, these would be badly served and we would all be worse off. Hence, altruism must motivate at most a fraction of our actions. But the skepticism about economists seems to go further than this, to the idea that they really oppose altruism. Indeed, they are thought to rely heavily on the market economists' bible, Adam Smith's Wealth of Nations ([1776] 1976), although Ayn Rand (1971) comes closer to holding the harsh view that self-interest should be the only motivator.

When Smith wrote in support of market organization of economic life and therefore in support of self-interest, he was opposing religious and group-based values that he thought burdened and even wrecked society. For example, towns and guilds excluded many laborers and severely restricted markets in order to protect their own from competition and to protect their local way of life. For Smith, self-interest spoke for egalitarianism and against especially favored interests. He was not fighting altruism, and his brilliantly quotable line about interests as motivating production is a claim for the possibility that interest can make things work well and beneficially. This was once a startling claim, because interest was formerly known as avarice, the greatest of all evils and the opposite of charity, the greatest of all virtues. In virtue theory, interest had been evil, while in Smith's view it was enormously beneficial to the larger society. Smith's claims for the beneficial superiority of interest still apply today to many values, such as the group-based values of racism and sexism and other in-group biases (Brewer, chapter 8; Rothbard, chapter 7; Tyler and Belliveau, chapter 9). Letting individual interest dominate these values would be widely beneficial, as it was also in Smith's time and has been ever since.

Then where must the Smith of The Wealth of Nations have stood on altruism? First, he cannot have opposed altruism of our first category, individual-to-individual, which is not related to his argument for the market economy. Second, he had the beginnings of an understanding of the logic of collective action, and he might have suffered our own ambivalence over altruism in the second category, individual action in the face of frequent demand. Or, given his views of the poor laws, one might suppose he would have opposed such altruism as not efficacious in at least some instances; this, of course, is a contingent matter of social science, not merely a matter of principle. Finally, whether he had an interpretable view of the third category of systematic altruism beyond the cases in which it would not be efficacious may be unclear. But he certainly favored public expenditures for broad public benefit and even for some relatively narrow benefits. These policies can, of course, be seen to serve mutual advantage rather than to be redistributive. Hence, we cannot unambiguously conclude where Smith stood on what is today the most important category of altruism.

There was, of course, an earlier Smith, of A Theory of Moral Sentiments. That Smith strongly argued for "universal benevolence," or altruism (A. Smith [1759]1976, 235). Possibly his views had changed in the nearly twenty years between these two works, so that by the time of The Wealth of Nations he held the views often attributed to economists. It seems implausible that he made a complete switch, although it seems likely that he made at least a partial switch. The later Smith was motivated to argue, as Bernard Mandeville ([1714]1924) had done earlier, that self-interest is a profoundly good motivation in the sense that it produces the best results we can expect. The struggle against botched values and their destructive intrusions may have been more
important to the later Smith than the concern with character and morality had been to the earlier Smith. Even more important than this, however, was the clearer understanding he seemed to have of the real weight of economic progress in making lives better. If I were to argue that their own altruism is what, say, Bangladeshis need, I would rightly earn Smith's dismissal and ridicule. Bangladeshis primarily need action from their own self-interest, without which they will remain in dreadful poverty. As the third category of problems for altruism swamps the other categories in our society, so Smith had evidently come to hold that the value of economic motivations swamped the value of normative motivations in his society. His view is arguably still true even in the most advanced societies.

When we study other motivations, moreover, we may find that, when they seem to be typically effective, they are often congruent with self-interest. This means that actions are overdetermined by supposedly contrary motivations. The hard task in analyzing motivations is to avoid simplistic claims of direct effects and to see how complex the relationships may be. When we do so, we may find that acting normatively well is simultaneously acting in our interests. This is what Hume ([1739–40]1778) showed in his account of promise keeping, with his implicit argument from the incentive structure of an iterated prisoner's dilemma. In parallel, we might find that actual instances of what we first think of as altruistic actions are actions of sensible generosity toward friends and associates, with whom we are in iterated prisoner's dilemma interactions. The altruistic act that is strictly against one's interest is substantially rarer, although surely common (Frank, chapter 4).

