10

MAKING LAW WITH RULES AND STANDARDS

10.1 The Basic Distinction

In most American states, the law instructs family court judges to make determinations about child custody and visitation rights in such a way as to further the "best interests of the child."[1] The judge is expected to take testimony about the facts, hear argument, and then make the decision that in his relatively unconstrained discretion will promote the best interests of the child, as opposed to, say, the best interests of the parents or the best interests of society. Similarly indeterminate is the Fourth Amendment to the Constitution, which provides that "[t]he right of the people to be secure . . . against unreasonable searches and seizures . . . shall not be violated." Courts applying the Fourth Amendment thus find themselves with the task of deciding what constitutes a "search," what constitutes a "seizure," and especially which searches and seizures are to be considered "unreasonable."

By way of contrast to open-ended legal terms like "best interests" and "unreasonable," consider the specificity of the regulation under the Occupational Safety and Health Act requiring that on all construction sites with "[m]ore than 20" and "fewer than 200" employees there shall be no less than "[o]ne toilet seat and one urinal per 40 employees."[2] In much the same way, a Securities and Exchange Commission rule promulgated under the Securities Act of 1933 directs registrants to file "three copies of the complete registration statement" on "good quality, unglazed, white

1. E.g., Ariz. Rev. Stat. §25–403 (2006); Cal. Fam. Code §3011 (Parker's 2001); Mass. Gen. L., ch. 119, §23 (2003); Mich. Comp. Laws Ann. 722.23 (West 2001); Wash. Rev. Code §26.10.160 (3) (1994).

2. 29 CFR §1926.65, Table D-65.2 (2002).

paper no larger than 8½ x 11 inches in size."[3] And Article III of the Constitution mandates that "[n]o Person may be Convicted of treason unless on the Testimony of two Witnesses to the same Overt Act."[4]

The difference between the former set of provisions and the latter should be obvious. The first group, including language such as "best interests" and "unreasonable," is broad, vague, general, and imprecise. The second group, with its use of terms like "more than 20" and "unglazed, white paper," is detailed, specific, concrete, and determinate. Conventionally, the difference between the two groups is described as a difference between precise *rules,* such as "more than 20," and vague *standards,* such as "unreasonable." The distinction between rules and standards appears everywhere, and no discussion of legal reasoning would be complete without a careful consideration of this centrally important distinction.[5]

Although rules such as the ones set out above are highly precise, and although the standards just described are vague in the extreme, the difference between rules and standards is actually a matter of degree. Or, to put it differently, the extreme of vagueness—"the best interests of the child" or "unreasonable"—is close to the vagueness/standards end of a spectrum, and the extreme of precision—"two witnesses," "white paper no larger than 8½ x 11 inches in size"—is close to the precision/rules end. In between these extremes there is a continuum, on which some legal directives are more at the rules end and others closer to the standards end. "Drive prudently" is pretty clearly a standard, and "Speed Limit 55" is equally clearly a rule, but "Slow Down for Children," "Yield to Oncoming Traffic," and "Lights on at Dusk" are all somewhere in between. Lo-

3. Rules and Regulations Under the Securities Act of 1933, Rules 402(a) and 403(a), 17 CFR §§230.402(a), 230.403(a) (2002).

4. U.S. Const., art. III, §3, cl. 1.

5. Among the leading analyses are Clayton P. Gillette, "Rules, Standards, and Precautions in Payment Systems," 82 *Va. L. Rev.* 181 (1996); Joseph R. Grodin, "Are Rules Really Better than Standards?," 45 *Hastings L.J.* 569 (1994); Louis Kaplow, "Rules versus Standards: An Economic Analysis," 42 *Duke L.J.* 557 (1992); Duncan M. Kennedy, "Form and Substance in Private Law Adjudication," 89 *Harv. L. Rev.* 1685 (1976); Russell B. Korobkin, "Behavioral Analysis and Legal Form: Rules vs. Standards Revisited," 79 *Or. L. Rev.* 23 (2000). See also Colin Diver, "The Optimal Precision of Administrative Rules," 93 *Yale L.J.* 65 (1983). And an influential application of the distinction to some constitutional issues is Kathleen M. Sullivan, "The Supreme Court, 1991 Term—Foreword: The Justices of Rules and Standards," 106 *Harv. L. Rev.* 22 (1992).

cated more or less midway between the Constitution's requirements of two witnesses to convict for treason and the minimum age of thirty-five to serve as president, on the one hand, and the prohibitions of cruel and unusual punishments and unreasonable searches and seizures, on the other, are the Seventh Amendment's guarantee of a jury trial in "suits at common law" and Article I's prohibition on "bill[s] of attainder." Accordingly, rather than thinking of a distinction between rules and standards, it might be better to think of a location on the vagueness-precision continuum, and "How vague?" or "How precise?" may be far more useful questions to ask than whether some directive is more like a rule or more like a standard.

10.2 Rules, Standards, and the Question of Discretion

The rules-standards continuum is important for many reasons, but foremost among them is the way in which selecting the point on the spectrum can be a highly effective device for the management of discretion. Discretion is, of course, a central concern of the law. Under what circumstances will which officials be given the freedom to exercise their own judgment and make their own choices, and under what circumstances will that freedom be constrained or even mostly eliminated? When a judge is determining which custody decision is in the best interests of the child, for example, she might be faced with a choice between a wealthy mother who can provide the child with high-quality education, housing, recreation, and culture and a less wealthy father who appears to understand the child better than the mother. Under these circumstances, some judges would prefer the mother and others the father, but the basic idea of discretion is that neither of these decisions, given the "best interests of the child" standard, would be legally incorrect. People might criticize one or the other decision for being morally wrong, psychologically ignorant, or based on erroneous factual premises, and so it would not be correct to say that the two decisions are equally right. But it would be correct to say that the two decisions are equally *legally* right, and thus that either decision would, in the ordinary course of things, be upheld on appeal. By contrast, if instead of the "best interests of the child" standard there existed a rule stating that custody was always to go to the wealthier parent, then a judge would no longer have the discretion to award custody to the less wealthy but more caring parent, and we would expect that a deci-

sion to do so would be overturned on appeal as outside of the judge's discretion.[6]

This idea of discretion has many different labels. Sometimes it is called leeway, sometimes it is seen as a variety of deference, and in European Community law and European human rights law, and sometimes elsewhere, it is referred to in terms of a *margin of appreciation,* the latitude that member states are given in making their own judgments about matters ultimately controlled by European law.[7] But regardless of the label, the idea of discretion is that some institution with the power to control or review will let stand a multiplicity of quite different decisions, including some that the controlling institution might think wrong. Just as the "no vehicles in the park" rule may permit a judge discretion about whether to include or exclude bicycles, skateboards, and baby carriages, so does any other official with discretion have the power to make any one of a number of different decisions.

Analyzing the idea of discretion is not our primary goal here. But from the perspective of the devices of legal reasoning, it is important to appreciate that a common way of granting discretion is to couch the governing law or regulation as a standard. When Montana eliminated its numerical speed limit for a few years in favor of the vague requirement—a standard—that driving be "reasonable and prudent," it gave police officers the discretion to decide what speeds under what conditions were reasonable and prudent, and it similarly gave to traffic court judges much the same discretion.[8] The Montana Supreme Court subsequently ruled that "reasonable and prudent" was too much of a standard to support a statute that provided for criminal penalties, but the basic idea here

6. "Abuse of discretion" is a very common phrase, and indeed abuse of discretion is commonly a ground for judicial invalidation of an administrative decision. See generally Charles E. Koch, "Judicial Review of Agency Discretion," 54 *Geo. Wash. L. Rev.* 469 (1986). The phrase only makes sense, however, if it is assumed that an official or judge has considerable decisional freedom, or leeway, short of abuse.

7. See Douglas Lee Donoho, "Autonomy, Self-Governance, and the Margin of Appreciation: Developing a Jurisprudence of Diversity within Universal Human Rights," 15 *Emory Int'l L. Rev.* 391 (2001).

8. Mont. Code Ann. 61–8–303 (1996), *invalidated as vague 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).

should be clear. When a directive is expressed in broad and vague terms, it grants discretion to and thus empowers not only those, like police officers, who are charged with enforcing the directive, but also those, like judges, whose task it is to interpret it.

Conversely, therefore, the use of directives at the rules end of the continuum can be understood as a device for constraining or withdrawing discretion. If a filing deadline for, say, the signatures necessary to put a candidate on the ballot is 5:00 P.M., the official who decides whether to accept the filing has far less discretion than would be the case if the requirement were only that the filing be "timely" or "sufficiently prior to the election as to permit the ballots to be printed."[9] Similarly, Article 35(1) of the Constitution of South Africa provides that

> [e]veryone who is arrested for allegedly committing an offence has the right . . . to be brought before a court as soon as reasonably possible, but not later than 48 hours after the arrest; or the end of the first court day after the expiry of the 48 hours; if the 48 hours expire outside the ordinary court hours or on a day which is not an ordinary court day.

This provision is, to put it mildly, vastly more precise than the American counterparts of mandating a "speedy" trial in the Sixth Amendment and "due process of law" in the Fifth, and this should come as no surprise. Police, prosecutorial, and even judicial abuses of power were widespread in the apartheid era, which ended with the 1995 enactment of the new South African constitution, and we can thus appreciate a reluctance at the time the constitution was drafted to grant discretion to mostly the same police officers, prosecutors, and judges who had been in power during apartheid. And even after the makeup of officialdom began to change, we can still understand the residual skepticism about official discretion in the criminal justice system. This skepticism about official dis-

9. See the unreported Vermont case of *Hunter v. Norman,* described in Frederick Schauer, "Formalism," 97 *Yale L.J.* 509 (1988).

Insofar as administrative officials or judges believe they have some flexibility in the matter, would they be more likely to be flexible with a "5 pm" deadline than with a "5:00 P.M." deadline, and more likely to be flexible with a "5:00 P.M." deadline than with a "5:08 P.M." deadline? Are drivers more likely to exceed a posted speed limit of 55 than one of 54 or 57? The very familiarity of 55—the conventionality of round numbers—will lead drivers to expect the normal leeway, but 54 cries out, "We mean it!"

cretion would naturally lead to the choice of rules over standards, and the highly specific provisions quoted above are the consequence of just this skeptical attitude.

Skepticism about discretion is not always a function of distrust of officials. Sometimes we worry about excess discretion because we are concerned with the lack of predictability that too much discretion will bring. If I want to know how fast I can drive, being told that I must drive reasonably and prudently will not answer my question, any more than being told that I must not enter into contracts in restraint of trade will help me to know how to comply with the antitrust laws. Indeed, this is one of the reasons why the Supreme Court has traditionally created so-called per se antitrust rules, under which schemes such as price-fixing, resale price maintenance, and tying arrangements were understood to violate the Sherman Act without regard to their economic effects in the particular case.[10] Some of these per se rules have been eliminated or weakened,[11] but the basic point persists that a per se rule is likely to give far more information about likely consequences of action than a less determinate standard.

Sometimes we wish to constrain discretion simply in the service of efficiency. A food inspection official at a border customs station can make a quick determination of whether a product being carried into the country is or is not "meat," but if the same official were empowered to determine on a case-by-case basis whether a product was "unsafe" or "unhealthy," he might have far less time to watch out for drug couriers and terrorists. Not only do precise rules require officials to spend less time on routine decisions, but they also allow the designers of decision-making environments to employ people with less skill or experience. Deciding whether a certain kind and level of factory waste dumped into a river is harmful to the environment probably requires an experienced environmental expert, but deciding whether the concentration level of a certain chemical is above or below a specified number can far more easily be entrusted to a less experienced and less well trained technician.

