

The Rationality of Rule-Guided Behavior: A Statement of the Problem

SCOTT J. SHAPIRO*

It is often thought that rules are problematic because of their over- and underinclusiveness. The worry presumably is that it can never be rational to guide one's conduct by rules because any rule will give this wrong recommendation in some cases, and it can never be rational to follow the wrong recommendation.

To my mind, this isn't the real problem with rules, if only because the problem exists even in those instances when rules are not over- or underinclusive—that is, even when they give the right result. For even when they give the right result, we can still ask: Why do I have a reason to guide my actions by the rule rather than just doing the right thing? If it is the right thing to do, then that is the only reason I need, not because the rule told me to do the right thing.

Let me try to state the full problem in the abstract. Consider any rule and any action required by the rule. Either the action is supported by the balance of reasons or it is not. Assume that it is not. In such a case, it seems as if a rational person should not follow the rule because a rational person should never act against the balance of reasons. On the other hand, if the action is supported by the balance of reasons, then he should perform that act, not because the rule requires this, but because

* Professor of Law, Benjamin N. Cardozo School of Law, Yeshiva University. The above text is a lightly edited version of the talk I gave at the AALS Symposium on Rules. I have chosen to retain the informal nature of the discussion in the published version, but readers wishing a more formal and polished presentation of the issues raised herein may consult Scott J. Shapiro, *Authority*, in *THE OXFORD HANDBOOK OF JURISPRUDENCE AND THE PHILOSOPHY OF LAW* 382 (Jules Coleman & Scott Shapiro eds., 2002) and S. J. Shapiro, *Judicial Can't*, in 35 *NOÛS* 530 (2001).

rational agents should always act according to the balance of reasons.

It would seem therefore that a rational person has no need for rules: when rules give the wrong results, it is irrational to follow them; when they give the right results, it is irrelevant that the rule tells the person to act in that way.

This is a real problem for the law, because the law contains rules. It would seem that the law serves no purpose for a rational person.

Because the dilemma is perfectly general, any legal rule will fall prey to it. This is true even for rules that are not over- or underinclusive—it is true for standards as well. But let me illustrate the dilemma by using two examples: one slightly sensationalistic and the other more mundane.

At the beginning of Philip Howard's book, *The Death of Common Sense*, he tells the story of Mother Teresa who took over three abandoned buildings in the South Bronx and converted them to homeless shelters.¹ Her shelters would not be ordinary ones—they would be run by ascetic principles. No modern appliances would be used. The homeless would learn how to do everything the "old fashioned" way.

After most of the renovations had been done, her organization applied to the Housing Board for an occupancy certificate. The Board denied it on the grounds that the regulations require that any renovated building taller than three stories must have an elevator, and these buildings did not have one.

Mother Teresa explained that no one would use these elevators, as the shelter was to be run by ascetic principles. "Too bad," the Board said. The law makes no exceptions for ascetics.

I don't know if Mother Teresa sued, but had she filed suit, it seems clear that she would have failed. The Board's ruling would have been affirmed in court.

This is strange—the judge who affirmed the denial is making everyone worse off. Isn't this irrational?

Let's say, on the other hand, that the rules were different and made exceptions for homeless shelters. Would we want to say that the judge should allow the shelter to be renovated because the rules require this? Shouldn't we just say that the judge should allow the renovation just because it is the sensible outcome?

A more mundane example might be the Statute of Frauds. Suppose someone comes before a judge with a videotape and one hundred witnesses that saw him and someone else make a contract for the sale of hammers totaling \$500.

A judge would be obligated, under the UCC's Statute of Frauds, to dismiss this action. The oral contract is not enforceable because it is not in writing.

1. PHILIP K. HOWARD, *THE DEATH OF COMMON SENSE* 3–5 (1994).

But this seems unfair and therefore irrational for a moral person—the rule requires the judge to cause more harm than good by having him throw out a perfectly good bargained-for agreement.

On the other hand, if the contract was for \$499, then the judge would be allowed to enforce the contract. But would we say that he has reason to apply the law in this case? Genuine agreements should be enforced regardless of what the law says.

What is the ultimate problem here? My sense is that the problem has to do with a clash of perspectives. Rules are essentially backward-looking: rules give you reasons, when they give you reasons, because they were adopted in the past. Not only that, but they require you to ignore the future—you are supposed to follow what you did in the past, regardless of the future.

Practical rationality, on the other hand, is essentially forward-looking. You are supposed to ignore sunk costs and simply act so as to make the rest of your life better. So we get very different pieces of advice: rules require us to look to the past, whereas rationality requires us to look to the future.

As I said before, this is a completely general dilemma and it applies to all rules. But it is even more general than that. It applies to all commitment-like attitudes, such as, promises, intentions, policies, projects, and the like. So the dilemma is not just the concern of the legal philosopher—it is a problem for the decision theorist, philosopher of action, moral philosopher, political philosopher, and so on. It is also such a deep problem that legal philosophers are going to have to look elsewhere to find solutions, inspiration, analogies and the like. For this reason, I believe, Larry invited Michael Bratman, who is a philosopher of action, to talk about the same problem as it applies to personal rules. The hope is that enough people can talk to each other and compare notes so that we can crack this problem.