INTUITIONS

In moral debates, intuitions often play large roles. There are at least three kinds of intuitions that are of interest in moral theory: (1) factual intuitions, (2) formal intuitions, and (3) substantive moral intuitions. Factual intuitions include both direct and indirect or inferential intuitions. For example, I may have the direct intuition that the world is flat. And I may have the inferential intuition that if $x$ is true then $y$ is true, but I know $x$ is true, therefore I intuit that $y$ must be true. Aristotle is justifiably notorious for having faulty factual intuitions about trivially checked matters, such as his belief that heavier objects fall faster than lighter objects. For a distressingly long period of time, leaders of the Roman Catholic Church suffered the inferential intuition that if Aristotle said it, it is true.

Consider a couple of factual intuitions. At the conference at which the papers in this book were presented, Robert Bies asked for a show of hands on who would rather live in a world in which everything about everybody was known to everybody and who would rather live in a world in which nothing about anybody was known to anyone else. Many people had an instant intuition in favor of one or the other. Unfortunately, those two worlds would be so radically unlike the world we know that we probably cannot begin to think through the implications, even if we had weeks to work them out. Instant intuitions can be of no value in such a matter. This conclusion suggests that credible intuitions are often or perhaps even typically grounded in experience that makes them seem to be learned.

To cite a second example, I arrived at that conference immediately after being abroad for more than three weeks. Most of the people I knew at the conference I had previously seen more in Europe than in the United States. Since Evanston was not home for me, it seemed rather like a continuation of my prior sojourn in Europe, and throughout the conference I kept having the odd intuition that I was abroad. Or perhaps for me, after more than a decade at the University of Chicago, Evanston was something of a foreign country. In any case, whatever its origin, when the intuition intruded I was able to think through what I knew about where I was and realize that the intuition was wrong. Discovering that it was wrong once, however, did not block it from coming again. Each time it came, I had to rethink what I knew—for example, that after landing at O'Hare, I had got to Evanston in a taxi whose driver was from Chicago's South Side.

Formal intuitions are necessary in any field of inquiry. Mathematicians, logicians, linguists, scientists, musical creators, and many others have formal intuitions that underlie their arguments. Before the rise of utilitarian moral and political theory, virtually every moral theorist had the formal intuition that ethical principles are universal in their scope; if they apply to anyone, then they apply to everyone. Utilitarians further have the intuition that welfare is the general good of humanity (or of all sentient beings). Other theorists, such as John Rawls, have the intuition that fairness is what is generally right for humanity. Or they intuit that autonomy, agreement, human rights, or whatever is the basic principle of morality. For all these principles, there is no further argument. These are ground principles that are like the rules of arithmetic in that they must first be accepted before we can go further in our deductions.

The problem of formal intuitions did not arise in the discussions of the conference on business ethics. Perhaps the lack of comment was evidence that such intuitions are not a contested or difficult issue. They are, however, a difficult issue in philosophical treatments of morality. This issue arises as the central question of modern moral theory, perhaps of all moral theory. That question is, Why be moral? If we simply stipulate the ground principles of our theory as, say, Kantian or utilitarian, we then leave ourselves open to this question. To answer it, we would have to say why some ground principles must be accepted by a sensible person. Answers in recent centuries have
included evolutionary arguments, religious claims, and intuition pure and simple—as well as, of course, occasional arguments that beg the question. None of these is adequate in the sense of convincing most participants in the discussion.

At one time, logicians and mathematicians may have implicitly assumed that all their ground-level intuitions were valid and consistent. Now logicians and mathematicians are willing to stipulate radically different axioms from which quite different theoretical conclusions follow. Unfortunately, moral theorists have always been in the position of contemporary logicians and mathematicians. One theorist starts from axioms that another theorist dislikes, but neither theorist can give a compelling reason for accepting one or another set of fundamental principles.

