THINKING LIKE A LAWYER

A NEW INTRODUCTION TO LEGAL REASONING

Frederick Schauer

HARVARD UNIVERSITY PRESS

Cambridge, Massachusetts

London, England

2009

Library of Congress Cataloging-in-Publication Data

Schauer, Frederick F., 1946–
Thinking like a lawyer : a new introduction to legal reasoning / Frederick Schauer.
p. cm.
Includes bibliographical references and index.
ISBN 0-674-03270-5 (alk. paper)
1. Law—Methodology. I. Title.
K212.S325 2009
340'.1—dc22 2008035011

for Bobbie

Of course, we do not live in Plato's utopia,[17] and thus we understand that the values of legal reasoning and the Rule of Law may serve important goals in constraining the actions of leaders lacking the benign wisdom of Plato's hypothetical philosopher-kings. But even when we leave Plato's utopia and find ourselves in the real world with real leaders and their real flaws, the same dilemma persists. Legal reasoning in particular and the Rule of Law in general will often serve as an impediment to wise policies and to the sound discretion of enlightened, even if not perfect, leaders.[18] When and where the Rule of Law might turn out to serve the wrong interests, or simply to be so concerned with preventing abuses of individual discretion that it impedes sound discretion, is not the focus of this book. Evaluating law and assessing the Rule of Law is the work of a lifetime, and indeed not just the lifetime of any one person. The far more modest goal of this book, therefore, is to identify, describe, analyze, and at times evaluate the characteristic modes of legal reasoning. Determining, in the aggregate, whether and when those modes are worth having, a question whose answer is far from self-evident, is best left for other occasions.

17. Nor did Plato, as he well recognized.

18. See Morton J. Horwitz, "The Rule of Law: An Unqualified Human Good?," 86 *Yale L.J.* 561 (1977) (book review).

2

RULES—IN LAW AND ELSEWHERE

2.1 Of Rules in General

Reasoning with *rules* is perhaps the most common image of what lawyers and judges do. A widespread popular conception has it that lawyers argue their cases by appealing to abstruse rules not understandable by ordinary people, and that judges make their decisions by consulting books full of such rules. Having found the right rule, so it is thought, the judge proceeds to apply it mechanically to the case at hand, and that is the end of the matter.

Legal sophisticates commonly mock this image, which strikes insiders to the law as being far removed from the realities of actual practice. And for a host of reasons it is, not least being that most controversies or events involving a straightforward application of existing rules will not wind up in court at all.[1] But for all the inaccuracies and exaggerations built into this ubiquitous caricature of what lawyers and judges do, it nevertheless captures a genuinely important part of law. Rules actually do occupy a large part of law and legal reasoning. Lawyers frequently consult them, and judges often make decisions by following them. Law may not be all about rules, but it is certainly a lot about rules, from the

1. Because straightforward or easy applications of legal rules are rarely litigated, the cases that come to a court are predominantly and disproportionately ones that are in some way hard. The litigated hard cases thus represent a biased sample of all legal events, a phenomenon typically referred to as the *selection effect*. See George L. Priest & William Klein, "The Selection of Disputes for Litigation," 13 *J. Legal Stud.* 1 (1984). We will take up the selection effect and its consequences in the next section of this chapter, and return to it in Chapters 7 and 8.

Rule Against Perpetuities in property to the "mailbox rule" in contracts to the felony murder rule in criminal law to the Federal Rules of Civil Procedure to innumerable others. And because rules loom so large in what law does and how it does it, figuring out in a noncaricatured way what rules are and how they work will take us some distance toward understanding legal reasoning, legal argument, and legal decision-making.

Consider the typical speed limit, which is a rather uncomplicated example of a rule. The sign says SPEED LIMIT 55, and our first reaction is that the speed limit is 55 miles per hour.[2] But why 55? Presumably the speed limit was set at 55 because someone in authority—possibly the legislature, but more likely the highway department, the county commissioners, or the state police—believed that driving faster than 55 on this road would be unsafe. All well and good, and probably right for most circumstances, but an important feature of the speed limit sign is that it is there *all* the time. And equally important is that the speed limit applies to virtually everyone.[3] The speed limit is 55 when it is raining and 55 when it is clear. It is 55 when there is heavy traffic and 55 when there is none. It is 55 for cars designed to go up to120 and 55 for cars that start to shake at 50. And although 55 is the speed limit for safe drivers, it is also the speed limit for the reckless and the inexperienced. The speed limit of 55 is designed to achieve safety, but in some circumstances 55 might be too high to achieve that goal, and in others it might be unnecessarily low.

So suppose that you are out driving your new and carefully maintained car one clear, dry, traffic-free Sunday morning. And suppose that you are an experienced and cautious driver. Indeed, you have never been in an accident and have never been cited for a moving traffic violation. Because you are a good driver and because the conditions are ideal,

2. Some people will respond to this example by pointing out that when the posted speed limit is 55, the "real" speed limit is somewhat higher. For many drivers, perhaps even most drivers, SPEED LIMIT 55 means you should not drive over 64, because they know that typically the police will not stop you unless you have exceeded the speed limit by at least 10 miles per hour. The discussion of Legal Realism in Chapter 7 will address this issue, examining more carefully the implications of the fact that official practice often diverges from the literal meaning of a written rule. This divergence raises important and complex questions, but the typical speed limit is more straightforward. Most drivers know what the acceptable leeway is, and almost all of the point of the example in the text is preserved even when there is a widespread knowledge that the actual speed limit is the posted speed limit plus 9 miles per hour.

3. We need not worry for the moment about fire engines, ambulances, and police cars.

you decide to drive—perfectly safely—at 70. Having made that decision, however, you look in your rearview mirror and are disturbed to see the flashing lights of a police car signaling you to pull over. The next thing you know, the police officer is informing you that you have been clocked at 70 miles an hour in a 55-miles-per-hour zone. "I know," you say to the officer, "but let me explain. The fifty-five-miles-per-hour limit is designed to ensure safety, but actually I am driving very safely. There is no traffic. The weather is clear. The highway is dry. My car is in good condition. And I have a perfect driving record. You can check. You and I know that fifty-five is just an average for all drivers and all conditions, but the real goal of the speed limit is to make sure that people drive safely, and you can't deny that I was driving very safely."