10. E.g., Times-Picayune Publishing Co. v. United States, 345 U.S. 594 (1953) (tying arrangements); United States v. Socony-Vacuum Oil Co., 310 U.S. 150 (1940) (price-fixing); Dr. Miles Medical Co. v. John D. Park & Sons, Inc., 220 U.S. 373 (1911) (per se rule against resale price maintenance).

11. See, most recently, Leegin Creative Leather Products, Inc. v. PSKS, Inc., 127 S. Ct. 2705 (2007), eliminating the per se rule for resale price maintenance.

Although these efficiency concerns are important, the use of directives toward the rules end of the rules-standards continuum will typically be a product of a determination that making individualized judgments in each case will produce too many unacceptable errors. That was the impetus behind the Federal Sentencing Guidelines,[12] a highly precise and constraining sentencing system, since softened somewhat by the Supreme Court,[13] intended to replace a regime of sentencing discretion that had generated wide disparities in sentencing for the same crimes. The use of vaguer directives at the standards end, conversely, will typically reflect a judgment that individualized or case-by-case determination is what is most important, preferring to endure the errors of individualized judgment to the errors that will come from the over- and underinclusiveness of rigid and precise rules. Indeed, something like this goal is what led the Supreme Court to make the Federal Sentencing Guidelines somewhat less rulelike. But there is no strategy that will be best in all contexts, and thus the lesson may be that the determination of how much officials should be allowed to look at the particular context of a particular instance—how much the official should be operating under a standard rather than a rule, or vice versa—will itself be a contextual determination.

10.3 Stability and Flexibility

It is true that location on the rules-standards continuum is an important way of allocating discretion between the issuer of the directive and those who must apply it, enforce it, or interpret it. But the difference between rules and standards is also a way of allocating decision-making between the present and the future. When a legislature, agency, or court sets forth a rule, it is making a decision *now* about what is to be done in the future. And when instead it moves toward the standards end of the continuum, it holds things open for the future and allows for a flexible approach to the problems of tomorrow.

Accommodating to a future that we can at best dimly perceive is a re-

12. Pub. L. No. 98–473, 98 Stat. 1837, 2017, *codified as* 18 U.S.C. §§ 3551–3673, 28 U.S.C. §§ 991–998 (2002). For an authoritative account of the original goals, see Stephen Breyer, "The Federal Sentencing Guidelines and the Key Compromises on Which They Rest," 17 *Hofstra L. Rev.* 1 (1988).

13. See Rita v. United States, 127 S. Ct. 2456 (2007); United States v. Booker, 543 U.S. 220 (2005).

curring problem in the law, and one that is almost certainly growing. With respect to rapid changes in electronic communications, for example, any attempt to write specific laws at this time is as destined to obsolescence as were those specific rules of only a generation ago that did not and could not have anticipated widespread cell phone usage, the development and dominance of the Internet, and a density of satellite-based communication technologies that makes a list of communications media limited to radio, television, and motion pictures seem archaic. So too in many other areas. Developments with respect to cloning, genetic modification, DNA identification, and much else will almost assuredly doom to failure, or at least to predictable obsolescence, any attempt to use the scientific knowledge of today as a basis for the categories of the rules that we expect to have effect in the future.

Just as the location of the rules-standards continuum can help to allocate discretion, therefore, so too can it allocate decision-making authority between the present and the future. A communications regulation that is drafted as a highly specific rule and that therefore inevitably draws on current and not future knowledge allocates power today to determine outcomes in and for an uncertain future, while a much vaguer standard—"medium of communication," for example—would let the future make decisions for itself, but at the cost of less guidance and less precision for the present.

The allocation of decision-making authority between the present and the future thus presents the fundamental tradeoff in the question about rules and standards. Highly specific directives—rules—will maximize certainty, constraint, and predictability, but they will do so at the cost of retaining the ability to achieve exactly the correct result in some currently unanticipated case or situation. And much vaguer directives—standards—will hold open the ability of some future decision-maker to make just the right decision (assuming for the sake of argument that that decision-maker will in fact do so), but at the cost of providing very little certainty, predictability, and decision-maker constraint.

There is, to repeat, no right solution to this inevitable tradeoff. Nor can it be said that one or another approach is more or less consistent with the values of the Rule of Law, because although some Rule of Law values are served by precise, predictable, and understandable rules, others are served by relatively open-ended standards that will allow judges and other official decision-makers the discretion to do justice in the individual case. But although there is no easy or consistent answer, understand-

ing the way in which the rules-standards continuum is a valuable tool for allocating authority among officials, allocating decisions between the present and the future, and allocating our concern between predictability and individualized justice will enable participants in the legal system to understand one of the most important devices of legal and regulatory institutional design.

10.4 Rules and Standards in Judicial Opinions

The distinction between rules and standards is central to questions of legislative control and administrative discretion, and thus it is common to think of rules and standards in terms of the degree of specificity and vagueness in a constitutional provision, in a statute, and especially in an administrative regulation. But judicial opinions can also be sources of guidance, command, and authority, and as such they are equally susceptible to analysis in terms of the extent to which the guidance offered by a judicial opinion is more or less like a rule or a standard.

Consider, almost at one extreme of the rules-standards continuum, the Supreme Court's 1966 decision in *Miranda v. Arizona,*[14] in which the Court held that the Fifth Amendment (either directly, with respect to the federal government, or through the Fourteenth Amendment, as applied to the states) mandated that a confession or other statement by a suspect subject to custodial interrogation could be used against him only if the suspect had been warned of his constitutional right to remain silent as well as his right to a lawyer. Having reached this conclusion, the Court could have announced what was in effect a standard for the lower courts to follow in evaluating subsequent claims under *Miranda*. The Court could have said, for example, that a confession or other statement will be admissible if and only if it is found that the statement had been made *voluntarily,* and it could have made clear that the determination of voluntariness was a contextual judgment that should consider all of the circumstances in which the statement had been made. Or the Court could instead have said that statements taken *unfairly* or *unjustly* or *coercively* would be inadmissible, again leaving it to lower courts in particular cases to decide in light of the circumstances whether a statement was taken unfairly, unjustly, or coercively. Had the Court proceeded in this way, by the use of these kinds of flexible, contextual, and vague requirements, we

14. 384 U.S. 436 (1966).

would have said that the Court had elected to set forth a standard to be followed by the police and lower courts alike.

As is well known to viewers of television and movies, to say nothing of lawyers and judges, the Court did not proceed this way in *Miranda*. Instead, it told police officers more or less exactly what to say: "You have the right to remain silent. Anything you say can and will be used against you in a court of law. You have the right to a lawyer. If you cannot afford a lawyer, one will be appointed for you." If the police did not say something pretty much exactly like this prior to a custodial interrogation, the Court emphasized, any statements made by the suspect to the police would be excluded from the trial. And in deciding whether the police had acted consistently with the requirements of *Miranda* and the Fifth Amendment, lower courts were charged with determining little more than whether the police has uttered something very close to these magic words.

The Supreme Court's actual approach, therefore, was very much a rulelike one. The Court told the police and the lower courts exactly what they should do, and it is remarkable that police officers now essentially read from a Supreme Court opinion when they are interrogating a suspect. The words on a so-called *Miranda card,* the text that police officers carry with them and read aloud to those they have apprehended, are almost word-for-word taken from the Supreme Court's opinion. And in so laying down exactly what police offers should do, the Court acted very much like a legislature or highway department deciding that drivers should drive no faster than a posted numerical speed limit, rather than telling drivers that they should simply drive prudently or carefully or reasonably.

It is controversial whether courts should proceed as the Supreme Court did in *Miranda,* or as the Court did in *Roe v. Wade*[15] when it set forth the precise trimester approach to the restrictions that states might permissibly place on the right of a woman to an abortion. For some critics, laying down detailed rules is for a legislature, and a court exceeds its authority and goes beyond the particular competence of a *court* when it does things that look legislative.[16] But this criticism seems odd, at least if

15. 410 U.S. 113 (1973).

16. See Robert F. Nagel, "The Formulaic Constitution," 84 *Mich. L. Rev.* 165 (1985). See also Akhil Reed Amar, "The Supreme Court, 1999 Term—Foreword: The Document and the Doctrine," 114 *Harv. L. Rev.* 26 (2000).

we assume that other courts, policymakers, and ordinary people should be guided by what the courts have done.[17] If we think that courts are only in the business of making decisions for the parties, then perhaps it is not so bad that others cannot be guided by those decisions. But especially for the Supreme Court, which these days decides so few cases, such a view of the Court's role seems as inefficient as it is unrealistic. There are rules that are addressed to courts and purport to tell courts how to decide cases, and there are rules that are addressed to the citizens and officials who wish simply to know what to *do,*[18] and judicial opinions that resemble the directives at the rules end of the rules-standards continuum are often quite plausibly focused on providing a source of guidance for citizens and nonjudicial officials alike. Especially where a judicial opinion deals with conduct that is repeated daily by numerous individuals, rulelike precision brings the virtue of providing reasonable advice to large numbers of people, and the advantages of doing so may often outweigh the disadvantages that come from relinquishing standardlike flexibility.[19]

Joining those who criticize courts when they issue rulelike directives are others who have argued that appellate courts should decide only "one case at a time,"[20] often insisting that this is what courts do best or that this is the only thing that courts legitimately ought to do.[21] The question of legitimacy may be somewhat remote from issues about legal reasoning and the rules-standards continuum, but not so for the other dimensions of the argument for deciding one case at a time—for deciding a particular controversy and not setting forth broad prescriptions for deciding others. And one argument for this approach is that making broad decisions—making decisions that in effect decide cases not before the

17. See Henry P. Monaghan, "The Supreme Court, 1974 Term—Foreword: Constitutional Common Law," 89 *Harv. L. Rev.* 1, 20–21 (1975); Frederick Schauer, "Opinions as Rules," 62 *U. Chi. L. Rev.* 1455 (1995).

18. See Stephen McG. Bundy & Einer Elhauge, "Knowledge about Legal Sanctions," 92 *Mich. L. Rev.* 261 (1993); Meir Dan-Cohen, "Decision Rules and Conduct Rules: On Acoustic Separation in Criminal Law," 97 *Harv. L. Rev.* 625 (1984).

19. See Frederick Schauer, "Abandoning the Guidance Function: *Morse v. Frederick*," 2006 *Sup. Ct. Rev.* 205.

20. See Cass R. Sunstein, *One Case at a Time: Judicial Minimalism on the Supreme Court* (1999).

21. See Edward A. Hartnett, "A Matter of Judgment, Not a Matter of Opinion," 74 *N.Y.U. L. Rev.* 123 (1999).

court—requires fact-finding capabilities that are beyond what we can expect in the appellate process. When a court decides an issue broadly—when it decides that *all* custodial interrogations by *all* police officers of *all* suspects under *all* circumstances must be preceded by a warning of the kind the Supreme Court delineated in *Miranda*—it is in effect deciding a whole bunch of cases, but in the process it is required to guess about what those other cases might look like. It is likely, for example, to assume that many of these other cases will look like the case before it, an assumption consistent with what psychologists refer to as *availability*[22]—the belief that that which is most cognitively accessible to us is representative of some larger class of acts or events or cases. But often the facts on which an appellate court must focus are not at all representative, and to that extent making broad rulings in the context of concrete cases may not be the best way to lay down broad principles of law.[23]

Once again, there is no right or wrong answer to the question of whether appellate courts should lay down broad rules in the process of deciding cases or whether instead they should focus on producing the best answer for the case at hand, leaving other cases for other occasions and other decision-makers. The decision about how much to decide, however, and thus of how much law to make—it is far too late in the day to think that courts do not make law—is much like the distinction between rules and standards. It involves the allocation of decision-making authority among potential decision-makers; it involves the assignment of decision-making responsibility between the certain present and the uncertain future; and it involves the pervasive tension between the advantages of flexibility and the competing virtue of letting citizens, officials, lawyers, and other courts know what the law is, even if the law they

22. The original insight is in Amos Tversky & Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," 185 *Science* 1124, 1127 (1974), and there is now a voluminous literature. See, e.g., *Heuristics and Biases: The Psychology of Intuitive Judgment* (Thomas Gilovich, Dale Griffin, & Daniel Kahneman eds., 2002); Scott Plous, *The Psychology of Judgment and Decision Making* (1993); John S. Carroll, "The Effect of Imagining an Event on Expectations for the Event: An Interpretation in Terms of the Availability Heuristic," 14 *J. Exp. Soc. Psych.* 88 (1978).