The unique problem of moral theory, however, as compared with theories of other matters, is that moral theorists often have substantive moral intuitions—for example, to lie or break a promise is wrong, or failure to cooperate with others who are cooperating in some joint venture is wrong, or doing harm even to accomplish good is wrong. As Jonathan Baron says, they do psychology with an n of 1, since it is only their own intuitions that interest them. Such intuitions are the curse of modern moral theory. At earlier moments in philosophical history, it was supposed that we could trust such intuitions, because they must have been implanted by god. For many people today that is not a credible argument. In any case it cannot fit the facts of our intuitions, unless god is playing vicious tricks on us, because we too often radically disagree about what is intuitively true. To whitewash this dispiriting fact, intuitionists in moral theory often talk about the “common” morality, as though there were agreement when there is not. Ridiculously many people—philosophers and others—think they have firm, correct intuitions about substantive moral matters. Furthermore, they solipsistically think that their having these intuitions makes them right, or proves the intuitions are right, or at least gives them standing in moral debate and moral theory.

To some extent, substantive moral intuitions might seem similar to factual intuitions. But consider the odd factual intuition that I was in a foreign country in Evanston. I could demonstrate to myself that the intuition was in fact false. Suppose I have the intuition that it is always wrong to tell a lie, and I now face the dreadful problem put to Kant. An assassin appears at my door and asks whether you, his intended victim, are in my home. Alas, you are. To protect you, I can lie. Or to protect my moral purity, I can answer truthfully and then do my best to try to prevent your murder—but consequences may not matter much to me. Kant said that morality requires truth even in this grim instance. Any utilitarian would likely say it requires saving you, if possible—if necessary, by lying. Suppose I am a utilitarian (I have the formal intuition that welfare is the human good), but I seem to suffer from Kant’s odd intuition. I can readily see that these two intuitions conflict, but I have no way to show that one of them is inherently wrong. There are no facts of the matter to help settle the issue. I could correct my false factual intuition about being abroad in less than a minute (except when I had it upon awakening—then it took longer). I cannot settle my substantive moral intuition except at the level of pretheory. I simply reject the authority of such intuitions and accept the authority of my formal utilitarian intuition.

Some moral theorists have made the opposite choice and assert the authority of their ragbag of substantive moral intuitions over formal intuitions. The misfortunes of this move are, again, that these theorists do not agree with each other about which intuitions are morally correct and which morally wrong and that there is no way to ensure that the intuitions escape contradiction. Intuitionist ethics is therefore typically casuist and sophistic, requiring many words to rationalize away apparent contradictions. And it is largely impenetrable for systematic moral theorists, such as Kantians and utilitarians, who can never count on reaching the correct answers, either because they do not happen to have the intuitions or because when they do seem to have the intuitions, they too readily associate these with bits of their upbringing in a church, community, or family. If their intuitions are learned (perhaps by indoctrination), there is no compelling reason to hold them in principle to be valid. They are nothing more than a particular group’s social conventions. In some cases they are vicious and depraved.

It takes very little thought about substantive moral intuitions to begin to wonder how they come into our heads. But moral philosophers have spent most of their effort on analyzing and identifying the intuitions we supposedly have and very little on how we come to have them. Social psychologists may have special talents and methods for learning how we acquire them and, especially important, why they carry their sometime conviction. (People have knowingly died for holding to their intuitions.) Political psychologists have tended to suppose that certain values get inculcated early and then stay relatively fixed thereafter. That would be a dispiriting conclusion, precluding any hope of moral behavior, because, according to the literature of political psychology, among the values that get fixed are values of racism, sexism, and vicious variants of partiality of other kinds, such as familial, communal, and ethnic. Perhaps because children are often taught principles of behavior that would be beneficial for the larger family, however, the inculcated values may also tend to include values of cooperativeness, face-to-face honesty, and even modest generosity.

These two bundles of partly contradictory values may complicate the moral lives of most of us. For example, Gunnar Myrdal (1944) worried that the white children of the American South were taught contradictory values of equality and racial superiority that must eventually come into severe political and perhaps even psychological conflict. Myrdal was clearly not concerned that moral theory was inherently contradictory. Rather, he was noting that our moral education can be inconsistent. What we would like to know, in trying to respond to claims of substantive moral intuition, is how
they work psychologically and socially. It is conceivable that if we understood them better we would find them more credible or, alternatively, less credible. In this odd problem, such facts could help us to determine what moral force to give to such intuitions, because, as Kant said, *ought* implies *can*. If people cannot be having justifiable moral intuitions about specific substantive matters, then they cannot be morally bound by the claims of such intuitions.