We all know what would happen next. The officer would point to the speed limit sign, if one were visible, and then say something like, "The speed limit on this road is fifty-five. Fifty-five means fifty-five, not what *you* think is safe driving." And that would be the end of it. You would receive a speeding ticket, and you would get that ticket even though the goal of the speed limit rule was to make people drive safely, and, most importantly, you would get the ticket even though you *were* driving safely.

This example may seem trivial, even silly, yet it illustrates a larger and central point about the very idea of a rule. Every rule has a background justification—sometimes called a rationale—which is the goal that the rule is designed to serve.[4] Just as the typical speed limit is designed to promote safety on the highways, so the goal of the Rule Against Perpetuities is to limit to a plausible time the period of uncertainty in the possession and disposition of property. The goal of Rule 56 of the Federal Rules of Civil Procedure—the summary judgment rule—is to eliminate before trial those cases in which there is no legally serious and factually supportable claim. The goal of the parol evidence rule is to effectuate the intention of parties to reduce their agreement to writing. And so on. Every rule has a rationale or background justification of this variety, and thus every rule can be seen as an attempt to further its background justification.

In theory, it would often be possible for the rule simply to be a restatement of the background justification. A few years ago, for example,

4. For a lengthier discussion, see Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (1991). See also Larry Alexander & Emily Sherwin, *The Rule of Rules: Morality, Rules, and the Dilemmas of Law* (2001).

the state of Montana eliminated all fixed speed limits, requiring instead only that driving should be "reasonable and prudent."[5] But drivers have widely divergent ideas of what is reasonable and what is prudent, and so do police officers and judges. As a result, there developed wide variations in speed limit enforcement, the consequence being that drivers became highly uncertain about just how fast they could go without running afoul of the law. This much uncertainty was too much for the Montana Supreme Court, which struck down the "reasonable and prudent" rule as excessively vague. Indeed, even had the rule not been declared unconstitutional under the Montana Constitution, it was likely that the legislature would itself have reinstituted numerical speed limits and eliminated the "reasonable and prudent" rule. In Montana, as elsewhere, people understand that the background justifications themselves are often too vague to be helpful, too fuzzy to give people the kind of guidance they expect from the law, and too subject to manipulation and varying interpretation to constrain the actions of those who exercise power. So although in theory a speed limit rule could simply restate these abstract rationales—Drive Safely, or Drive Prudently, or Drive with Care—in practice the abstract rationales or background justifications are typically reduced to concrete rules. These concrete rules are designed to serve the background justifications, but it is the rule itself that carries the force of law, and it is the rule itself that ordinarily dictates the legal outcome. That is why the safe driver gets a ticket when she is driving safely at 70 miles per hour, and this example is just one of many that illustrate the way in which it is the concrete manifestation of a rule and not the abstract justification lying behind it that normally represents what the law requires.

Consider, to take another example, the somewhat technical rule (Rule 16(b) of the Securities Exchange Act of 1934) in the American law of securities regulation that prohibits certain corporate insiders from buying and then selling (or selling and then buying) shares in their own company within a period of six months or less.[6] Lying behind this rule is the goal—the rationale—of preventing corporate insiders, who are presumed to have inside information typically unavailable to the public and unknown by those with whom insiders might trade, from trading on that inside information. But the rule itself says nothing about the actual possession of

5. Mont. Code Ann. 61–8–303 (1996), invalidated on grounds of excessive vagueness in State v. Stanko, 974 P.2d 1132 (1998). See Robert E. King & Cass R. Sunstein, "Doing Without Speed Limits," 79 *B.U. L. Rev.* 155 (1999).

6. 5 U.S.C. §78p(b) (2000).

inside information, and instead simply prohibits any officer, director, or holder of 10 percent or more of a company's shares from buying and then selling or selling and then buying the company's shares within a six-month period. The thinking that produced such a specific rule was that by prohibiting people from engaging in so-called short-swing transactions, the rule makes it that much more difficult for insiders to profit from the knowledge they have gained just because of their position as insiders. The rule does its work, therefore, by prohibiting short-swing transactions regardless of whether the person engaging in the transaction actually has insider knowledge, just as the speed limit rule prohibits driving at a speed in excess of the limit regardless of whether the driver is in fact driving unsafely. The short-swing purchaser or seller who qualifies as an insider under the highly precise definition of an insider has violated the rule and is required to "disgorge" his profits even if he has no inside information whatsoever. And although a person who trades on inside information without being an insider as defined by this rule may well find himself in trouble under some other rule,[7] it is noteworthy that he is not liable under *this* rule, just as the person driving unsafely but below the speed limit has not violated the speed limit rule.

Still another example comes from the laws in many jurisdictions prohibiting the possession of burglar tools.[8] The law does not *really* care about burglar tools—it cares about burglaries and about limiting their frequency. But although the rule serves the background justification of preventing burglaries, it puts that background justification into effect by prescribing something more specific. The rule prohibits possessing burglar tools rather than just anything that might increase the risk of burglary, just as the typical speed limit is an explicit numerical rule and not a mandate that everyone drive safely or prudently, and just as the short-swing transaction rule prohibits all transactions by defined insiders in a defined period of time and not all or only those transactions in which a person trades on inside information.

The lesson to be drawn from these examples is that one of the principal features of rules—and the feature that makes them rules—is that

7. In particular, Rule 10b-5 promulgated by the Securities and Exchange Commission, 17 C.F.R. §240.10b-5 (2007), which, among other things, makes it unlawful in many securities transactions to omit to state a material fact to other participants in the transaction.

8. E.g., Conn. Gen. Stat. Ann. 53A-106 (West, 1999); Cal. Penal Code 466 (West, 1999).

what the rule *says* really matters. That is why the police officer will give you a ticket if you are driving above the speed limit even when you are driving carefully and safely, and that is why corporate officials and major shareholders are liable for damages if they trade in their company's stock within a six-month period even when they have no inside information at all. Recall from Chapter 1 the discussion of *United States v. Locke*,[9] in which the Supreme Court enforced the "prior to December 31" Bureau of Land Management filing rule even though it was obvious that what Congress really meant to say was something like "on or prior to December 31." The decision in *Locke* seems to some commentators mistakenly to take the importance of the actual language of the rule to absurd extremes,[10] and perhaps it does, but the fact that six Supreme Court Justices were willing to enforce to the letter the literal language of the "prior to December 31" rule demonstrates the way in which a big part of a rule's "ruleness" is tied up with the language in which a rule is written. Central to what rules are and how they function is that what the rule *says* is the crucial factor, even if what the rule says seems wrong or inconsistent with the background justifications lying behind the rule, and even if following what the rule says produces a bad result on some particular occasion. When we take up statutory interpretation in Chapter 8, we will delve more deeply into these issues, including considering the circumstances under which what a statute literally says is not the last word in interpreting its meaning and application. But even when what a rule says is not the last word, it is almost always the first word, and understanding what rules are and how they work entails understanding that the rule, as written, is important in itself, rather than being merely a transparent window into the rule's background justification.