23. See Neil Devins & Alan Meese, "Judicial Review and Nongeneralizable Cases," 32 *Fla. St. L. Rev.* 323 (2005); Jeffrey J. Rachlinski, "Bottom-Up versus Top-Down Lawmaking," 73 *U. Chi. L. Rev.* 933 (2006); Frederick Schauer, "Do Cases Make Bad Law?," 73 *U. Chi. L. Rev.* 883 (2006).

know may not always be the best law that could be developed for every individual dispute and every particular act.

10.5 On the Relation between Breadth and Vagueness

We commenced this chapter by describing the conventional distinction between precise rules and vague (or indeterminate) standards, and we have moved gradually to the question of whether courts should issue broad rulings or narrow ones. This is the appropriate place to point out, therefore, that the scale of broad to narrow has little to do with the scale of precise to vague.

We can start with a simple example. Take the category of insects. It is a huge category. In the first place, there are lots of insects. Trillions of them. Lots more than there are mammals, fish, and birds combined. And there are lots of *kinds* of insects. Again, there are far more species of insects than of all the other members of the animal kingdom put together. Yet although there are lots of insects and lots of kinds of insects, the category of insects is very determinate, and so is the word "insect." Like any other word, it is not perfectly determinate, and there are certainly contexts in which we can imagine struggling with whether a toy insect or a dead insect is really an insect, just as we would struggle with whether a bicycle or a skateboard is a vehicle for the purpose of the "no vehicles in the park" rule. Yet for the category of insects, the proportion of borderline cases is very small. It is not nonexistent, but it is much smaller compared to the class of nonborderline cases than would be the case for vehicles, say, or for schools, or, to take an example made famous by the philosopher Ludwig Wittgenstein, for games.[24] What the example of in-

24. "Someone says to me: 'Shew the children a game.' I teach them gaming with dice, and the other says 'I didn't mean that sort of game.'" Ludwig Wittgenstein, *Philosophical Investigations* 33 (G. E. M. Anscombe trans., 3d ed., 1958). The example has been used and misused countless times and has spawned a voluminous literature, but the basic idea is that the word "game," as used in the language, refers to many different things that may have what is called a *family resemblance* among them. That is, there is no list of necessary and sufficient features that all proper applications of the word "game" share. Even so, however, there are clear cases of games—baseball, for example, or chess—and there are marginal cases—playing the stock market, perhaps—and the inability to specify the necessary and sufficient conditions for what makes the clear cases clear does not mean that such clear cases do not exist.

sects teaches us, therefore, is that broad categories can be extremely precise. And of course narrow categories can be vague. Even though there are far fewer cases of heroism than there are insects, and though most people would agree that heroism is quite rare, the term itself is highly vague, and there would be little agreement about what counts as a heroic act and what does not.

In law, the distinction between the precise and the vague is largely a question about discretion, a question about flexibility, and a question about the competing aims of predictability and individualized justice. And the distinction between the broad and the narrow is largely a question about how much courts or legislators or other rulemakers should do when they are deciding a case or laying down a rule. Both the distinction between the broad and the narrow and the distinction between the precise and the vague—and both, of course, are questions about a place on a continuum and not simply about one thing or another—are important, but they are not the same. A court concerned with not deciding too much but also concerned with giving guidance might, for example, make a decision containing a precise but narrow rule. The Supreme Court might still have specified, as in *Miranda,* close to the exact words that a warning would have to contain, but might have limited, as it did not, the cases to which those words would have to be given to a narrow class of police interrogations or to a narrow class of crimes. In such case, it would have selected an approach to rulemaking that was on the rules end of the rules-standards continuum but on the narrow end of the broad-narrow one. And in doing so, it would have done something quite different from what *was* done in *Miranda,* where the Court set forth a rule that was, like the class of insects, both broad and precise.

Conversely, the Court in a different hypothetical variation on the real *Miranda* case might have decided that flexibility was more important than guidance. It might have decided, for example, that it could not in 1966 predict all of the future possibilities for police conduct and misconduct, or that the case it had before it—an actual controversy involving the state of Arizona and a particular individual named Ernesto Miranda in the context of a particular criminal case with particular facts—did not provide sufficient information for the Court to be making broad rules for the future. It might, therefore, have chosen to prefer something more like a standard than a rule—confessions would be inadmissible if they were "uninformed," for example. But it would still have had to make a choice about the scope of application of that standard. It might again have de-

cided that the "informed" (or "uninformed") standard would apply to all custodial interrogations, or it might have concluded that it would be best to restrict its application to a much narrower class of police actions. At times the arguments for proceeding by way of a standard will be similar to those for deciding narrowly rather than broadly—both are ways of leaving some determinations for the future, for example—but it is still the case that the question of the scope of a rule or a ruling is different from the question of its precision.

Entire books could be written about the techniques of rulemaking,[25] and indeed some of the manuals on legislative drafting come close to having this aim. But rulemaking is not just for legislatures, and once we recognize that courts as well as legislatures make law and that courts as well as administrative agencies make rules, we can begin to think of judicial rulemaking as a task worthy of far closer attention than it has received to date.

25. Surprisingly, and perhaps disturbingly, very few such books have in fact been written. Maybe it is thought that such things are self-evident, but they are not. And maybe it is thought that the techniques for rulemaking can be picked up from other places, but again there is quite a bit of evidence to the contrary.

11

LAW AND FACT

11.1 On the Idea of a Fact

Most discussions of legal reasoning and legal argument, including much of this book, tend to focus a great deal on law and not very much on facts. The standard treatments assume that the interesting issues in *Donoghue v. Stevenson*[1] are about whether Mrs. Donohue ought to be able to recover against the ginger beer bottler despite the absence of privity, and mostly ignore the question of whether it really was a decomposed snail that came out of the bottle or just how ill, if at all, the sight of the snail actually made her. We know after *Raffles v. Wichelhaus*[2] that when both of two contracting parties are fundamentally mistaken about the object of the contract, there is no contract at all, but how do we know that there were two ships named *Peerless,* and how do we know that each of the parties really was mistaken? *R. v. Dudley & Stephens*[3] is a staple of criminal law classes, but just how hungry really were the survivors, and just how close to death was the cabin boy before he was killed for the alleged survival of the others? And although the Supreme Court in *Brown v. Board of Education*[4] appeared to base its conclusion on the proposition that racially separate but physically equivalent educational facilities impaired the education of black children, how did the Court obtain that information, and was the information it obtained correct?

All of these questions are *questions of fact:* Was it a decomposed

1. [1932] A.C. 562 (H.L.).
2. 2 H. & C. 906, 159 Eng. Rep. 375 (Ex. 1864).
3. 14 Q.B.D. 273 (1994). For an engaging and important account, see A. W. B. Simpson, *Cannibalism and the Common Law* (1984).
4. 347 U.S. 483 (1954).

snail? Were there two ships named *Peerless,* or only one, or maybe even three? How close to death were the shipwrecked sailors? Do black children get a worse education in an all-black, legally segregated school whose physical facilities and teacher training are the same as those in the all-white schools? These questions are traditionally contrasted with *questions of law:* Is a manufacturer (or bottler) directly liable to the consumer when there is a decomposed snail in a ginger beer bottle? Is there a contract when the contracting parties have different beliefs about what they are contracting for? Is dire necessity a defense to a charge of murder? Does a separate but nominally equal racially segregated school system violate the Fourteenth Amendment? The typical legal decision involves an initial assessment of what happened—the question of fact—and then moves on to a determination of what the law should *do* in light of what has happened—the question of law.

The distinction between questions of law and questions of fact is not without difficulty. A controversy about how to explain the difference between law and fact has generated a substantial body of commentary,[5] even including the view that the distinction is entirely illusory.[6] Much of the debate centers on the implications of the way in which, conventionally, the jury (or a judge explicitly serving as the trier of fact) is charged with determining the facts, while the judge has the job of interpreting and (perhaps) applying the law. In reality, however, juries make many decisions that partly involve determinations of law, such as whether someone's actions were "reasonable" or whether the defendant's actions "caused" the plaintiff's injury. Conversely, judges commonly make factual determinations when they are reaching legal conclusions, sometimes just by virtue of having to make the factual determination that some rule or precedent is or is not the law, and sometimes because, especially with respect to constitutional issues, making determinations about facts is part of what we want judges to do in order to ensure that constitutional values are preserved.[7]

5. See, e.g., Richard D. Friedman, "Standards of Persuasion and the Distinction between Fact and Law," 86 *Nw. U.L. Rev.* 916 (1992); Henry P. Monaghan, "Constitutional Fact Review," 85 *Colum. L. Rev.* 229 (1985); Stephen A. Weiner, "The Civil Jury and the Law-Fact Distinction," 54 *Cal. L. Rev.* 1867 (1966).

6. Ronald J. Allen & Michael S. Pardo, "The Myth of the Law-Fact Distinction," 97 *Nw. U.L. Rev.* 1769 (2003).

7. See David Faigman, "'Normative Constitutional Fact-Finding': Exploring the Empirical Component of Constitutional Interpretation," 139 *U. Pa. L. Rev.*

Yet although the fact-law distinction in law can become muddied quite quickly, the confusion does not always stem from the lack of a fundamental distinction between fact and law, which becomes far less mysterious if we just think of it as a variation on the venerable distinctions between fact and value, is and ought, and description and prescription.[8] Rather, the confusion comes from the way in which the law has traditionally insisted that facts are for juries and the law is for judges, when in reality many of the things that juries do by way of law application involve making legal determinations, and many of the things that judges do involve making factual ones. If we accept that the distinction between law and fact does not and could not track the distinction between what judges do and what juries do, then we need not reject the basic distinction between what happened and what someone ought to do about it in order to recognize that making factual determinations is a central part of reasoning and argument at all stages of the legal system.

Thus, although legal decisions, even those made by judges and even those made in appellate courts, typically involve both factual and legal elements, discussions of legal reasoning have traditionally focused overwhelmingly on the latter only.[9] They have assumed that thinking about factual questions is for the law of evidence or that making factual determinations is not really a matter of legal reasoning at all. But given that questions of law almost always turn on determinations of fact, and given that determinations of fact are in numerous ways structured by legal rules and by characteristic ways of reasoning, to exclude questions of fact from the topic of legal reasoning seems peculiar. In this chapter, therefore, we shall take up the question of questions of fact and examine the

541 (1991); Monaghan, *supra* note 5; Note, "Corralling Constitutional Fact: De Novo Fact Review in the Federal Courts," 50 *Duke L.J.* 1427 (2001).

8. There are, of course, controversies about and challenges to these venerable distinctions as well, some but not all of which come from perspectives loosely labeled as "postmodern." And it is true that many purported descriptions have a normative element to them, with values being smuggled in under the cover of purported neutral description. Nevertheless, it is sufficiently implausible to insist that there is no difference between "John fired a gun whose bullet entered Mary's heart and caused her death" and "John ought to go to prison for murdering Mary" that allegedly sophisticated challenges to any of the distinctions in the text need not detain us any further here.