**INSTITUTIONAL MORALITY**

So far, the discussion here has largely been about morality at the level of the individual, although, as I have noted, the bulk of our concern today may well be the morality of institutional and organizational actions, as in the third category of problems of altruism. This is surely the bulk of the problem of ethics in business. In an earlier age, the conditions of the market and the technology of the times made simple self-interest in mutually beneficial market transactions a sufficient condition for generally good and moral outcomes. Today, it can still be true that self-interest suffices only if organizations that dominate production and exchange can similarly control their participants with nothing more than self-interest to generate results that are profitable or otherwise desirable to the organizations—but even then, only if there are no deleterious external effects of organizational behaviors. In general, there typically are major external effects, such as pollution, skewed information and deception, social dislocation, discrimination, and so forth. And there are often opportunities for individuals to gain at the expense of the organizations for which they work and at the expense of the clients or customers of those organizations. We therefore face at least four possible devices for controlling individual behavior: (1) traditional Smithian self-interest, (2) individual moral principles, (3) legally backed enforcement, and (4) organizationally induced self-interest.

**Self-interest**

Traditional Smithian self-interest probably still successfully governs an enormous percentage of relevant behaviors, although we may have grown so accustomed to its working that we take little note of it. It is most obviously ineffective when individual incentives for gain conflict with organizational gain, as in various professional conflicts of interest. For example, consider recent accounts of the questionable practices of some mutual fund managers who invest to benefit themselves, their relatives, and their associates while risking their clients' funds (for example, *New York Times*, August 7, 1994). But it is also often ineffective when organizational interests conflict with broader social concerns.

**Individual Morality**

Much of the writing on organizations—and especially on business, and especially popular writing—focuses on inculcating relevant moral principles. It is probably for social psychologists to say whether such an effort can be expected to succeed, but it seems the least likely of the four devices to work. Even if we could inculcate various values, it is not obvious that doing so could resolve the problems organizations face. The crude theory that an aggregation of individual behaviors can produce organizational results will typically fail, for at least two general reasons.

First, it is a fallacy of composition to suppose that a cluster of people individually aiming at some organizational goal, such as reaching a particular moral outcome, will collectively achieve that goal. Rather, each may have to act in ways that are superficially unrelated to the organization’s goals. (It is a fallacy of composition to suppose without argument that the characteristics of a group are the same as those of its members. For example, although the members of a group may be rational, it does not follow that the group will be rational in the same sense.) Second, for many purposes we wish to control individuals’ behavior for stochastic, not directly individual, reasons. For example, it seems clear that we have far greater success in reducing road accidents by arresting those who test as drunk than by merely punishing those who are in fact personally responsible for accidents. On some moral theories, obtaining the result of reducing fatalities to innocent others is good reason to commend punishing a driver who drives drunk whether or not that person has actually caused harm.

A very important special case of attempting to inculcate moral principles to override the incentive of self-interest is the effort to get certain professionals to act according to a code of ethics. For example, accountants, auditors, and other agents of corporations may typically be in positions in which their interests and the interests of the corporation or of its owners or managers are in conflict. There is an enormous literature on such professionals, and the issue is too big to address here. It is plausible that this is the biggest unresolved area of conflict between individual interest and role requirements in all of public life. One way to resolve the conflict, at least in certain cases, is to change, say, the audit function to a form that would be much more a matter of public record and would make it easier to hold individuals liable for wrongdoing.