2.2 The Core and the Fringe

Although a large part of how rules work is a function of what the words of a rule say, it is often difficult for lawyers and judges, and even more for

9. 471 U.S. 84 (1985).

10. Richard A. Posner, "Legal Formalism, Legal Realism, and the Interpretation of Statutes and the Constitution," 37 *Case West. Res. L. Rev.* 179 (1986); Nicholas S. Zeppos, "Legislative History and the Interpretation of Statutes: Toward a Fact-Finding Model of Statutory Interpretation," 76 *Va. L. Rev.* 1295, 1314–16 (1990).

law students, to appreciate this feature of rules. This is because so much of what judges, lawyers, and law students do takes place at the edges of rules rather than at their centers. The English legal philosopher H. L. A. Hart famously drew a distinction between the clear center (he called it the "core") of a rule and its debatable edges (which Hart labeled the "penumbra"), and in the process offered a hypothetical example that has become legendary.[11] In his example, Hart asked us to imagine a rule prohibiting "vehicles" from a public park. This rule, Hart observed, would plainly prohibit automobiles, because automobiles clearly count as vehicles according to the widely accepted meaning of the word "vehicle." And Hart would undoubtedly have reached the same conclusion with respect to trucks, buses, and motorcycles, all of them being core examples of "vehicles" as well. But what, Hart asked, if we were considering whether bicycles, roller skates, or toy automobiles were also prohibited by the "no vehicles in park" rule? And what, he might have asked, about baby carriages? And, these days, what about skateboards or motorized wheelchairs? Now we are not so sure. We are no longer at the core of the rule, where things appeared pretty straightforward. Instead we have moved out to the fuzzy edge or penumbra of the rule, where we might be required to look to the purpose behind the rule to see whether some particular fringe application should be included or not. If the rule's background justification had been to promote safety for pedestrians, for example, then perhaps baby carriages but not bicycles or roller skates would be allowed in the park. But if instead the rule had been aimed at keeping down the noise level, then maybe there would be no reason to exclude bicycles, roller skates, or baby carriages, although there might be good grounds for wanting to keep out gas- or electric-powered toy cars.[12]

11. H. L. A. Hart, *The Concept of Law* 125–26 (Joseph Raz & Penelope Bulloch eds., 2d ed., 1994). The example first appeared in H. L. A. Hart, "Positivism and the Separation of Law and Morals," 71 *Harv. L. Rev.* 593, 608–15 (1958). For an extended analysis, see Frederick Schauer, "A Critical Guide to Vehicles in the Park," 83 *N.Y.U. L. Rev.* 1109 (2008). We will return to the example in Chapter 8 when taking up the subject of statutory interpretation.

12. In a memorable debate in the pages of the *Harvard Law Review*, Lon Fuller, Hart's American contemporary, challenged the idea that the plain meaning of words alone could *ever* produce a clear outcome without consultation of the purpose lying behind the rule. Lon L. Fuller, "Positivism and Fidelity to Law—A Reply to Professor Hart," 71 *Harv. L. Rev.* 630 (1958), *replying to* H. L. A. Hart, "Positivism and the Separation of Law and Morals," 71 *Harv. L. Rev.* 593 (1958). The

That rules have debatable fringes where there are good arguments on both sides of the question whether the rule apples or not is hardly news to lawyers. Indeed, such disputes are a large part of the lawyer's stock in trade. But the clear and undebatable core of a rule is often neglected by lawyers and law students because plain or easy applications of rules so rarely get to appellate courts. For that matter, they rarely get to court at all, or even to lawyers. If the driver of a pickup truck with family in tow and picnic aboard arrived at the park and observed the NO VEHICLES sign, we would expect him in the normal case simply to turn around and drive somewhere else, producing no controversy at all. Similarly, although there might be difficult and contested questions at the edges of even a rule specifying a precise time limit, in the ordinary course of things a defendant in federal court will answer a complaint or request an extension prior to the expiration of the twenty-day period specified in Rule 12(a)(1)(A)(i) of the Federal Rules of Civil Procedure. These straightforward applications of legal rules rarely appear in casebooks or law school classes, and as a result much that is important about legal rules tends to operate invisibly to law students, invisibly to lawyers, and especially invisibly to judges.[13]

The distinction between the clear core and the fuzzy edge of a rule can be illustrated by way of a real case decided by the United States Supreme Court a few years ago, a case intriguingly similar to Hart's hypothetical example of the vehicles in the park. In *Stewart v. Dutra Barge Company*,[14] the question before the Court was whether a large dredge called a Super Scoop was a "vessel" as that word is used in federal maritime law. In fact, the Super Scoop was the largest dredge in the world at the time, and was being used to excavate Boston Harbor as part of the project known as the Big Dig. Willard Stewart, a worker on the Super Scoop, was injured while on the job, and he sued the owners of the dredge, claiming that the company's negligence was the cause of his injuries. It turned out, however, that whether Stewart could bring such a suit depended on whether the Super Scoop was a "vessel." If it was, then a federal statute called the Jones Act[15] would allow and provide the basis

relationship between text and purpose is important and will be among the central themes we deal with in Chapter 8.

13. See Frederick Schauer, "Easy Cases," 58 *S. Cal. L. Rev.* 636 (1985).

14. 543 U.S. 481 (2005).

15. 46 U.S.C. App. §688(a)(2000).

for the suit. But if the Super Scoop was not a vessel, then another federal statute—the Longshore and Harbor Workers' Compensation Act[16]—would allow people like Stewart to claim the equivalent of workers' compensation payments but would preclude a suit against the company for negligence. So whether Stewart had a right to bring an action for negligence against the barge company turned on whether the Super Scoop was a vessel.