9. A noteworthy exception by a prominent Legal Realist is Jerome Frank, *Facts on Trial: Myth and Reality in American Justice* (1949).

reasoning processes that legal decision-makers use to determine in the first instance simply what happened.

11.2 Determining Facts at Trial—The Law of Evidence and Its Critics

In the normal course of things, determining what happened is for the trial court. Did the defendant shoot her husband? Was that the testator's authentic signature at the bottom of a document that appears to be a will? What kind of damage did the overflowing water cause in *Rylands v. Fletcher,*[10] and how much would it cost the plaintiff to repair it? These issues are normally determined at trial and not on appeal, and they are determined by the person or institution we call the "trier of fact." The classic trier of fact in common-law legal systems[11] is the jury, although it turns out that in many criminal cases and most civil ones the determination of what happened is made by the presiding judge.

If we set aside the law for a moment, we can appreciate the fact that there are multiple ways of finding out something about the world. Outside of the legal system, for example, a common method of determining what happened in the past is to go out and investigate, just as police detectives do when a crime has been committed, and just as congressional investigators do when Congress initiates an inquiry into the cause of a disaster such as the explosion of the *Challenger* space shuttle or the nuclear leaks at Three Mile Island. Investigation itself takes many forms, but all share the idea that the investigators go out into the field, ask questions, poke around, interview witnesses, examine physical evidence, and then make the decision themselves.

In other contexts, particularly in science, the way to find out about something is to conduct an experiment. Sometimes the experiment will be conducted in a laboratory, sometimes it will involve some variation on giving some people a drug and others a placebo, and sometimes scientists and others can analyze a natural experiment, the situation in which the world rather than the scientist creates the conditions in which almost ev-

10. 3 L.R.E. & I. App. 330 (H.L. 1868).

11. In general there are no juries in civil-law systems, and judges both determine the facts and apply the law. In some civil-law systems, however, judges will occasionally try cases, especially criminal cases, in conjunction with several laypeople typically known as "assessors."

erything is the same except for some consequence or symptom whose cause we wish to identify. And empirical social scientists often find out about the world by collecting and analyzing data, often in the large computerized arrays of information called data sets. They run regressions using different variables, typically in an attempt to locate the causes and consequences of various social phenomena.

There are, to be sure, other forms of discovering facts of the world, but cataloging all of them here would serve no purpose. The point of mentioning of few of the more widespread fact-finding methods, however, is to highlight the fact that the law's characteristic way of determining what happened is hardly universal and hardly the only way of finding out about things, even the things that the law would need to know for its own purposes. Indeed, the fact-finding methods that we associate with the law in the common-law world—adversarial trials in which whatever information the judge or jury has on which to base its decision is supplied by the parties—are themselves hardly universal. In France, for example, judges play an active role in managing and conducting the more serious criminal investigations,[12] and variations of this approach are seen in many other civil-law countries. In England prior to the fifteenth century, jurors were largely self-informing, expected to rely in part on their personal knowledge of the litigants, in part on their personal knowledge of the situation, and in part on what they could find out by their own investigations.[13] The idea that a jury—or the judge serving as the trier of fact—should be largely ignorant of the specific litigants and the specific facts prior to the trial itself is a relatively modern invention and hardly a universal one. But even apart from the question of the jury's prior knowledge, the view that the best way to make a factual determination is to allocate to the parties all of the burden of coming forth with evidence and then to have a group of nonexperts evaluate that evidence in an adversary mode, rather than, say, an investigative or collaborative one, and rather than relying on people who might have relevant expertise, is hardly self-evident. Nor is the common law's adversary method self-

12. An intriguing and instructive narrative is Bron McKillop, "Anatomy of a French Murder Case," 45 *Am. J. Comp. L.* 527 (1997).

13. See Sanjeev Anand, "The Origins, Early History and Evolution of the English Criminal Trial Jury," 43 *Alberta L. Rev.* 407 (2005); Thomas A. Green, "A Retrospective on the Criminal Trial Jury, 1200–1800," *in Twelve Good Men and True* 358 (J. S. Cockburn & Thomas A. Green eds., 1988).

evidently wrong, and indeed it has its counterparts in other decision-making environments. The Roman Catholic Church, after all, has institutionalized the concept of the devil's advocate as a way of ensuring that the initial impression of a candidate's sainthood is not accepted as final before hearing the best argument against the proposed saint's actually having been one. Thus, the determination of facts in most common-law countries is premised on the belief that adversarial procedures in which the parties have the primary responsibility for coming forth with evidence are valuable ways of determining the truth, even if they are not the only ones. Just as one argument for a system of freedom of speech is based on the assumption that a good way of finding out the truth is through the clash of opposing ideas, the adversarial process relies on similar assumptions.[14] Let the parties bring forward their evidence, let that evidence be subject to the particular form of scrutiny we call cross-examination, and then let the truth, or at least the closest approximation of it we can achieve, emerge. Or so we believe.

This is not the place to evaluate the adversary system as a method of discovering the truth, whether for the Catholic Church, for public deliberation, or for the law. But contrasting the law's methods of fact-finding with others that are or have been used in other contexts or other countries does put the law's method of fact-finding in proper perspective. Moreover, contrasting the adversary system of fact-finding with others reminds us that that jurors or even judges are not only at the mercy of the parties in terms of what evidence they can consider, but are also prone to a host of cognitive failures—bias, inattention, and countless others—that affect most human decision-makers. Indeed, a large body of social science research, mostly by psychologists, concentrates not only on how jurors—and judges, for that matter—might be subject to many of the same cognitive failures that we observe in all decision-makers,[15] but also on the

14. This is not necessarily to say that adversarial epistemology is a particularly reliable way of determining the truth, whether in public debate (see Frederick Schauer, *Free Speech: A Philosophical Enquiry* 15–34 [1982]) or even in the courtroom (see Frank, *supra* note 9, at 80–81; Leon Green, *Judge and Jury* [1930]; David Luban, *Lawyers and Justice* 68–92 [1988]; John H. Langbein, "The German Advantage in Civil Procedure," 52 *U. Chi. L. Rev.* 823 [1985]).

15. A large, growing, and highly valuable body of research focuses on the cognitive failings of juries, of judges as fact-finders, and of judges as interpreters and appliers of law. As to juries, for which the literature is by far the largest, see, e.g., Dennis J. Devine et al., "Deliberation Quality: A Preliminary Examination in

fact that even some of the law's characteristic methods are potentially more flawed than the law has traditionally assumed. Eyewitness testimony, for example, is far less reliable than many people have traditionally thought,[16] and even reliable scientific methods such as DNA identification are subject to the imperfections of the human beings whose job it is to administer the tests and analyze the results.[17]

Legal fact-finding is not only subject to the myriad problems of an adversarial approach to locating the truth, but is also framed by the odd set of rules that are called the law of evidence. Space does not permit providing here even a brief summary of the substance of evidence law, but it is nevertheless important to highlight its peculiar assumptions. In part because of the special needs of the adversary system, in part because of a

---

Criminal Juries," 4 *J. Empirical Legal Stud.* 273 (2007); R. Hastie, D. A. Schadke, & J. W. Payne, "A Study of Juror and Jury Judgments in Civil Cases: Deciding Liability for Punitive Damages," 22 *L. & Human Behavior* 287 (1998); R. J. MacCoun & N. L. Kerr, "Asymmetric Influence in Mock Jury Deliberations: Jurors' Bias for Leniency," 54 *J. Personality & Social Psych.* 21 (1988). On judges as fact-finders, see Paul H. Robinson & Barbara A. Spellman, "Sentencing Decisions: Matching the Decisionmaker to the Decision Nature," 105 *Colum. L. Rev.* 1124 (2005); Barbara A. Spellman, "On the Supposed Expertise of Judges in Evaluating Evidence," 155 *U. Penn. L. Rev. PENNumbra* (2006); Andrew J. Wistrich, Chris Guthrie, & Jeffrey J. Rachlinski, "Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding," 153 *U. Pa. L. Rev.* 1251 (2005). On judges and the law, see Chris Guthrie, Jeffrey J. Rachlinski, & Andrew J. Wistrich, "Inside the Judicial Mind," 86 *Cornell L. Rev.* 777 (2001); Frederick Schauer, "Do Cases Make Bad Law?," 73 *U. Chi. L. Rev.* 883 (2006); Dan Simon, "A Third View of the Black Box: Cognitive Coherence in Legal Decision Making," 71 *U. Chi. L. Rev.* 511 (2004); Dan Simon, "Freedom and Constraint in Adjudication: A Look Through the Lens of Cognitive Psychology," 67 *Brooklyn L. Rev.* 1097 (2002); Dan Simon, "A Psychological Model of Judicial Reasoning," 30 *Rutg. L.J.* 1 (1998).

16. See, e.g., Elizabeth F. Loftus & James Doyle, *Eyewitness Testimony: Civil and Criminal* (3d ed., 1997); Elizabeth F. Loftus & Edith Green, "Warning: Even Memory for Faces May Be Contagious," 4 *L. & Human Behavior* 323 (1980); Gary L. Wells & Elizabeth F. Loftus, "The Malleability of Eyewitness Confidence: Co-Witness and Perseverance Effects," 79 *J. Applied Psych.* 714 (1994).

17. See, e.g., Brandon L. Garrett, "Judging Innocence," 108 *Colum. L. Rev.* 55, 63, 84 n.109 (2008); Edward J. Imwinkelried, "The Debate in the DNA Cases over the Foundation for the Admission of Scientific Evidence: The Importance of Human Error as a Cause of Forensic Misanalysis," 69 *Wash. U.L.Q.* 19 (1991); William C. Thompson & Simon Ford, "DNA Typing: Acceptance and Weight of the New Genetic Identification Tests," 75 *Va. L. Rev.* 45, 66–67 (1989).

substantive concern for the rights of criminal defendants, and in very large part because finding the facts has traditionally been the province of a jury with no specialized training either in law or in factual analysis, a body of law developed whose principal function has been to keep even relevant evidence away from a frequently distrusted jury. For fear that jurors would make too much of some evidence and too little of other, the law of evidence has a host of exclusionary rules that often seem strange. Although we often give some weight to what people hear other people say, for example, the law has traditionally prevented jurors from taking such hearsay evidence into account. And despite the fact that we commonly think that what someone has done in the past might help us determine whether they have done something similar now, much of this evidence of "bad character," "prior bad acts," or even previous convictions for the same type of crime is excluded at the typical trial.

The exclusionary rules of the law of evidence generated no small amount of ire in our old friend Jeremy Bentham, who would pretty much have eliminated all of the rules of evidence in favor of what he called the "natural" (as opposed to "technical") system, which has now come to be known as a system of Free Proof.[18] Under a natural or Free Proof approach, one that Bentham thought not that different from what ordinary people do in their daily lives, evidence is not excluded at the outset by rules that exclude entire categories of evidence, such as hearsay and prior criminal convictions. Rather, virtually all relevant evidence is admitted and then sifted, weighed, and evaluated in light of other evidence in order to give each piece of evidence the weight to which it is entitled. Some evidence will seem unreliable and will be discarded, while other pieces of evidence will be given a bit of weight but discounted. The basic point is that when we are trying to find out what happened, we do not set up a system that will keep potentially relevant evidence from our fact-finding process just because it fits some category of imperfect evidence.