Incidentally, to suppose that good individual actions can simply be aggregated to produce good collective organizational results is to suppose that there is nothing distinctive about business, or professional, or government ethics, that they are all wholly the same in content as ordinary personal ethics. In that case, there would be little point in holding a conference on behavioral research and business ethics. All this might be a correct view, but one should probably suppose the contrary—that the structure of roles for various purposes entails distinctive principles for action in those roles.
Legal Enforcement

Legal sanctions have, of course, been a major device since ancient times, and they can be used in many business contexts, as in governmentally imposed conflict-of-interest standards for the securities industry. For three reasons, however, law is apt to be a clumsy device for organizational ethics in general. First, it is too expensive and cumbersome to bring into play for minor matters, which are common and collectively important. If organizations are to be efficient, therefore, they must find ways to handle most of their own moral problems. Second, law will typically lag the state of the art. For example, many of the most familiar problems of organizational behavior that have caused harms externally were problems that taught us what the law should be. Wrongs come first, and then law follows. Third, and arguably most important, law runs up against the problem of complex causation within organizations. Let us consider this problem in more detail.

The usual causal model for determining responsibility in the law works reasonably well for criminal law. That model is spelled out by H. L. A. Hart and A. M. Honore (1959); a similar, far more accessible account is presented by Ann McGill in chapter 12. In these accounts, the cause, for the purposes of assigning responsibility, is what seems anomalous in the series of actions and conditions that lead to a perverse result, such as a death or an organizational wrong. The model takes for granted that many actions and conditions likely contribute to any result of interest and that therefore we do not literally want a full causal account; we want only a somehow relevantly focused account. This model of causation will not work for moral theory, for the trivial reason that there is no moral agency analogous to the legal system to enforce it (R. Hardin 1988). Hence, in moral applications it lacks a value—decisive finality—that strongly commends it to legal use: In the law it tends to get definitive results despite the fact that its application is highly conventionalized. This is important in the law because we want order and finality in order to make our plans and investments. We want the legal system to settle things and let us move on. (Even better, of course, would be to have things settled in our favor.) The law tells us in advance what will determine culpability and thereby gives us incentive to act well in our own longer-run interest. A moral analog without a credible, stable enforcer cannot do this.

Despite its specifically legal peculiarities, the model of legal responsibility is a good starting point for a model of organizational responsibility. It sidesteps the problem of the fallacy of composition that undercuts simplistic causal accounts of such responsibility; if we are to get anywhere, we must get past this problem. Hence, we turn from the fallacious composition to a full causal model of our organization. Unfortunately, a full causal account of a significant organization's actions would require massive understanding far beyond what is possible for any individual agent. But if this is true, then the agent can hardly be held accountable for fitting all actions precisely to the organization's goals. Instead, the agent must fit actions to standard routines and even rules, which have been designed with the hope that they fit the desired result. It is this design that constitutes the fourth of the devices for controlling individual behavior in an organization.

Organizationally Induced Self-interest

This last device, organizational design, has been the least well studied of the four. Its rise in importance has paralleled the rise of relatively specific legal sanctions in the era of the modern corporation. Organizational design for achieving the central purposes of organizations is, of course, an old art that goes back several millennia in military, agricultural, navigational, and other activities. But the additional concern of achieving moral outcomes more or less by the way is a more recent concern. Typically, all organizational design takes the form of building in incentives for proper behavior. Hence, there need be little or no direct oversight or sanctioning. Rather, each agent of the organization simply faces a schedule of incentives that make the agent's self-interest and the organization's purpose congruent. Each individual agent acts as a Smithian producer for the market would.

In essence, morality must be designed into institutions and organizations (R. Hardin 1996). This is essentially what Willem Wagenaar in chapter 18 recommends for safety in the workplace. His example of the monkey man, who leaps about dangerously on an elevated platform, is, in an odd sense, structurally equivalent to some immoral behavior that is standard practice in the organization. Both are also structurally equivalent to organizational design to achieve any purpose, such as to produce its goods or services.