This was a hard case. Although the Super Scoop spent most of its time in a stationary position while dredging out the channel, and although its almost total lack of capacity for self-propulsion required that it be towed from one location to another, it did have a captain and crew, and it did float, both while dredging and while being moved from place to place. Consequently, Stewart made the plausible argument that the Super Scoop's normal floating position, combined with its captain and crew, made it a vessel, while the barge company offered the equally plausible argument that the Super Scoop's lack of self-propulsion and resemblance in appearance and function to a piece of stationary land-based construction equipment made it something other than a vessel. At the end of the day, the Supreme Court decided unanimously that the Super Scoop was indeed a vessel, but the actual outcome need not detain us. What is important here is that although the case before the Supreme Court was a hard one in which there were nonfrivolous arguments on both sides, the Supreme Court case is likely to paint a false picture of the routine and unlitigated operation of this particular set of legal rules. Unlike the question of the Super Scoop in Willard Stewart's lawsuit, most of the questions—virtually all of the questions, for that matter—about whether something is or is not a vessel would almost certainly never reach the Supreme Court, would probably not get to an appellate court, and likely would not even have been litigated. If the question had been whether a thousand-passenger cruise ship was a vessel, there would be no serious argument that it was not, and no competent lawyer would argue otherwise. There might be other good arguments available in the overall dispute, but it is unlikely that a court would be called on to adjudicate the question of whether the ocean liner was a vessel. The rule would be applied, but it would never see the inside of a courtroom. Similarly, if the edge of a harbor were being dug out by a land-based excavating machine that did not and could not enter the water, the machine's status as some-

16. 33 U.S.C. §902 (2000).

thing other than a vessel would in all likelihood not have been challenged and again would not have come before a court at all.

Because genuinely easy cases and straightforward applications of legal rules are so rarely disputed in court, the array of disputes that *do* wind up in court represents a skewed sample of legal events. The effect, known as the selection effect,[17] is such that the cases that wind up in court are only—or almost only—the ones in which two opposing parties holding mutually exclusive views about some legal question *both* believe they have a reasonable chance of winning. If one of the parties thought it had no reasonable likelihood of prevailing in the case as a whole, or even just prevailing on this particular issue, it would not, barring unusual circumstances,[18] contest the matter at all. It would follow the law, or pay the claim, or settle the case, or rely on some other argument in litigation. The cases and arguments that are seriously contested in court, therefore, are the ones in which both parties think they might win, and this situation typically occurs only when they both have plausible legal arguments. With respect to legal rules, therefore, both parties will reasonably think that they might win when, ordinarily, the relevant question lies at the edges and not at the core of the pertinent rule. And thus the selection effect is so called because the incentives of the legal system create a world in which only certain applications of law or rules are selected for litigation, and the ones selected have the special characteristic of being at the fringes of legal rules, or in some broader way at the edges of the law.

17. There is a large literature on the selection effect in law, but the seminal article is George L. Priest & William Klein, "The Selection of Disputes for Litigation," *supra* note 1. See also Richard A. Posner, *Economic Analysis of Law* §21 (3d ed., 1986); Frederick Schauer, "Judging in a Corner of the Law," 61 *S. Cal. L. Rev.* 1717 (1988). An excellent overview of the issues and the literature is Leandra Lederman, "Which Cases Go to Trial?: An Empirical Study of Predictions of Failure to Settle," 49 *Case West. Res. L. Rev.* 315 (1999). And it is worth noting Karl Llewellyn's much earlier observation that litigated cases bear the same relationship to the underlying pool of disputes "as does homicidal mania or sleeping sickness, to our normal life." Karl N. Llewellyn, *The Bramble Bush: On Our Law and Its Study* 58 (1930).

18. One such unusual circumstance occurs when a party litigates or threatens litigation even when it knows the law is contrary, simply for the purpose of wearing down an adversary by delay or expense. In theory the legal system has devices to prevent this—summary judgment, for example—but in practice parties do pursue losing causes for strategic reasons more than the pure theory of the selection effect would predict.

The selection effect is the major factor in determining which disputes or law-controlled events wind up in litigation, but the effect is even greater as we proceed up the appellate ladder. In its 2007 Term,[19] for example, the United States Supreme Court, which has almost total power to decide which cases it wants to hear, was asked to hear more than nine thousand cases from the federal courts of appeals and from the highest courts of the states, but agreed to take and decide, with full briefing and argument and opinions, only seventy-one.[20] These seventy-one cases were almost all ones in which there was no clear legal answer, and taking these seventy-one as representative of how law works or how rules work would be a major blunder.

Very much the same dynamic applies to the cases selected for law school casebooks. What makes those cases interesting and pedagogically valuable is, usually, that they are hard cases, ones in which the lawyers on both sides can make strong arguments and in which the students can analyze and evaluate the opposing positions. And because these are hard cases, the opinions of the deciding courts can almost always be questioned, which is a big part of what case-based law classes do. In itself, there is nothing wrong with this. Learning how to make good arguments on both sides is part of becoming a lawyer, and so is learning how to expose the weaknesses in a judicial opinion. But it is nevertheless an error to suppose that all or even most cases are hard, that most legal events are disputable, and that legal rules never or rarely give clear answers. Appellate courts and law school classrooms have good reasons for operating in the gray areas of rules—on the fuzzy edges. But it is a big mistake to assume that rules are nothing but gray areas and fuzzy edges.

19. The Supreme Court hears and decides cases in what is called a Term, traditionally starting on the first Monday of October and ending when the Court finishes deciding the cases it has heard, typically in June. The Term is designated by the year in which it starts, so sometimes it is called the October 2007 Term, for example, and sometimes just the 2007 Term.

20. The exact count for the 2006 Term, the most recent Term for which exact statistics were available at the time of publication, is that the Court received 8922 appeals or petitions for review, decided 278 of those by summary order without opinion, and agreed to hear and decide 77, of which 73 wound up actually being decided, after briefing and oral argument, with full opinions. "The Supreme Court, 2006 Term: The Statistics," 121 *Harv. L. Rev.* 436 (2007).