In objecting to a system of factual determination largely structured around a series of what Bentham thought were artificial and categorical exclusions, Bentham was joined then, and even more so since, by many others, including not a few philosophers whose concern is epistemology.[19] And in important respects Bentham and his allies have been carry-

18. Jeremy Bentham, *Rationale of Judicial Evidence* (1827).

19. E.g., Alvin I. Goldman, *Knowledge in a Social World* (1992); Larry Laudan, *Truth, Error, and the Criminal Law: An Essay in Legal Epistemology*

ing the day. Especially with the decline in the importance of the jury—juries have for all practical purposes disappeared throughout the common-law world in civil cases, except in the United States, where the Seventh Amendment and its state constitutional counterparts have rescued the civil jury from oblivion—the formal rules of evidence have been consistently relaxed. Judges sitting without juries often treat the rules of evidence casually and appear to have little hesitancy in announcing that because there is no jury, most of the exclusionary rules of evidence will simply be ignored.[20] Moreover, exclusionary rules such as the hearsay rule and the original documents rule (often called the "best evidence" rule) are increasingly subject to a host of exceptions, and various other exclusionary rules have been officially eliminated or unofficially ignored. We may still be a long way from Bentham's preferred system of Free Proof, but we are also a long way from the highly rule-based and largely exclusionary system that generated Bentham's anger in the first place.[21]

The somewhat peculiar institution of the adversary system, the even more peculiar institution of the jury, and the especially peculiar idea of rigid exclusionary rules of evidence are all of a piece with the larger themes of this book. Law does things differently, for better or for worse, and the difference between how law determines the facts of a case and how other decision-makers find out about the world around them is consistent with law's use of the unusual devices that we have considered earlier, such as stare decisis and a commitment to the sometimes suboptimal control of rules. As with some of the other tools of legal reasoning, law's methods of fact-finding are not totally unique to law. Adversary determinations can be seen in other decision-making environments, as can even exclusionary evidentiary rules. But the fact that law's methods are not unique to law does not mean that law is no different from anything else,

(2006); Susan Haack, "Epistemology Legalized: Or, Truth, Justice, and the American Way," 49 *Am. J. Jurisp.* 43 (2004).

20. See Frederick Schauer, "On the Supposed Jury-Dependence of Evidence Law," 155 *U. Pa. L. Rev.* 165 (2006).

21. This is especially obvious once we realize that Bentham allocated a large part of his considerable capacity for outrage to the rules of competency—the rules that made it impossible for most women, most minors, most convicted felons, and most of the litigants themselves to be witnesses at trial. In large part the rules of competency have been eliminated in the United States, and although there are still things to which witnesses may not testify, there are few blanket exclusions based on a witness's status or personal characteristics.

and thus it should come as no surprise that when it comes to facts as well as to law, it is a mistake to fail to recognize how decision-making within the legal system is, at the very least, a little bit different.

Law's commitment to its own methods of factual determination is reflected even in the structure of the legal system's decision-making about questions of law. Because law is committed to the distinction between the trier of fact and the determiner of law, findings of fact are typically separated from conclusions of law when the same trial judge takes on both tasks. More importantly, findings of fact are typically, except in the most egregious of instances, treated as sacrosanct in the appellate process. It is only slight hyperbole to say that if a jury were to find that the moon was made of green cheese, an appellate court ruling on a legal question about the moon or about green cheese would be expected to take the jury's false conclusion as true. We have seen throughout this book that questions of jurisdiction in the broad sense—what is important is not only what is decided but who has the authority to decide it—are a ubiquitous feature of legal analysis. And jurisdiction in this broad sense has much to do with determining the facts. It is the job of a jury, or the trial judge acting as the trier of fact, to determine the facts. Even if the facts which that trier of fact has found seem wrong to an appellate court, the fact- finder's seemingly erroneous factual conclusions must nevertheless be taken as true. This will seem odd at times, but it may be part of a larger and pervasive characteristic of law itself. What makes law different is that legal decision-making, whether about law or about fact, differs from the simple mandate to judges and other legal decision-makers that they simply "do the right thing." Just as rule-based and precedent-based decision-making often requires legal decision-makers to do something other than the right thing, the strong obligation to accept the fact-finder's factual finding sometimes produces the same kind of suboptimality. To some this may be a bad thing, but to others it is simply part of law's commitment to achieve the greatest good in the aggregate, even if that requires giving up the aspiration to do what particular decision-makers think is the right thing in particular cases.

11.3 Facts and the Appellate Process

At the beginning of this chapter we made reference to *Brown v. Board of Education* and the way in which the Supreme Court in that case relied on psychological studies showing that segregated African-American chil-

dren suffered educationally from their exclusion from the schools attended by whites. This aspect of *Brown* generated much controversy, and for several reasons.[22] First, it was not clear that the conclusions of the studies were necessarily correct. Other psychologists had come to different conclusions, and there was a worry about whether litigation was the best way to resolve disputed questions of scientific fact.

More importantly, the Supreme Court appeared to make its *own* evaluation of the question rather than simply relying on the trial court's resolution of the factual issues. It may be, as we discussed in the previous section, that litigation and the adversary system are not the best ways to resolve some or all factual questions, but that is the way of the law, and it has been for centuries. Not so, however, for appellate courts, and for just as many centuries the assumption has been that determining the facts is for the trial court and evaluating the trial court's handling of the law is for appellate courts. If, barring blatant error or prejudice, the trial court, whether by judge or jury, has found *x*, then *x* must be accepted as true. The lawyer who tries to argue before an appellate court that *x* is not true will quickly find himself on the wrong end of a scolding from the court for trying to use the appellate courts as the forum for relitigating factual determinations that appellate courts are expected to take as final.

This is a nice model, but it may not capture fully the extent to which appellate courts are themselves engaged in determining questions of fact. *Brown v. Board of Education* may have highlighted the issue because of the prominence of the case and because the Supreme Court's footnote reference to the relevant studies made it quite obvious what was going on, but *Brown* turns out not to be all that unusual.

Consider, for example, *New York Times Co. v. Sullivan,*[23] the 1964 case in which the Supreme Court constitutionalized and revolutionalized the law of defamation throughout the United States by holding that public officials (and, later, public figures)[24] could succeed in a libel case only if they could prove with convincing clarity not only that what had been said about them was false, but also that it had been said with knowledge

22. For descriptions of the controversy, see John Monahan & Laurens Walker, *Social Science in Law: Cases and Materials* 84–99, 106 (1985); Mody Sanjay, Note, "Brown Footnote Eleven in Historical Context: Social Science and the Supreme Court's Quest for Legitimacy," 54 *Stan. L. Rev.* 703 (2002).

23. 376 U.S. 254 (1964).

24. 388 U.S. 130 (1967).

of its falsity. In other words, plaintiffs had to prove not only intentional publication, but also intentional falsity. This was a dramatic change in the common law, and the Court justified the change by concluding that criticism of public officials would be "uninhibited, robust, and wide open" only if publishers were relieved from liability for even their negligent untruths. This empirical conclusion may well be true, but it is not at all clear how the Supreme Court knew that it was true. Some might think the proposition self-evident, but once we realize that uninhibited, robust, and wide-open press criticism of public officials exists in countries with far more restrictive defamation doctrines (Australia, for example) than exist in the United States, it becomes less clear that the factual proposition that provided the linchpin for the Court's conclusion was as self-evident as the Supreme Court thought it was. Nevertheless, this factual proposition about press behavior was an essential element of the Court's conclusion. Whether the Court was right (probably) or wrong (possibly) in its assessment is not the important issue here. The important issue is the question of the extent to which a potentially contestable factual proposition—and not one that had been part of the trial proceedings at all—turned out to be central to the Court's legal conclusion. Perhaps because the Court in *Sullivan* did not cite to nonlegal sources, as it did in *Brown,* the factual link in the Court's argumentative chain was less obvious, but no less than in *Brown,* the Court in *Sullivan* rested its conclusion on a contestable factual proposition as to which there had been no finding of fact below.

Much the same was true, and with a level of controversy closer to *Brown* than *Sullivan,* with respect to the Supreme Court's conclusion in *Mapp v. Ohio*[25] that illegally obtained evidence could not be used at a subsequent criminal trial regardless of its reliability. If an illegal search, for example, actually did lead to the discovery of drugs plainly belonging to the defendant, after *Mapp* those drugs would be excluded as evidence from the trial. The Court based its conclusion on the belief that an exclusionary rule would deter the police from engaging in unconstitutional behavior, but once again this is an empirical conclusion with which reasonable people can and did disagree.[26] Maybe the police do not much

25. 367 U.S. 643 (1961).

26. See Yale Kamisar, "Does (Did) (Should) the Exclusionary Rule Rest on a 'Principled Basis' Rather than an 'Empirical Proposition'?," 16 *Creighton L. Rev.* 565 (1983).

worry about what goes on at trials and are concerned mainly with apprehending perpetrators, or maybe unconstitutional police behavior would be deterred more by threats of internal sanctions against police officers personally. But whatever the fact of the matter, the important point is that once again the Court's route to a new legal rule was one that took it through the making of a factual determination as to which most of the evidence appeared to come from the Justices' own beliefs, experiences, hunches, intuitions, and armchair sociology.

Finally, consider the plurality opinion in *Bush v. Gore*.[27] In concluding that the Supreme Court of Florida had erred in rejecting George W. Bush's equal protection challenge to the Florida vote-counting procedure, the Supreme Court found it important that the casting of invalid ballots was not in fact a historically infrequent occurrence and that many invalid presidential ballots had been cast in most previous elections. Whether this should or should not have been important to the Court is not pertinent to our discussion of factual determination, but what is germane here is the fact that on this factual proposition there was again virtually no finding below, and the Court reached its conclusion, as discussed at somewhat greater length in Chapter 4, on the basis of several newspaper articles, presumably located by the Justices (or, more likely, their law clerks) through a Nexis search.

Brown, Sullivan, Mapp, and *Bush v. Gore* are all constitutional cases in the Supreme Court, but it would be a mistake to think of the phenomenon as restricted to constitutional law. When Holmes insisted that the "life of the law has not been logic; it has been experience,"[28] he made it clear that appellate judges, in both following and creating "the path of the law," would have to rely on empirical and factual determinations, a phenomenon extensively theorized almost a century later by Melvin Eisenberg in showing how reliance on what he called "social propositions" is an essential element in common-law reasoning.[29] *Henningsen,* for example, was premised on a view about the nature of consumer transactions that came largely from the Court's own impressions, and when the New York Court of Appeals in *Adams v. New Jersey Steamboat Co.*[30] concluded that a stateroom on a steamboat was more like an inn

27. 531 U.S. 98 (2000).
28. O. W. Holmes, Jr., *The Common Law* 1 (1881).
29. Melvin A. Eisenberg, *The Nature of the Common Law* (1988).
30. 45 N.E. 369 (N.Y. 1896). The case has become a staple of discussions

than like a sleeping compartment on a train, it relied heavily on what *it* believed about contested factual propositions regarding the normal uses and expectations with respect to steamboats, inns, and trains.

But if social propositions—which are conclusions of fact, albeit about general social conditions and not about the particular facts of the particular case—play such a large role in appellate decision-making, then how is an appellate court to find out about the facts necessary to reach such conclusions? This has been a recurring issue, and it is one that Justice Breyer of the Supreme Court, more than anyone, has brought to the forefront of legal debate, especially in the context of questions about science.[31] Justice Breyer himself is hardly reticent about going far beyond the record to make factual determinations he believes necessary to resolve the cases before him, and his dissenting opinion in *Lopez v. United States*[32] is replete with scores of references to economic, sociological, and political materials directed at the question of whether the possession of weapons in the public schools has an effect on interstate commerce. Similarly, Justice Breyer's (again dissenting) opinion in the high school affirmative-action case of *Parents Involved in Community Schools v. Seattle School District No.1*[33] drew heavily not only on his own research about the factual background of *that* case[34] but also on far-reaching em-

about the use and misuse of analogy in legal reasoning. See Richard Posner, *How Judges Think* 169–70 (2008); Lloyd Weinreb, *Legal Reason: The Use of Analogy in Legal Argument* (2005); Scott Brewer, "Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy," 109 *Harv. L. Rev.* 923 (1996). And see Chapter 5, *supra*.