A well-designed organization might leave no room for anything other than specific individual responsibility for bad outcomes; then we could use McGill's causal account to establish culpability. Few organizations of substantial complexity and scale are likely to be so well designed, however. Investigations into bad behavior and bad outcomes should therefore focus sharply on organizational design. For example, while it is true that, in the moment when it really counts, it is the monkey man's own personal misstep that leads to his fall and likely death, it is also true that his risk of doing so was designed into the drilling procedure. Worker safety could be greatly enhanced by designing a new procedure that did not entail putting anyone at such risk. So too, a person's immorality on the job might be a risk to which the person is exposed by the organization's poor design for motivating action. True, a monkey man who is more agile and has quicker reactions might face less risk than another, and a saint might be less at risk of immorality on the job than a person of average fallibility. But in both cases, the organization would commonly be better served by finding procedures that did not call for exceptional people rather than trying to find exceptional people.
Note an implication of the argument from design. Suppose that our organization is to produce x and that this is a morally good purpose (according to our moral theory). If we design the organization to produce x in a particular way, with various roles and expectations for behavior in them, then it follows that in each of our roles it is moral for us to do what the role specifies. Might there ever be an exception to this conclusion? Yes. Suppose our background moral theory is utilitarianism and I can assess that, by violating some specification of what I should do in my role, I could cause a much better general result for the organization. That is to say, it is much better in the organization’s own terms. My organization might hold it correct for me to act in such a way and might even make that a part of my role. The organization would then be right to chastise me should I fail to act in this way, to go beyond my narrow role specification.

Indeed, even the law might be invoked in favor of going beyond narrowly defined roles. In tort law, the so-called reasonable-person rule is sometimes invoked. If a reasonable person would have taken some step to prevent a potential harm, then I should have done so. And if I did not, then I am tortiously liable (or my adversary is less liable than he or she might otherwise have been). That is to say, the court can decide, ex post facto, what would have been good or right behavior and hold someone accountable to this previously unstated criterion. In criminal law in the United States, laws cannot be applied ex post facto to actions taken before the laws were passed. My actions in a corporation similarly cannot be held criminally liable if there was no law already in effect to proscribe such action. But I and my corporation can be held financially liable for tort actions under the reasonable-person doctrine.

We might even go further and suppose that a corporation could be held liable under an analogous reasonable-corporation rule. It might then be up to the corporation to decide whether to hold me to account for being unreasonable in actions in my role in the corporation. The corporation’s own reasonable-person rule would stipulate that role holders should act with at least some discretion to disobey their role’s rules (Kadish and Kadish 1973). Some organizations might strongly discourage and punish such discretion; others might encourage it. Virtually all might expect more discretion to be exercised in some roles than in others. As though to parallel the law of torts, organizations can and do punish people for failing to do what is “reasonable” even to the point of disobeying their normal rules if necessary to be reasonable. Hence, they implicitly judge from a standard of the reasonable agent.

It is striking that the combination of a reasonable-corporation rule in the tort law and a reasonable-agent rule within the corporation would be roughly adequate to produce morally good action that is congruent with both agents’ self-interest and morality. Since it is the opposition of these two that is widely thought to drive most immoral behavior, corporations could achieve moral outcomes over a broad range. The central difficulty in organizational morality, therefore, is perhaps not the possible immorality toward others, as in actions with grossly harmful external effects, but the violation of organizational interests. Organizations typically must want to prevent such harmful actions but they fail to (Darley, chapter 1; Wagenaar, chapter 18). Exxon will have lost more from the Exxon Valdez oil spill than it will have made in oil transport overall, and the company has therefore sold off its tankers to eliminate the risk of future disastrous losses (Wagenaar, chapter 18). There might have been an incentive problem for Exxon in that case, but if so it was the failure of incorporating the new incentive into old routines for individual employees as tort responsibility for oil spills was expanded. Increasingly, the reasonable corporation will come around to being reasonable about redesign of processes and procedures to match them better with incentives.

We are left with the problem of immorality within organizations. It is at least conceivable that, while our moral theory would typically commend productivity or profit as an organization’s principal goal, it would also require fairness and other values in dealing with employees. And if both these principles are to apply, it is conceivable that on occasion they will conflict. Then we might expect to see profit override fairness. At this point, one might think we could resort to a variant rule of reasonableness about the trade-off between these two concerns. Unfortunately, if our theory does not give us a trade-off function, reasonableness will not either. In the reasonable-corporation and reasonable-agent rules, the standard of reasonableness is tested against actual results. In balancing two normative principles, there is no fact of the matter against which to test reasonableness. Anyone who invokes reasonableness in such a context (many moral philosophers do) is essentially resorting to substantive moral intuitions.