2.3 The Generality of Rules

Although the application of legal rules to the world is characterized by easy cases, adjudication is dominated, for the reasons just examined, by the hard ones. These hard cases, however, come in several varieties. One is the case at the fuzzy edge of a rule, of which *Stewart v. Dutra Barge Company* is a typical example. A very different type of hard case, however, resembles the speed limit scenario more than it does the Super Scoop case. When you are pleading to the police officer that you were not actually driving unsafely, you are not claiming that the rule is unclear in this application, as you might be if you were stopped for not having your lights on after dark if it were dusk or if you were stopped in Montana during the regime of the "reasonable and prudent" speed limit. Rather, the typical attempt by a driver to talk her way out of a ticket involves acknowledging that the rule according to its literal terms plainly applies to her—she really was going 70 in a 55-miles-per-hour zone—but she nevertheless claims that literal application of the rule's terms to *this* case would not serve the background justification lying behind the rule. She admits she was going more than 55, but she certainly wasn't driving unsafely. Or so she says.

Such conflicts between the outcome that the words of a rule indicate and the outcome indicated by the rationale behind the rule are ubiquitous. For example, the Seventh Amendment to the Constitution provides the right to a jury trial in any civil case at common law in a federal court in which the amount in controversy is "twenty dollars," and it is obvious that the purpose behind the twenty-dollar minimum was to limit jury trials to cases in which substantial sums were involved. But although twenty dollars was a substantial amount of money in 1791, when the Seventh Amendment was adopted, it is hardly substantial anymore. Much the same can be said about the requirement in Article II of the Constitution that the president have attained the "Age of thirty-five Years," a requirement created when the life expectancy at birth for a male (almost no one at the time contemplated that women could even vote, let alone be president) was under forty, as compared to the current average life expectancy at birth for American men and women combined of over seventy-five.[21] But as with the effect of inflation on the twenty-dollar threshold

21. Many children died from disease in the eighteenth century, so the raw figures can be a bit misleading, because most adult males did live into their fifties and sixties. But even for those who reached adulthood, the differences between 1787 and now are still substantial.

for jury trials, the fact that the literal meaning of the "Age of thirty-five Years" rule fails to serve the rule's background purpose does not change the meaning of the rule itself, a meaning that remains tethered to the meaning of the words in which the rule is written. If you are only thirty-two years old, you cannot be president, and it verges on the fantastic to imagine circumstances in which that would not be true, regardless of the underlying rationale for the rule.[22] So too with the more controversial requirement, also in Article II, that the president be a "natural born citizen." This rule, which has precluded secretaries of state such as Madeline Albright and Henry Kissinger and governors such as Arnold Schwarzenegger and Jennifer Granholm from seriously contemplating running for president, is almost certainly a poor embodiment of the original background justification of ensuring loyalty and commitment, but the words of the rule prevail nevertheless.

Although the words of a rule triumph over its purpose in these and many other instances, it is not always so. An often-cited example of purpose prevailing over literal meaning is *United States v. Church of the Holy Trinity*.[23] There, a church had been prosecuted for violating a federal law prohibiting any American employer from paying the passage of an alien employee from a foreign country to the United States for the purpose of taking up employment. The defendant church had done just that, the payment being part of the process of hiring a new pastor, and so the church was in literal violation of the statute. Nevertheless, the Supreme Court held that the statute should not be literally applied in this case. The law, Justice Brewer reasoned, was aimed at employers who were importing large quantities of cheap foreign labor into the United States. And because the church's payment of its new pastor's ocean passage was well removed from what the Court saw as Congress's purpose in enacting the

22. It is not clear what age in 2009 would be equivalent to thirty-five in 1787. In an era in which it is possible—by virtue of television, the Internet, technological advances in publishing, and air travel, for example—to learn far more far earlier than was previously possible, it could be argued that the purpose behind the thirty-five-year rule would be served by lowering the minimum age. But if the framers of the Constitution wanted to ensure that the president was drawn from the older and more experienced segment of the population, then perhaps the underlying rationale would now counsel an age threshold substantially higher than thirty-five.

23. 143 U.S. 457 (1892).

law, the Court concluded that the literal meaning of the words of the statute should yield to the statute's actual rationale, and as a result the church was deemed not to have violated the rule at all.

In reaching this conclusion, Justice Brewer relied on an even earlier case to the same effect, *United States v. Kirby.*[24] In *Kirby,* the defendant was a Kentucky law enforcement officer who had been convicted under a federal law making it a crime to interfere with the delivery of the mail. And that was exactly what Kirby had done. He had unquestionably interfered with the delivery of the mail, but he had done so in the process of boarding a steamboat to arrest a mail carrier named Farris who had been validly indicted for murder by a Kentucky court. As in *Church of the Holy Trinity* twenty-four years later, the Supreme Court in *Kirby* held that the literal words of the statute should not be applied when, as here, applying those words would hardly serve the underlying purpose of the statute.

We will examine additional examples of the tension between language and purpose in Chapter 8, when we take up issues of statutory interpretation. For now, however, these few examples are sufficient to illustrate an important feature of rules—their generality. In contrast to specific *commands—you* take out *this* bag of trash *now*—rules do not speak merely to one individual engaging in one act at one time. Instead, rules typically address many people performing multiple acts over an extended period of time. The speed limit applies to all drivers on all days under all circumstances, just as the rule promulgated by the Occupational Safety and Health Administration (OSHA) requiring hearing protection for workers applies to all factories of a certain type and to all employees in those factories.

Rules are characterized by being general in just this way, but like most generalizations—even statistically sound ones—they might not get it right every time. It is a pretty good generalization that Swiss cheese has holes, but some of it does not. And few people would disagree with the generalization that it is cold in Chicago in January, but warm January days in Chicago are not unheard of. And so too with the generalizations that are part of all rules. But precisely because rules are general, there is always the risk that the generalization that a rule embodies will not apply in some particular case. Even if it is true in *most* instances that drivers should not drive at greater than 55 miles per hour, there will be some cases in which the generalization that driving at more than 55 is unsafe

24. 74 U.S. (7 Wall.) 482 (1868).

will not apply, and when that eventuality arises the rule can be said to be *overinclusive.* The rule includes or encompasses instances that the background justification behind the rule would not cover, as in the *Kirby* and *Church of the Holy Trinity* cases, as with the driver driving safely at 70, and as with an ambulance which might fall within the literal scope of the "no vehicles in the park" rule. In such cases the reach of a rule is broader than the reach of its background justification, and so we say that the rule is overinclusive.