31. See Stephen Breyer, "Introduction," *in Reference Manual on Scientific Evidence* (2d ed., 2000); Stephen Breyer, "The Interdependence of Science and Law," an address at the American Association for the Advancement of Science Annual Meeting and Science Innovation Exposition, Feb. 16, 1998, available at www.aaas.org/meetings/scope/breyer.htm *and in* 280 *Science* 537 (1998).

32. 514 U.S. 549 (1995).

33. 127 S. Ct. 2738 (2007).

34. This practice is both unusual and controversial. There is a traditional distinction between legislative facts and adjudicative facts, the former being the facts necessary to make or support a legal rule and the latter being the facts of a particular controversy or rule application. This is a distinction that is of some import with respect to questions of due process and the right to a hearing, because it is accepted that individuals have due-process rights to notice and hearing with respect to adjudicative facts that will produce adverse consequences to them, but not to legislative facts that will produce adverse consequences to them only in respect to which they are members of a class adversely affected by the legislative rule. See Bi-

pirical inquiry about the history, sociology, psychology, and politics of student assignment in American public schools. For Justice Breyer, managing appellate factual and scientific inquiry has been for some time a pressing question, but it may be that we are not especially close to an answer.

If what Holmes called experience and what Eisenberg calls social propositions are a pervasive and indeed necessary component of common-law *legal* decision-making, then where are appellate judges (or trial judges making legal and not factual determinations) to get the information necessary to reach their factual and empirical conclusions? Justice Breyer's opinions, the social science data in *Brown v. Board of Education,* and the newspaper reports in Justice Kennedy's opinion in *Bush v. Gore* have the virtue of displaying the sources on which the Justices were relying, but *Sullivan, Mapp, Henningsen,* and *Adams* are for just that reason more important. Even when a judge does not cite to nonlegal academic journals or newspapers or anything else, she is still, although less obviously, relying on sources of information that are importantly factual, that may very well be contested, and that wind up being part of the law in a somewhat under-the-table manner, even apart from the way in which such propositions may produce adverse consequences for one of the parties without that party having much or any opportunity to challenge those propositions by the normal adversarial processes, including but not limited to cross-examination.[35]

Metallic Investment Co. v. State Board of Equalization of Colorado, 239 U.S. 441 (1915); Londoner v. Denver, 210 U.S. 373 (1908). Related to but somewhat distinct from the notice and hearing question is the question of whether an appellate court should investigate adjudicative facts not found below or even reevaluate findings about adjudicative facts made by the trier of fact. The answer to these questions has traditionally been a clear no, and the extent to which judges may or should do research about the facts of *this* case, outside of the formal adversary processes of trial with the rules of evidence, is more controversial and far less accepted than the idea that judges can and must do their own research with respect to legislative facts.

35. It is worth noting here that traditional English practice, now softening somewhat, prohibits judges from doing their own research outside of the presence of counsel, even as to the law. Cases and statutes not cited and argued by the parties or discussed in open court might as well not even exist. This practice may seem unusual to Americans, but it is part of a tradition of *orality* that stresses that nothing should happen in litigation that is not transparent and available for argument by all parties. See Delmar Karlen, *Appellate Courts in the United States and En-*

To the extent that contested factual propositions are increasingly "flagged" by citation to nonlegal materials,[36] the issue is becoming more patent, but the deeper question is not about the materials that judges consult or cite in order to make legal, as opposed to adjudicative factual, determinations. Even with no explicit consultation and no citation, judges making law, and often just applying law, must rely on empirical conclusions that lurk scarcely beneath the surface. When the existence of such conclusions is not announced by means of, for example, citation to newspapers or nonlegal books or periodicals, there is a risk that we may ignore the extent to which such conclusions are open to contest, which may well be a function of what the judges think of as common knowledge but which others may wish to challenge. Citation to materials outside of the traditional legal canon may be for some a source of alarm, but it may as well be a way in which the empirical propositions that are necessarily a part of all judicial lawmaking and much judicial law application can be subject to argument and challenge, rather than simply being clothed in the disguise of common knowledge or what judges believe, not always correctly, and not necessarily unrelated to their own backgrounds, to be the common wisdom of humanity.

---

gland (1964); Suzanne Ehrenberg, "Embracing the Writing-Centered Legal Process," 89 *Iowa L. Rev.* 1159 (2004); Robert J. Martineau, "The Value of Appellate Oral Argument: A Challenge to the Conventional Wisdom," 72 *Iowa L. Rev.* 1 (1986); Richard A. Posner, "Judicial Autonomy in a Political Environment," 38 *Ariz. St. L. Rev.* 1, 10 (2006). The tradition can produce extraordinarily lengthy appellate arguments (which often take days, rather than the typical thirty minutes per side in American appellate courts), because the expectation is that everything on which judges rely is open to argument by both sides, and it produces a tendency to rely on only a narrow range of widely accepted legal sources. But it does forestall most objections that judges are making decisions based on information not known to or argued by both parties.

36. See the discussion in section 4.4, *supra*. See also Frederick Schauer & Virginia J. Wise, "Non-Legal Information and the Delegalization of Law," 29 *J. Legal Stud.* 495 (2000).

12

THE BURDEN OF PROOF AND ITS COUSINS

12.1 The Burden of Proof

Law navigates through a fog of uncertainty. In attempting to figure out what happened in the past, the legal system must deal with faulty recollections, lost documents, missing witnesses, inattentive jurors, and countless other impediments to knowing with very much confidence what actually took place months or years earlier. And even in trying to assess what law should apply to the facts so imperfectly perceived, lawyers and judges face a world of conflicting precedents, vague statutes, substantive disagreement, and a host of additional obstacles to being able to determine just what it is that the law requires.

Not only does law do its work under conditions of uncertainty, but the legal system is also a complex one in which the separation of powers, in the nontechnical (or at least non-constitutional) sense of that term, is a dominant consideration. Appellate courts must take account of the tasks assigned to trial judges and trial juries, federal courts are constitutionally required to be cognizant of the proper domains of state courts and vice versa, and courts engaged in judicial review of legislative or administrative action need to recognize the distinct responsibilities of legislatures and the specialized expertise of agencies. The question before a judge is rarely simply the question of what is right but is almost always imbued with the issue of whose job it is to determine what is right. *Jurisdiction* may be an important component of the law of civil procedure, but it is even more central to the very structure and idea of law itself, for law is characterized by its concern not only for what is decided but for whose job it is to decide it.

Under such conditions of uncertainty and divided responsibilities, the burden of proof, and its companion concepts of deference and presump-

tion, play a huge role. These concepts tell us just how sure the legal system needs to be in order to reach a particular conclusion, and indirectly tell the system what is to happen in the event that it is not sufficiently sure. And by specifying how confident the law must be in order to produce a particular legal outcome, the burden of proof, especially, reflects deeper substantive and not just procedural values that vary depending on the consequences of that outcome.

The most familiar operation of the burden of proof is in criminal cases, where we encounter the well-known requirement that the prosecution must prove its case beyond a reasonable doubt. But what does this mean, and what values does it reflect? In order to address these questions, and also for purposes of simplicity and clarity, let us assign a rough numerical probability to the burden of proof.[1] We can thus start with the premise, say, that the prosecution must prove its case such that the jury, to convict, must be 95 percent certain of the defendant's guilt. This percentage is in some sense arbitrary, or at the very least a rough estimate, but it captures the basic idea that the beyond-a-reasonable-doubt standard requires that the jury or judge have a very high degree of confidence in the defendant's guilt in order to convict. If they are not 95 percent sure, then they must render a verdict of "not guilty."

An important feature of such a high burden of proof is that if taken seriously—and there is scant reason to believe that it is not in most criminal cases—the 95 percent "beyond a reasonable doubt" standard can be predicted systematically to let a large number of guilty people go unpunished. Suppose we have ten defendants, and suppose further that each of them is 90 percent, but only 90 percent, likely to have committed a crime. If the jury is doing its job properly, then all ten defendants will go free, because in none of these cases will the prosecution have met the requisite burden of proof of 95 percent confidence in the defendant's guilt. But because each of the ten defendants is 90 percent likely actually to have committed the crime with which he has been charged, we can expect that nine of those ten acquitted defendants will actually be guilty of that for which they are on trial, in spite of which they are now going free.

In letting so many of the guilty go free, the "proof beyond a reasonable doubt" standard may seem like a bad idea. And of course it certainly does seem so to legions of politicians who claim that the legal system

1. See Frederick Mosteller & Cleo Youtz, "Quantifying Probabilistic Assessments," 5 *Statistical Sci.* 2 (1990).

coddles criminals and lets far too many of them escape their just deserts. But William Blackstone has given us the canonical justification for the law's approach. "The law holds, that it is better that ten guilty persons escape," he wrote in 1769, "than that one innocent suffer."[2] What Blackstone understood was that if the legal system employed a lower burden of proof in criminal cases, say 60 percent, then we could expect that for every ten defendants for whom the probability of guilt was exactly 60 percent and no more, all ten would go to prison, but only six of them would be guilty.[3] No guilty people would go free, but four innocents would be punished. And that, for Blackstone and the legal system he celebrated, was just too much. It is unfortunate when guilty people go free, Blackstone thought then and we still think now, but it is far worse when innocent people are condemned. Consequently, the legal system calibrates the burden of proof in such a way that the law can serve the social interest in convicting the guilty while keeping the number of innocents that it punishes very low.[4]

Blackstone's solution—our solution—to the problem of uncertainty is far from perfect. If we really wanted never, ever to convict the innocent, we would set the burden of proof astronomically high—absolute certainty, or 99.99 percent determined by three consecutive juries, or something of that sort—and we would have solved most of the problem of convicting the innocent. We would have done so, however, at the cost of convicting far too few of the guilty. The standard that common-law legal systems have chosen is a balance, but a balance heavily weighted in favor of the social judgment that convicting the innocent is a great deal worse

2. 4 William Blackstone, *Commentaries* *358. Blackstone was not the first to express the idea, John Fortescue having written in 1471 that "I should, indeed, prefer twenty guilty men to escape through mercy, than one innocent to be condemned unjustly." Sir John Fortescue, *De Laudibus Legum Angliae* 65 (Dr. Chrimes ed., 1942) (1471). And in 1824 it was said that "it is better that ninety-nine . . . offenders shall escape than that one innocent man be condemned." Thomas Starkie, *Evidence* 756 (1824). Fortescue, Blackstone, and Starkie were all expressing the same principle, but the differences among ten to one, twenty to one, and ninety-nine to one reflect different views about the comparative harms of the two types of errors. See generally Alexander Volokh, "n Guilty Men," 146 *U. Pa. L. Rev.* 173 (1997).

3. For the statistically inclined, it is worth noting that Blackstone's ten-to-one ratio reflects an underlying burden of proof of 0.91.

4. See John Kaplan, "Decision Theory and the Factfinding Process," 20 *Stan. L. Rev.* 1075 (1968).

than freeing the guilty, but also that avoiding convicting the innocent is not the only social value there is.

Now let us compare to the criminal standard the standard that is typically used in civil cases. Here the plaintiff will prevail simply if she proves her case by a preponderance of the evidence; the British call this the "balance of probabilities." If the plaintiff proves her case by a bare preponderance—50.000001 percent, say—she will win, but if she falls below that, she will not.