Such trade-off problems might be resolved in the law, which might stipulate how corporations are to handle such matters as fairness and discrimination. Lesser problems may not be suited to legal resolution and might give moral critics of organizations a great deal of difficulty. In such cases, we might readily expect organizations to limit their concern with these issues and to concentrate on what is legally regulated and on productivity. It will be hard for moral critics to say that this is a wrong policy.

In general, it seems likely that those matters that are regulated by legal mandates and restrictions and those that are regulated by the threat of tort actions together constitute the bulk of what matters morally in corporate behavior. Hence, it is clear that moral action of and within corporations can be made largely congruent with self-interest. If so, then Smith’s reliance on self-interest is still adequate for most of what matters in business today.

It is also clear that this conclusion would not follow if there were nothing equivalent to a strong tort regime to make it the interest of corporations not to harm others far beyond the benefits they bring them. One might suppose that during the Second World War Johns-Manville acted as badly as it did in
exposing workers to harm from asbestos in large part because its officers did not foresee the disastrous tort settlement that eventually bankrupted the firm. (Tobacco company executives today may have the same vision of their likely limited responsibility.) That is to say, they did not face an adequate incentive of self-interest to act well. In this, they differed from the Exxon executives and perhaps from the executives in the oil-drilling company with the badly designed monkey man role and, indeed, from executives in most corporations in legally advanced nations today.

There remain at least two major problems. First, today's executives might face no cost from tomorrow's tort action and might therefore severely discount it in deciding what is in their personal interest, as Johns-Manville executives may have done. Second, a corporation might organize with very little capital to be seized in a tort action while making great profits until being bankrupted by such tort action, as the small operators who have taken over Exxon's oil tankers perhaps do. Limited liability reduces the incentive of self-interest. Regulation might be necessary for such cases, although regulation is apt to be a clumsy device and it is likely to lag the problem even longer than tort action would.

INTEREST IN MORALITY

In sum, note again that the role account is essentially institutionalist and functional. It is not essentially utilitarian. We could work up a functional account of business ethics if we could first stipulate a general principle for which particular actions or classes of actions might be functionally relevant. For example, a Rawlsian might stipulate as the background normative principle some principle of fairness. Or a utilitarian would stipulate some principle of welfare. There might be other background principles, but these two may be the only systematic and coherent ones that cover most of what might be of interest. Such supposed general principles as rights and virtues seem to require grounding in manifold intuitions and are not systematic.

There are two conspicuous issues that the functional account of roles does not appear to address. First, perhaps it is true, as is sometimes supposed in the literature on professionalism, that many professionals inherently must be normatively driven, at least in part because they cannot be monitored. Even for such people, however, organizations and the law can be designed to bring moral action more nearly into congruence with self-interest to keep George Loewenstein's balance of self and others (chapter 11) from tipping too heavily toward self.

Second, for some values, such as certain environmental values (as Bazer- man and his colleagues describe in chapter 14), there may be no one whose interests are adversely affected who could be expected to enter a tort claim. Hence, these values plausibly cannot be secured with the combination of the reasonable corporation and its reasonable agents that are given incentive to make their actions congruent with the relevant values and their own interests. The environmental lawyer Christopher Stone (1972) has proposed that we give trees legal standing, so that any lawyer could defend a tree against assault and gain the lawyer’s fee in a successful suit. In such a legal regime, the reasonable corporation would have to take environmental values into account. Hence, environmental values would be brought into congruence with corporations' and their agents' interests. There are difficulties in this resolution, however. How is a court to decide the monetary value of a tree's suffering in order to assign a lawyer's fee? And should, for example, whales, the snail darter, the Grand Canyon, and rain forests have standing? And how do we handle the interests of future generations whose particular members would come into existence only because of long-past actions? Can those people reasonably hold someone responsible for harming them when they would not even have existed but for the ostensibly harmful action?