At other times a rule's generalization will be *underinclusive,* failing to reach instances that the direct application of the background justification would encompass. If the purpose of the "no vehicles in the park" rule is to prevent noise, it will be overinclusive with respect to quiet electric cars (which are certainly vehicles) but underinclusive with respect to musical instruments, political rallies, and loud portable radios, all of which are noisy but none of which are vehicles. So too with the rule at issue in *Kirby,* for we can imagine all sorts of impediments to reliable postal service that would not count as an "obstruction" of the mails.

A modern example of both over- and underinclusiveness can be seen in the efforts of an increasing number of states to prohibit driving while talking on a cell phone.[25] The justification for these laws—a justification apparently well supported by the available evidence—is that people who are talking on their cellular phones while driving pay less attention to their driving than they would if they were not on the phone and that this practice is a significant cause of automobile accidents. But those who have objected to such laws say that the laws are overinclusive with respect to drivers who are talking on the phone but still paying attention, and thus the objectors insist that the reach of a "no cell phone" rule is broader than its "no distraction" justification. Moreover, the critics contend, the proposed bans are underinclusive with respect to other sources of distraction while driving, such as eating or listening to an exciting sporting event on the radio. These objections have sometimes prevailed, and sometimes they have not,[26] but it is important to recognize the way

25. See, e.g., Cal. Stat. Ch. 290 (2006), Cal. Vehicle Code § 23123 (2006); N.J. Stat. Ann. 39:4–97.3 (West 2004); N.Y. Vehicle & Traffic Law § 1225-c (Consol. Cum. Supp. 2004).

26. See Note, "The 411 on Cellular Phone Use: An Analysis of the Legislative Attempts to Regulate Cellular Phone Use By Drivers," 39 *Suffolk U.L. Rev.* 233 (2005); Note, "Driving While Distracted: How Should Legislators Regulate Cell Phone Use Behind the Wheel," 28 *J. Legis.* 185 (2002).

in which, as relatively uncontroversial examples like that of the speed limit illustrate, at least *some* degree of both over- and underinclusiveness is an inevitable part of governing human behavior by general rules.[27]

That rules, because of their intrinsic generality, could produce bad results in particular cases was noticed by Aristotle long before there were cell phones and long before the Supreme Court decided cases like *Kirby* and *Church of the Holy Trinity*. In explaining why there needed to be a way of avoiding the mistakes of under- and overinclusion, Aristotle pointed out that "all law is universal," and that "the law takes account of the majority of cases, though not unaware that in this way errors are made. And the law is none the less right; because the error lies not in the law nor in the legislator, but in the nature of the case; for the raw material of human behavior is essentially of this kind."[28]

Aristotle's solution for this problem—equity—will occupy some of our attention in Chapter 6, but for now what is important is only to understand that rules are inevitably general. Rules work as rules precisely because of their generality, and even if it were possible to anticipate every possible application of a rule and incorporate the right result for every application into the rule, such a rule would be too complex to provide the guidance we expect from rules. And even if we were willing to sacrifice intelligibility and useful guidance for precision, we would still be unable to predict the future perfectly. Just as we cannot fault the original drafters of the patent laws for being unable to anticipate in the late eighteenth century that living organisms could be created in the laboratory,[29] so must we recognize that even the most careful of drafters cannot possibly predict what will happen in the future, nor can they predict how we

27. An even more controversial example comes from the efforts of some municipalities to ban certain breeds of dogs—pit bulls, most commonly—on the grounds that some breeds tend to be more aggressive and dangerous than other breeds. Because most pit bulls are not dangerous, however, the ban would be overinclusive, and because dogs of other breeds can be dangerous, the ban would also be underinclusive. In this respect, pit bull bans are little different from rules of any kind, but the opponents of breed-specific bans have nevertheless had considerable success, often by borrowing the language of civil rights and objecting, for example, to "breedism" and "canine racism." For a more extensive discussion and analysis of the controversy, see Frederick Schauer, *Profiles, Probabilities, and Stereotypes* 55–78 (2003).

28. Aristotle, *Nicomachean Ethics* 1137a–b (J. A. K. Thomson trans., 1977).

29. See Diamond v. Chakrabarty, 447 U.S. 303 (1980).

will want to deal with that future when we get there. It is precisely in the inevitable generality of rules, therefore, that we are forced to confront the tension between what a rule says and what it might best be interpreted to do, a tension that pervades the use of rules both in law and outside of it.

2.4 The Formality of Law

There is no uniform answer to whether and when the language of a rule will or should yield to the goal of reaching the best result in the particular case. Nor does the law always give the same answer when there is a conflict between the outcome that would be produced by a rule's background justification and the outcome indicated by the literal meaning of the rule's words. Although cases like *United States v. Locke* show that taking the words at face value even at some sacrifice to reaching the best result for the particular case is common in American law (and even more common elsewhere),[30] so too is the opposite result. Yes, it would be a mistake to ignore the numerous instances such as *Locke* in which what the words most literally say carries the day in legal decision-making. But it is just as much of a mistake to ignore the descriptive importance, in the United States and even elsewhere, of the *Church of Holy Trinity* principle: that achieving a rule's purpose even at some sacrifice to literal meaning is the appropriate course of action.[31] Indeed, if we understand this characterization of the two positions as another way of describing the frequent tension between the letter and the spirit of the law, it is impossible to conclude, especially in the United States, that one approach is more dominant than the other.

Legal arguments for preferring the letter to the spirit of the law are often criticized as *formalistic,* and judicial decisions like *Locke* routinely attract charges of formalism. Yet although it is true that nowadays to call

30. In the United Kingdom, for example, courts are somewhat less likely to ignore the words of a legal rule even when doing so is necessary to serve the rule's background justification. See Patrick Atiyah & Robert S. Summers, *Form and Substance in Anglo-American Law: A Comparative Study in Legal Reasoning, Legal Theory and Legal Institutions* (1987). For a more comprehensive comparative analysis, see D. Neil MacCormick & Robert S. Summers, *Interpreting Statutes: A Comparative Study* (1991).