This is a very different standard from that used in criminal cases, and that is because the values at stake are very different. Yes, it is regrettable when a defendant who is not actually at fault, say, is held liable in a tort action and has to pay damages. But in the eyes of the law this is no more regrettable than someone who is injured through the fault of another not being able to recover because the defendant has been mistakenly found not to be liable. Unlike in the criminal case, where we deem the mistake of imprisoning the innocent far worse than the mistake of freeing the guilty, in the civil context we consider the mistake of a wrongly uncompensated plaintiff to be no less serious than that of a defendant wrongfully held liable. The two errors being equal, the burden of proof selected reflects this underlying equality of values.[5]

Although proof beyond a reasonable doubt and proof by a preponderance of the evidence are the best-known and most widely used of law's burdens of proof, there are in fact many others. People may not be committed to mental institutions, for example, unless they have been proved by "clear and convincing" evidence to be dangerous to themselves or others, and it is generally accepted that this standard is somewhat higher than proof by a preponderance of the evidence, yet somewhat lower than proof beyond a reasonable doubt.[6] The police may obtain a search warrant only if they can establish "probable cause"—the exact words used in the Fourth Amendment—to believe that the search will yield evidence of the crime they are investigating. And although it

5. In fact the matter is somewhat more complicated than this, because the plaintiff must prove *each* element of a cause of action by a preponderance of the evidence. In other words, the plaintiff must prove *all* of the elements by a preponderance of the evidence to prevail, but the defendant will prevail if she proves only one element to such a standard. The effect of this is to place a burden on the plaintiff to actually prevail that is somewhat higher than a bare preponderance. See Ronald J. Allen, "The Nature of Juridical Proof," 13 *Cardozo L. Rev.* 373 (1991).

6. See Addington v. Texas, 441 U.S. 418, 432–33 (1979).

may be even harder to translate this standard into a number, it is well understood that this places a higher burden of proof on the police than the "reasonable suspicion" standard applicable when the police would stop and question a person but not seek to search his residence.[7]

Although the burden of proof is most visible with respect to the final determination of the outcome of a civil or criminal case, it is even more prevalent when used in conjunction with the various decisions that have to be made prior to final judgment, or when the burden is assigned to one party or another for various elements of a case. For the numerous legal and factual questions that are heard by judges leading up to the trial itself, for the procedural and evidentiary issues that arise during a trial, and for numerous different substantive issues, one party or another will have what is called the burden of persuasion. In criminal cases, for example, the defendant will have the burden of persuasion on defenses such as insanity, alibi, and self-defense. And in tort actions for trespass or battery, a defendant who claims consent must persuade the trier of fact that such consent existed, as opposed to the plaintiff having to show that there was no consent. And on most procedural issues arising at or before a trial, it will be the moving party who bears the burden of persuasion.

The burden of persuasion needs to be distinguished from the burden of production, sometimes called the burden of going forward with the evidence. Largely because one party or another is more likely to have the relevant information, the law will often assign to a particular party the burden of producing sufficient facts to put some matter at issue, and this is what is called the burden of production. Sometimes the party with the burden of production will also have the burden of persuasion, but often the two will be distinct. In contract cases to enforce a contract, for example, it is common for the defendant to have the burden of coming forward with sufficient evidence to put into issue the possibility of mutual mistake, or impossibility, or release, but once the defendant has met its burden of production, the burden of persuasion still rests with the contract-enforcing plaintiff to show the lack of release, or the lack of mutual mistake, or the lack of impossibility.[8]

Whether it is the burden of proof for the case as a whole, or the bur-

7. See Alabama v. White, 496 U.S. 325 (1990); Terry v. Ohio, 392 U.S. 1 (1968).

8. See Robert E. Scott & George C. Triantis, "Anticipating Litigation in Contract Design," 115 *Yale L.J.* 814 (2006).

den of persuasion on a particular issue, or the burden of production of evidence, it is important to bear in mind an important theme: these devices are all called "burdens" for a reason. From the point of view of an advocate, these burdens, like most other burdens, are almost by definition something that you do not want to have. The burden of proof is, for entire cases or for components parts of them, something a lawyer wishes her opponent to have, and often much of legal argument consists of attempting to persuade a judge (or jury) that your opponent has the burden of proof. It is an exaggeration to say that once you have saddled your opponent with the burden of proof on some issue, your job as an advocate is done. As we know in the criminal law, even high burdens of proof are met with great frequency. Still, the advocate who can persuade the decision-maker that the other side has the burden of proof has gone a long way toward success, and as a result, arguments about the burden of proof are as frequent in legal argument as arguments about the law and the facts themselves.

12.2 Presumptions

As we have just seen, allocating the burdens of proof, persuasion, and production is one way in which the law deals with decision-making under factual uncertainty. But it is not the only way. Often the law uses *presumptions,* which are the starting point from which factual inquiry takes off and which specify what facts will be taken to be true if the party with the burden of proof does not satisfy the burden of proof and thus overcome the presumption. Presumptions also specify what additional facts are taken to have been proved when one of the parties proves some particular fact.

In some respects, a presumption is prior to but closely related to the burden of proof, because it is a presumption that typically establishes who *has* the burden of proof, independent of just how much that party needs to demonstrate in order to satisfy it. And it is the presumption that specifies what is to happen when the party with the burden of proof on some issue does not meet it. When we say that a defendant in a criminal case is presumed innocent, for example, we are also saying that the prosecution has the burden of proof, with the failure to satisfy it resulting in the defendant's acquittal. Similarly, when it is held that a will is presumed to have been properly executed,[9] the effect is that a party claiming fraud,

9. See Slack v. Truitt, 791 A.2d 129 (Md. 2002).

duress, or lack of capacity must demonstrate that one of these defeating conditions has actually occurred, and in the absence of such a demonstration, the court will accept the validity of the will, even though the party claiming that the will was validly executed has proved nothing at all. In the same way, the common law has traditionally had a presumption of legitimacy, such that a child is presumed to be the legitimate offspring of his parents, and the burden is on someone challenging legitimacy—in the context of a dispute about inheritance, for example—to show that the presumed fact of legitimacy is not in reality so.[10] And perhaps the most familiar presumption of all is res ipsa loquitur ("the thing speaks for itself"), the effect of which, for example, is to establish that that a surgical patient does not need to prove negligence when he emerges from the surgery with a sponge or a surgical instrument in his abdomen, because negligence will be presumed from the very nature of the event.[11]

Quite often presumptions do not operate on the basis of a clean slate but require a party to prove something, by virtue of which something else will be presumed. A good example is the long-standing rule in most jurisdictions that proof that a letter has been mailed will create a presumption that it has been received.[12] The presumption of innocence may be unusual, therefore, because it is a background assumption of the legal system. But in the normal course of things, the legal system will require a party to prove *A*, and by virtue of *A* having been proved, *B* will be presumed. A beneficiary under an accidental death insurance policy may have to prove that the decedent died from something other than natural causes, but under what is called the "presumption against suicide," it will ordinarily be presumed that if the claimant so proves, then accident rather than suicide will be presumed. Consequently, the burden is on the insurance company to come forward and prove suicide, failing which the death will be presumed to have been caused by an accident.[13]

It is common to distinguish between rebuttable and irrebuttable

10. See, e.g., John M. v. Paula T., 571 A.2d 1380 (Pa. 1990); Joseph Cullen Ayer, Jr., "Legitimacy and Marriage," 16 *Harv. L. Rev.* 22 (1902).

11. See, e.g., Fla. Stat. § 766.102 (4) (2004); Armstrong v. Wallace, 47 P.2d 740 (Cal. 1939); Fink v. Bonham, 183 N.E.312 (Ind. 1932).

12. See, e.g., Hagner v. United States, 285 U.S. 427 (1932); Santana Gonzalez v. Attorney General, 506 F.3d 274, 278–79 (3rd Cir. 2007); Holt v. Mississippi Employment Security Commission, 724 So. 2d 466, 471 (Miss. Ct. App. 1998).

13. See Davison v. National Life & Accident Insurance Co., 126 S.E.2d 811 (Ga. 1962); Schelberger v. Eastern Savings Bank, 458 N.E.2d 1225 (N.Y. 1983).

(sometimes called "conclusive") presumptions. And it is the former that are most closely aligned with the burden of proof. When a defendant is presumed innocent, when a sponge is presumed to have been left in a patient's abdomen by virtue of the negligence of the nurse or physician, when a child is presumed legitimate, and when a will is presumed to have been validly executed, it remains open to the other party to prove, by the burden of proof that has been assigned to it, that the state of affairs embodied in the presumption is not in reality so in this case. The existence of a presumption of validity for a will, for example, does not prevent another party from attempting to prove that the will was made under duress, or as a result of fraud, or by someone not in full possession of her mental faculties at the time of execution. And if it is so proved, then the presumption can be said to have been *rebutted*. In this respect, the presumption is in effect an allocation of the burden of coming forward, a specification of the burden of proof to be borne by the party who has the burden of coming forward, and a statement of the facts that will be understood as having existed if the party with the burden either does not come forward or comes forward and does not meet his burden of proof.

Irrebuttable presumptions are quite different. In structure they are similar to rebuttable ones, because an irrebuttable presumption also specifies the state of affairs that the law will assume to exist (even if it does not). But because the law does not allow the other side to challenge the conclusion of an irrebuttable presumption, it remains unconnected with procedural issues such as a burden of coming forward or a burden of proof. Indeed, there is little difference, except in the form of expression, between an irrebuttable presumption and what we would simply call a rule of law. The state of Florida, for example, has, like all other states, a law prohibiting the sale and distribution of illegal drugs, including cocaine. The law imposes a much higher penalty on those who are trafficking in drugs than on those who merely possess them, and then goes on to say that anyone possessing 28 grams or more of cocaine is irrebuttably presumed to be in the business of and guilty of trafficking illegal drugs.[14] In other words, if you are found in possession of more than 28 grams of cocaine, you are presumed to be a cocaine dealer, whether you are or not. The burden of proof is not an issue, because the law simply does not permit you to try to rebut the presumption, even if it is totally false as to you.

14. Fla. Stat. §893.135(1)(b)(1) (Supp. 2002).

Such a presumption might seem unfair, and indeed this kind of presumption seemed very unfair to the Supreme Court for a few years back in the 1970s.[15] But as the Supreme Court came to realize upon further reflection,[16] an irrebuttable presumption is little different from any other legal rule. Yes, the possessor of 28 grams of cocaine might not be a drug dealer, but the person who drives at 70 miles per hour in a 55-miles-per-hour zone might not be driving unsafely, and the insider who buys and then sells shares in her own company in a less-than-six-months period may not be in possession of any inside information at all. But, as we saw in Chapter 2, rules do their work precisely by cutting off access to their background justifications, and few are surprised that exceeding the speed limit while driving very safely is still an invitation to a traffic ticket.

Once we see that rules have force even when they produce results that would not be produced by direct application of the rationales or background justifications, we can see irrebuttable presumptions in their proper light. If Florida, in order to address the problem of cocaine selling (the background justification), had simply prohibited possessing 28 grams or more of cocaine, the nondealing possessor of more than 28 grams of cocaine would have no better an argument than the safe driver who is exceeding the speed limit. Just as we would say to the safe driver that he has broken the law even if he is not within the class of people the law was designed to encompass, we can say the same thing to the nondealing drug possessor. An irrebuttable presumption is thus just a different way of characterizing an omnipresent feature of all legal rules, and there is no reason to suppose that there is anything deeply wrong with that.

The category of interest, therefore, is the category of rebuttable presumptions, for these are the ones that actually make a procedural difference. By allocating the burdens of proof and persuasion, and by specifying what we might think of as the factual "default," rebuttable presumptions are essential elements in the structuring of litigation, not only serving substantive legal goals, but also attempting to ensure that scarce legal and judicial resources are not wasted by forcing parties to spend time and money proving what is usually but not always true.

15. See Cleveland Board of Education v. LaFleur, 414 U.S. 632 (1974); United States Department of Agriculture v. Murry, 413 U.S. 508 (1973); Vlandis v. Kline, 412 U.S. 441 (1973).