When the functional account of roles and the morality of their occupants can bring agents' actions into congruence with the relevant values and the agents' own interests, the chief problem organizations then face is bringing agents' actions into congruence with the organizations' interests by inducing the agents not to make rational blunders that lead them astray. The problems surveyed by Jonathan Baron (chapter 10), John Darley (chapter 1), Robyn Dawes (chapter 16), Helmut Jungermann (chapter 17), Josh Klayman (chapter 13), Willem Wagenaar (chapter 18), and others then should dominate our concern. Such problems may be especially important when they are compounded by the dynamics of changing technology and the rigidity of past practice, as Darley and others suggest. As a first cut, it is plausible that there is far more failure of rationality than of morality in organizations and that much of what seems to be immorality is rather irrationality. What organizations typically need is not so much to alter their agents' values but to be, as Patricia Werhane (commentary, part 2) suggests, more creative in thinking through and resolving problems they face and, as Robert Cialdini (chapter 2) suggests, inclined to make their control systems simple and clear enough for workers who are neither moral nor rational specialists.

Finally, in the discussion of the fact/value distinction there is often a somewhat contrary issue flitting in the wings. Kant supposed that ought implies can. That is to say, if you cannot possibly do action x, then it is not the case that you ought to do it. But this means that the fact/value distinction is slightly violated. The fact of the matter—that you cannot do x—immediately tells us something about the morality of the matter—that it cannot be supposed that you ought to do x. This is, incidentally, one point at which social science matters fundamentally for the content of moral theory. Being able to do something is at least a necessary condition for being morally obligated or duty-bound to do it. Virtually all moral theorists think this principle is true, although their views of the constraints that qualify for saying you cannot do
x may vary enormously. If we accept the principle, there is an odd implication for us. People who are better able to reason through the moral arguments for what it is right to do are bound to a higher standard than are those less able to reason them through. For example, if you are relevantly smarter, better educated, or more knowledgeable about our problems than I am, you can plausibly know to act better. But, by Kant's principle, if you can act better, then you should.

This conclusion might have an important structural implication for organizations. They can achieve better results if they have specialists to think about their problems, plan their results, and design solutions. At the same time, because they then can do better, they can reasonably be held to higher standards. What I do in a corporation that is struggling to survive and that has no resources to plan and second-guess its actions may be less constrained morally than what you do in a very self-aware, highly organized corporation. Courts might hold, however, that there is a correct standard for a reasonable corporation and might hold any corporation to that standard. Giving corporations incentive to be reasonable according to dynamically changing standards may be the best to be expected for the morality of business.

ENDNOTES

1. There may be other differences as well. For example, one might be provoked to altruism by the opportunity to relieve suffering, somehow defined, but not by the opportunity to make someone simply better off. More generally, one's altruism might be constrained by one's sense of justice.

2. For further discussion, see R. Hardin 1988, especially pp. 22-29, 178-191.

3. Incredibly, Kant argued that violating the formal principle against lying "is much worse than to commit an injustice to any individual" ([1797]1909, 365; see also R. Hardin 1989b, 78-79). If he would not permit lying for such a grave matter as saving someone's life, he presumably also would reject it when it had mere functional value, as in negotiation.

4. A sometime response to such problems is to argue from what must be only contingent facts that it is unlikely that one would have to violate a particular moral rule, such as, in this instance, not to lie. This is a pointless response to rationalize a rule that cannot be universally consistent with the relevant other general principle, which in this case is welfare. All that must be true is that some utilitarian in some such case might genuinely and honestly conclude that lying would most likely produce the best result.

5. Despite his statement of the fact/value distinction (perhaps, alas, as an afterthought), Hume ([1739-40]1978, 569) thought this a compelling argument in defense of the rightness of these intuitions.

6. I will carry through the argument from utilitarian assumptions. But every moral theory must be fitted to a similarly institutional structure if it is to be relevant to organizational behavior.

7. The reasonable-person rule began life as the reasonable-man rule, and it is severely criticized by many scholars for having enforced middle-class white male reasonableness on those other than middle-class white males. But the way such a rule has been used need not determine how it will be used hereafter. For an engaging discussion, see Calabresi 1985, especially chapter 2.