31. See generally Aharon Barak, *Purposive Interpretation in Law* (2005).

a judge or opinion or decision formalist is rarely a compliment, it is not entirely clear what it is to be a formalist or just what is wrong with it.[32]

Often the charge of formalism is leveled against those who appear to deny the degree of choice available to a judge in some legal controversy. Under this view, judges are being formalistic when they believe that they are operating in the core of a legal rule when in reality they are at the fringe. When Justice Peckham in *Lochner v. New York,*[33] for example, concluded that the word "liberty" in the Fourteenth Amendment necessarily encompassed the freedom of a bakery employee to agree without state interference to work for more than sixty hours a week or ten hours a day, he acted as if no other meaning of "liberty" were even possible. We now know better, of course, and even those who would agree with Justice Peckham's ultimate conclusion would be unlikely to believe, as Justice Peckham appeared to believe, that the outcome was commanded solely by the plain meaning of the word "liberty." When legal decision-makers like Justice Peckham, who are actually (and perhaps, as in this case, necessarily) making a policy or political choice act as if there were no choice to be made—when they treat a policy choice as simply an exercise in knowing the plain meaning of a word—their behavior is sometimes described as formalistic. They act as if it is the form that matters, but in fact it is substance that is doing the work. And it is hard to deny that this form of judicial deception—or self-deception—is worthy of criticism.

Justice Peckham's formalism was the formalism of disingenuousness and fully entitled to the stigma it has attracted. When we look at another conception of formalism, however, the formalism of Justice Thurgood Marshall in *United States v. Locke,* for example, it is not so clear that formalism deserves to be treated as a vice at all. It is, to be sure, formalistic to take the literal meaning of the words "prior to December 31" in *United States v. Locke* as dictating a result other than what seems to be the most sensible one, because it is to treat the *form* of a legal rule as more important than its deeper purpose, or more important than reaching the best all-things-considered judgment in the particular context of a particular case. But although *Locke* is from this perspective formalistic, it is also formalistic in just the same way to use the 55-mile-per-hour speed

32. See Brian Bix, *Jurisprudence: Theory and Context* 179–90 (4th ed. 2006); Robert S. Summers, *Form and Function in a Legal System: A General Study* (2006); Frederick Schauer, "Formalism," 91 *Yale L.J.* 571 (1987).

33. 198 U.S. 45 (1905).

limit to penalize the driver who is driving safely at 70, to penalize the short-swing trader who in fact has no inside information, to allow those with twenty-one-dollar claims to demand a jury trial, and to prohibit otherwise qualified thirty-four-year-olds from becoming president. In all of these cases, law operates formally in treating the meaning of the words of a rule as more important than achieving the law's deeper purpose and reaching the ideal result in this particular case. Formalist this may be, but formalism is, as these and countless other examples demonstrate, a central feature of what makes law distinctive.[34]

That formalism is a part of legalism seems plain enough, but that does not mean that formalism is always desirable. Nor does it mean that a formalist approach to interpreting rules is what we do or should expect from all legal decision-makers at all times. Still, if we can get over the fact that the word "formalism" is typically used to condemn, we can see that formalism—in the sense of preferring the outcome dictated by the words on the printed page rather than the outcome that is best, all things considered—often has much to be said for it. Consider, for example, the numerous cases involving search warrants that turn out to have contained an erroneous address for the premises to be searched. Although many such cases uphold a warrant containing this kind of minor error,[35] there are many that reach the opposite conclusion. So in *United States v. Kenney,*[36] for example, the United States District Court for the District of Columbia invalidated a search of the premises at 2124 8th Street in Washington because the warrant had specified 2144 8th Street, and in *United States v. Constantino,*[37] the contraband actually found in a search of 710 Jacksonia Street was similarly suppressed because the warrant had specified 807 Jacksonia Street. For the courts in those cases, the formal, technical, and literal approach to interpreting the warrant was justified because the real issue was not whether the police officers had searched the right building but whether police officers should be empow-

34. "Of all the criticisms leveled against textualism, the most mindless is that it is 'formalistic.' The answer to that is, *of course it is formalistic!* The rule of law is about form." Antonin Scalia, *A Matter of Interpretation: Federal Courts and the Law* 25 (1997).

35. E.g., United States v. Lora-Sorano, 330 F.3d 1288 (10th Cir. 2003).

36. 164 F. Supp. 891 (D.D.C. 1958).

37. 201 F. Supp. 160 (W.D. Pa. 1962). See also United States v. Ellis, 971 F.2d 701 (11th Cir. 1992).

ered to decide for themselves which premises were *really* to be searched, the exact language of the warrant notwithstanding.

If the formalism of treating search warrants literally is seen as, at the very least, plausible, then it turns out that formalism itself is not necessarily or always to be considered a vice. Rather, the virtues of formalism are part of a larger consideration of whether decision-makers of a certain type should be empowered to decide when the literal language of some rule (and to which the search warrant description, while not exactly a rule, is analogous) should give way to a less constrained determination of purpose, reasonableness, or common sense, for example. Those who defend the result in *United States v. Locke,*[38] for instance, do not maintain that denying Mr. Locke's claim because he filed *on* December 31 rather than *prior to* December 31 is the best or most reasonable outcome in *that* case. Rather, they argue that the real question is whether and when judges should be empowered to decide when the literal language of an act of Congress should be set aside in the service of what the judges believe Congress must have intended or what outcome Congress would have preferred. And when the question is reformulated in this way, it is no longer clear that a formal approach to legal rules is necessarily or always to be criticized, even if the results of that formalism will in particular cases often seem strange and at times even ridiculous.

None of this is to say that law is always formal in this way, or that it should be. As we have seen, courts often do ignore or go around the literal language of a rule when that language is inconsistent with obvious legislative purpose, and it is a mistake to argue that *United States v. Locke* is more representative of legal analysis than *Church of the Holy Trinity* or *United States v. Kirby.* Both the formal and the nonformal (or *purposive*) approaches are professional, respectable alternatives for a judge or advocate in the American legal system, and countless examples can be found in support of both of them. As a result, it is not uncommon to see cases in which one of the parties is arguing on the basis of the letter of the law and the other is relying on a law's spirit, purpose, or rationale. But even when spirit or purpose or rationale prevails, the law remains pervasively formal. It is common for literal language to give way to the purpose behind a *particular* legal rule, as in *Church of the Holy Trinity,* but it is considerably rarer for the purpose behind a rule also to give way when a judge determines that enforcing even that purpose would be in-

38. Including this author. *See* Frederick Schauer, "The Practice and Problems of Plain Meaning," 45 *Vand. L. Rev.* 715 (1992).

consistent with justice, or with larger conceptions of fairness or good policy. When a court denies relief to a litigant with an otherwise valid claim because he has failed to comply with a rule of procedure, for example, the court is recognizing that its job is not merely to decide which of the parties, all things considered, is more worthy.[39] So too when a party is allowed to escape from a contractual promise because of the absence of a requisite contractual formality,[40] or when, prior to the rise in the doctrine of comparative negligence, a plaintiff who was slightly at fault was denied relief against a substantially at-fault defendant.[41] In all of these cases, the pervasive formality of law—its tendency to take its rules and their words seriously even though in some cases they might work an injustice—is what distinguishes law from many other decision-making contexts.