16. Weinberger v. Salfi, 422 U.S. 749 (1975).

Most presumptions arise in the context of facts. But presumptions can be legal as well as factual. Consider the structure of American equal protection doctrine, for example. It is now well settled that state classifications based on race, religion, ethnicity, and national origin are presumptively unconstitutional and will be upheld only if the state demonstrates a *compelling interest* in using such a classification, and demonstrates as well that there is no less restrictive (of equality) alternative that it can use to achieve that interest.[17] Here the question is one not of fact but of law, but we are still talking about a presumption. The law is presumed unconstitutional, but the state may rebut that presumption by satisfying a heavy burden of justification. Conversely, when a statute draws a classification within this category of "suspect" classifications, it is presumed to be constitutionally permissible and will be invalidated only if the challenger meets *its* burden of proving that the classification is irrational.

It turns out that this kind of presumption not only pervades constitutional law but may usefully characterize much of the operation of American law generally. As we have seen in various places throughout this book, it is sometimes the case in American law that judges will set aside the literal or plainest indications of formal law in order to serve the law's purpose or in order to achieve larger goals of justice. This is a fair characterization of *Riggs v. Palmer,*[18] *Church of the Holy Trinity v. United States,*[19] and *United States v. Kirby,*[20] for example, and it is a frequent occurrence on the American legal landscape. But it is also the case that the formal law often prevails even when the outcomes it produces are somewhat unjust or in other ways somewhat suboptimal. That is how we might explain *TVA v. Hill,*[21] constitutional separation-of-powers cases like *Immigration and Naturalization Service v. Chadha*[22] and *Bowsher v. Synar,*[23] the large number of cases in which unworthy beneficiaries who

17. See Palmore v. Sidoti, 466 U.S. 429 (1984); Loving v. Virginia, 388 U.S. 1 (1967). The standard in fact originated in the notorious Korematsu v. United States, 323 U.S. 214 (1944), in which the Supreme Court proceeded to conclude that internment of Americans of Japanese origin during the Second World War satisfied the "compelling interest" standard.

18. 22 N.E. 188 (N.Y. 1889).

19. 143 U.S. 457 (1892).

20. 74 U.S. (7 Wall.) 482 (1868).

21. 437 U.S. 153 (1978).

22. 462 U.S. 919 (1983).

23. 478 U.S. 714 (1986).

were in some way responsible for the death of the testator were, unlike Elmer Palmer, allowed to inherit, and possibly even *United States v. Locke*.[24] And perhaps the idea of a legal or normative presumption is the best way of at least partially reconciling these two seemingly opposed lines of cases and decisions. It may be, that is, that the best characterization of much of American law is that the formal side of law—what the rules or the precedents *say*—will be *presumed* to control, but the outcome indicated by the formal law will not be the final outcome of the case if the party burdened by the formal outcome can prove or persuade the court that the result so indicated will be *highly* unjust or in some other way not simply wrong but very wrong. In adopting this approach to the effect of rules and precedents, American law has perhaps used the idea of a presumption as a way of reconciling the stability and predictability needs that are satisfied by a formal approach to law while recognizing that formal law cannot always produce the right answer and that sometimes a wrong answer will be so wrong that it would be irresponsible and reprehensible were the legal system unable to do anything about it.

12.3 Deference and the Allocation of Decision-Making Responsibility

In the same cluster of ideas in which we find burdens of proof and presumptions, we also find the idea of *deference*. Law, perhaps even more than other institutions, is very concerned with *jurisdiction,* in both the technical and the nontechnical senses. In the technical sense, the law worries about whether and when a court has the power to make a decision and enforce a judgment against certain individuals and other entitles. Questions about *personal jurisdiction* are about whether, for example, a court can hear and decide a case against an individual with few connections to the location of the court, or at times about whether a court can exercise power over people holding certain kinds of positions.[25] And sometimes the jurisdictional question is not so much about *who* a court

24. 471 U.S. 84 (1984).

25. Marbury v. Madison, 5 U.S. (1 Cranch) 137 (1803), has achieved its fame by virtue of having established the power of judicial review, but at the time it was most controversial and raised President Thomas Jefferson's hackles for its assertion of jurisdiction over officials of the executive branch of government.

can reach as about *what* a court can adjudicate. Thus, the issue we refer to as *subject matter jurisdiction* is about whether a court can hear cases of this kind, as when, for example, a federal court must decide whether its authority to decide diversity of citizenship cases between citizens of different states includes the authority to decide a case between a citizen of a state and another state itself.[26]

But, as noted briefly above, the concept of jurisdiction looms larger in law than we might suspect if we were to focus only on the procedural questions of personal and subject matter jurisdiction. In this larger sense, jurisdiction is about who gets to decide what, and a pervasive concern of the legal mind is the question of whether some institution that made a decision had the authority to make it. It is not enough, especially for the law, that the decision was right. It must also be the case that the judge or other person making the decision, or the institution making the decision, was authorized by the system to make it. Whether it is constitutional concerns of federalism, or the question of when congressional legislation has preempted the states, or whether it is for a judge or a jury to decide questions of fact, most legal decisions involve, sometimes explicitly but usually implicitly, the question not only of whether some decision was right but of whether the right institution made it.

Many of these issues arise in the context of direct determinations of jurisdiction or decision-making authority. But sometimes they are embedded in the important question of *deference:* under what circumstances will a decision-maker respect the decisions of another body, even when the decision-maker thinks the decision of the other body is mistaken?[27] In some respects this resembles the questions of authority we took up in Chapter 4, but deference often operates in just the opposite direction. When we think of authority, we ordinarily think of it as a relationship from top to bottom. Those in authority—sergeants, parents, teachers, supreme courts—give orders or make decisions, and those below them are expected to obey even when they disagree with the decisions. To respect authority is to look up from the bottom to the top. Deference also involves respecting decisions with which the deferrer may well disagree, but by contrast, often it is from top to bottom rather than from bottom to top. Deference is the way in which those in the higher po-

26. It does not. Moor v. Alameda County, 411 U.S. 693, 717 (1973).

27. See generally Philip Soper, *The Ethics of Deference: Learning from Law's Morals* (2002).

sition in some hierarchy allow some leeway to those below them in the service of efficiency, respect, specialization, diversity, or separation of powers in the broad sense.

Thus, it is frequently the case that an appellate court will defer to the determinations of a trial judge, and in doing so it says that even though it exists above the trial judge in the hierarchy, it will accept some number of trial decisions as valid, even if it disagrees with them. So too when a judge defers to the jury, when the Supreme Court defers to Congress or an administrative agency, or when the European Court of Human Rights allows individual states a *margin of appreciation* in making their own individual rights determinations.

Deference is in important ways closely related to presumptions and burden of proof, and often the mediating idea is the idea of the *standard of review.* When an appellate court reviews the decision of a court below, or when a court reviews the actions of an administrative agency, it typically operates according to an explicit standard of review, which in a way is the mirror image of a burden of proof. It is often the case, for example, that a court will not overturn the action of an administrative agency unless it finds the action to be *arbitrary and capricious,*[28] and an appellate court will not generally overturn a trial court ruling with respect to the admission or exclusion of evidence unless it finds that there has been an *abuse of discretion.*[29] Such standards are highly deferential, because their implication is that most evidentiary and administrative decisions will remain standing, even by judges who disagree with them, unless the decisions are, say, outrageously or extremely wrong. Decisions that are perceived from above to be only somewhat wrong, or maybe even significantly wrong, will under these highly deferential standards be allowed to stand.

Similar standards apply to the review of jury verdicts, to appellate review of trial court findings of fact, and to myriad other questions. But at other times the standards of review are not nearly so deferential. When an appellate court reviews a finding *de novo*—"from the beginning"—the idea is that the appellate court will not defer to the court or agency below but will make its *own* decision, and this is frequently the standard

28. See William S. Jordan, "Ossification Revisited: Does Arbitrary and Capricious Review Significantly Interfere with Agency Ability to Achieve Regulatory Goals through Informal Rulemaking?," 94 *Nw. U. L. Rev.* 393 (2000).

29. See Old Chief v. United States, 519 U.S. 172 (1997).

explicitly or implicitly applied to review of questions of law. When an appellate court is reviewing a trial court ruling about the interpretation of a statute, for example, it will typically say or assume that its ability to interpret the law is no less than that of a trial court. There are no witnesses to observe or physical evidence to evaluate, and thus the standard reasons for appellate deference to trial decisions—apart from simple efficiency—are no longer present, and under such circumstances appellate courts become less hesitant to substitute their judicial judgment for that of another judge. So too, traditionally, for review of legal determinations by an administrative agency, where the presumed greater expertise of the agency entitles it to deference as to questions of fact and policy but not for questions of statutory interpretation.[30]

Deference, legal presumptions, and the burden of proof are thus concepts that not only work together but that are also in some respects and in some contexts different ways of saying the same thing from varying standpoints. It may be hard to understand the idea that a defendant in a criminal case is entitled to deference, and thus there may not be a corollary in the criminal law context to the fact that the prosecution bears the burden of proof. In other contexts, however, the relationship among the burden of proof, deference, and the standard of review is more apparent. It seems quite sensible, therefore, to say that because an administrative agency is presumed to have come to the correct conclusion within its domain of expertise, then it is entitled to deference, and thus that its decision may be overturned only if the challenger can meet the burden of proving that the agency decision was, say, arbitrary and capricious. There is only one relationship in the previous sentence, but deference, presumption, the burden of proof, and the standard of review are all different ways of describing that same relationship from different angles.

The relationships and concepts described in this chapter not only relate to each other, but also connect to the larger themes in this book. When a

30. After the Supreme Court decision in Chevron, U.S.A., Inc. v. Natural Resources Defense Council, Inc., 467 U.S. 837 (1984), reviewing courts are expected to defer to an agency even with respect to interpretations of the agency's principal governing statutes. The Court in *Chevron* justified this change on the basis of the agency's presumed expertise in understanding the ins and outs of the technical and specialized statutes with which it deals constantly, but *Chevron* might also be justified simply as a matter of conservation of scarce judicial resources.

court defers to a lower court or to an agency, or when it operates under a high standard of review, it in effect commits itself to accepting—maybe *tolerating* is the better word—some number of what it perceives as erroneous outcomes. So too with presumptions, which under conditions of uncertainty will sometimes compel factual conclusions that are simply not true. And when a party finds itself saddled with the burden of proof, it will on occasion not be able to prove something that is in fact the case. Just as the "beyond the reasonable doubt" standard in criminal law will acquit some number of people who are probably guilty, so too will any high burden of proof commit the system in which it operates to some number of mistakes, although it is to be hoped that they are, like the mistake of acquitting the guilty, mistakes of the right kind.

In accepting the inevitability and strategic or long-term desirability of some number of mistakes of mostly the right kind, the legal system's use of burdens of proof, presumptions, standards of review, and principles of deference, perhaps especially the last, resembles the system's use of rules and precedents, and resembles the legal system's at least partial commitment to formality. It may also resemble the legal system's willingness to make decisions on the basis of less than all of the best or available information. In all of these dimensions, the law, more than many other decision-making institutions, commits itself to accepting wrong or at least suboptimal answers, and it does so in the service of larger or longer-term institutional values, as well as service to the idea that the best way to get the largest number of correct decisions in the long term is often something other than attempting to make the best decision on every occasion. In operating in this way, law and legal reasoning may not be different in kind from other decision-making institutions, but they may differ in degree. At the heart of much of law's use of its characteristic reasoning devices is its acceptance of the fact that the best decision is not always the best *legal* decision. In operating in this fashion, law does not intend to be perverse. It does, however, intend to take institutional values especially seriously, and it does *that* in the hope that in the long run we may be better off with the right institutions than we are when everyone simply tries to make the best decision.