As I mentioned, in the philosophical literature, there have been numerous responses, but let me mention two. The first position, sometimes called particularism, concedes the force of the dilemma and claims that it is never rational to be guided by a rule. This is the position, for example, of the anarchists, who believe that rules and authority can never be legitimate. I think that this is Larry's position as well.

The second response tries to explain why a rational agent would follow a rule even when it gives the suboptimal result in an individual case.

This reply points out that sticking to a rule can have long-term benefits even though it has short-term costs. If a judge deviates from the law, this would have deleterious effects on future litigants.

What are the long-term benefits of sticking to rules? Let me mention one prominent one, namely, the benefit of deterrence.

As we all know, the law has rules, which are used to deter behavior that is socially undesirable. It threatens to punish acts that it wishes would not take place.

One example is the Statute of Frauds. The rule is there to prevent parties litigating evidentiary issues with regards to the existence of business contracts. If someone doesn't put his contract in writing, then regardless of its authenticity, the contract will not be enforced.

Deterrence is not a costless activity. Carrying out one's threat may harm innocent people or retard other socially beneficial goals. Consider the exclusionary rule. To prevent the constable from blundering in the future, we let the murderer go free now.

The argument is, however, that the benefits outweigh the costs. If we didn't pay the costs, we couldn't achieve the deterrence benefit. So sometimes we have to let genuine oral contracts go unenforced. That is the price of deterrence.

It would seem as if the dilemma is solved: rules are useful for rational agents because they can convince others that the rule-follower will act in certain ways, for example, by deterring them. An agent can be justified in following a rule when it gives a locally suboptimal result, because globally, it is maximizing to do so.

Rule-guided behavior, therefore, is not irrational—when we say that someone should follow a rule even when more harm than good comes from it, what we mean is that a rule should be followed when in the short term, more harm than good comes from following the rule, but in the long term it will be optimizing to do so.

This may sound commonsensical, but I think the argument as stated is perverse. To get at its perversity, let me rehearse a general problem about deterrence discussed by Thomas Schelling in *The Strategy of Conflict*.²

In his book, Schelling brought out the strange nature of threats. Threats are strange because they are expressions of intentions to act in the future in a costly manner, at a time that the action can no longer influence contact.

If the mugger says "Your money or your life," it is unclear why the victim would turn over his money. Why doesn't the victim reason as follows: "If he shoots me, then he can't get the money because he will have to run away and he will be up for murder if he does so. Hence, it

2. THOMAS C. SCHELLING, *THE STRATEGY OF CONFLICT* (1980).

would be irrational for him to carry out his threat.” And knowing that the victim will not treat his threat as credible, it is not clear why the mugger makes the threat in the first place.

Schelling offered the following solution: it may be rational for someone to try to convince another that he is irrational. The mugger, if rational, will want to convince the victim that he is irrational and hence the victim will succumb to his threats. And in the long run, it may be rational to cultivate and maintain a reputation for irrationality.

Now, when a judge follows through by applying the law in contexts of deterrence, that is, he is following through on the law’s threat, what is the judge trying to convince future litigants of? That he is irrational?

If at the present time the judge does more harm than good by following the law, won’t the same be true the next time the issue is raised in court? Won’t it be locally suboptimal to enforce the rule then too? What would make someone think that a judge would enforce the rule the next time around if that is the case?

Of course, if the judges were believed to be irrational or immoral, then it might be reasonable for others to rely on judges to apply the law. However, do we want to say that the Rule of Law rests on such beliefs and only on such beliefs?

I think the dilemma can be solved without saying that the Rule of Law is sustained by our thinking that the legal mind is irrational.

My strategy is to point out a hidden assumption of the dilemma. I call this the “autonomy assumption.”

It holds that people who have adopted rules are free not to follow them. Nevertheless, these people choose to follow the rules each time they deem the rules applicable. Each act of compliance implies choice.

If this is the case, then the dilemma is a knock-down argument—if a rule-follower has the ability not to follow his rules, then he shouldn’t do so when it gives the wrong results.

My suggestion, rather, is that rules operate as constraints on action—one who adopts a rule is constraining his future self to follow it. Adopting a rule is doing internally what Ulysses was able to do externally when he lashed himself to the mast. It is to forgo later choice by the operation of the Will, but it is as real as using some physical mechanism of commitment. The dilemma therefore dissolves: rules are not irrelevant because they constrain future action and it may be rational to prevent oneself from maximizing later on.

Obviously, I cannot defend my position here, but let me just clarify one thing that Larry mentioned earlier. He said that I believe that when people follow rules, they don't do so for reasons. This isn't right. To be guided by a rule requires that someone treat the rule as a reason and to act for that reason. So rule-guided behavior is intentional action—I am simply denying that it is free action. The rule constrains nonconformity, but it does not compel the intentional action. The agent still has to act on the rule and to treat it as a reason. Indeed, according to the Constraint Model, he does have such a reason, namely, because the constraint renders conformity the optimal option.