At times law does act otherwise. A prominent example is *Riggs v. Palmer,*[42] in which Elmer Palmer, named as the beneficiary in his grandfather's will, had attempted to accelerate his inheritance by the expedient strategy of murdering the testator. The case did not involve Elmer's criminal conviction for killing his grandfather. To this, Elmer had little defense, and he was duly sentenced to a lengthy prison term. Nevertheless, Elmer claimed that even though he was convicted of and was paying the penalty for murder, he was still entitled to the inheritance. The relevant rule, the New York Statute of Wills, said nothing about murderous beneficiaries and provided only that, upon the death of the testator, the beneficiary under a valid will was entitled to inherit. That was the case here, Elmer argued, and so although he knew that he had to go to prison, he also believed that he was entitled to his grandfather's estate.

The Court of Appeals famously[43] rejected Elmer's claim, concluding that the literal language of the Statute of Wills must yield to the principle

39. See, e.g., General Mills, Inc. v. Kraft Foods Global, Inc., 495 F.3d 1378 (Fed. Ct. 2007); Speiser, Krause & Madole, P.C. v. Ortiz, 271 F.3d 884 (9th Cir. 2001).

40. See Robert S. Summers, "Why Law Is Formal and Why It Matters," 82 *Cornell L. Rev.* 1165 (1997).

41. See, e.g., Miller v. United States, 196 F. Supp. 613 (D. Mass. 1961); Co-Operative Sanitary Baking Co. v. Shields, 70 So. 934 (Fla. 1916).

42. 22 N.E. 188 (N.Y. 1889).

43. The case is analyzed extensively by the legal philosopher Ronald Dworkin in *Taking Rights Seriously* (1977) and *Law's Empire* (1986). Dworkin resoundingly applauds the result and takes it as highly typical of the American (and, to him, better) approach to legal decision-making.

that no person should profit from his own wrong. But there was a dissent even in that case, and it is by no means clear that setting aside the result indicated by a concrete rule in the service of larger and less concrete conceptions of justice is an accurate characterization of the typical nature of legal decision-making. In extreme cases, of which *Riggs v. Palmer* seems an obvious example, specific rules are often set aside, but in cases less extreme than this it is far more common for the rule to be applied even when it seems as if some injustice is done in the process. Indeed, there are many cases in which beneficiaries who were responsible for the death of the testator were allowed to inherit, including one in which the beneficiary was convicted of voluntary manslaughter of the person from whose death he would benefit,[44] another in which the beneficiary was found guilty of being an accessory after the fact in the murder of the testator,[45] still another in which a remainderman had killed the holder of a life estate in order that the killer could take the estate sooner,[46] and, finally, a case in which a "selfish, angry, resentful, indignant, bitter, self-centered, spiteful, vindictive, paranoid, and stingy" woman whose gross negligence served to "shorten the decedent's life" was nevertheless allowed to inherit sooner than would otherwise have been the case.[47]

Just as there are cases in which a rule is allowed to prevail even when an injustice is done in the process, so too are there even more cases in which courts have enforced what they see as bad rules because of the view that changes in bad rules, at least those bad rules that have come from a legislature, are for a legislature and not a court to make. In *Blanchflower v. Blanchflower,*[48] for example, the Supreme Court of New Hampshire was faced with the question whether same-sex adultery could count as adultery for purposes of the New Hampshire at-fault divorce statute, a statute whose language made it clear that adultery could be committed only with a person of the opposite sex. The court appeared to believe that the statute was both anachronistic and morally dubious on equality grounds but nevertheless concluded that any change was to be made by the legislature and not a court. For the Supreme Court of New Hampshire, like the dissenting judge in *Riggs,* like the courts that differ

44. Bird v. Plunkett, 95 A.2d 71 (Conn. 1953).
45. Reynolds v. American-Amicable Life Ins. Co., 591 F.2d 343 (5th Cir. 1979).
46. Blanks v. Jiggetts, 64 S.E.2d 809 (Va. 1951).
47. Cheatle v. Cheatle, 662 A.2d 1362 (D.C. 1995).
48. 834 A.2d 1010 (N.H. 2003).

from *Riggs* and allow people to profit from their known wrongs, and like the Supreme Court in *Locke,* what a legal rule actually says in the literal or plain language of its words made a substantial difference. The letter of a rule may not, as the majority opinion in *Riggs* and the decisions in *Church of the Holy Trinity* and *Kirby* show, always make a difference, and it may not always make all of the difference, but to ignore the ubiquitous importance of what a legal rule literally says is to ignore something very important about rules.

The importance of what a rule actually says is not just a point about rules. More pervasively, to ignore the even more ubiquitous importance of what rules do even when what they do appears unfair is to ignore something very important about law itself. It is not law's purpose, of course, to be unfair for the sake of being unfair. But there is an important group of values—predictability of result, uniformity of treatment (treating like cases alike), and fear of granting unfettered discretion to individual decision-makers even if they happen to be wearing black robes—that the legal system, especially, thinks it valuable to preserve. These values often go by the name of the Rule of Law, and many of the virtues of the Rule of Law are ones that are accomplished by taking rules seriously as rules. In doing so, law remains irreducibly formal and thus at times seemingly unfair in particular cases. But law is more than simply doing the right thing in each individual case. At times law's unwillingness to do just that will seem wrong, but what makes law what it is—usually for better but sometimes for worse—is that it takes larger institutional and systemic values as important, even if occasionally at the expense of justice or wise policy or efficiency in the individual case. There are many ways in which law does this, but the principal one is by taking rules seriously. Understanding when, why, and how rules—as rules—are important in law will take us a long way toward understanding law itself.