with dogs, even that it is an example of racism. But as we also saw in the case of the dogs, on closer analysis the rhetoric is often seen to be misplaced.

Yet even as we recognize the way in which the rhetoric of unconstitutional, illegal, and immoral discrimination has seeped into far less morally problematic areas, the analysis of pit bulls may have created as many problems as it solves. For although the analysis cautions us against assuming that generalizations and stereotypes can (or should) ever be eliminated entirely, it raises questions regarding the distinction between permissible and impermissible stereotypes about people. If the principal argument against the opponents of breed-specific restrictions on dogs is that breed-specific restrictions are little different from a large number of inevitable and widely accepted bases for restriction, then is not the same true for people as well? If stereotyping dogs is permissible not only because dogs are dogs and not people, but also because stereotyping and generalizing are ineliminable, then what does this say about people, and what does it say about the distinction between permissible and impermissible stereotypes? It is this conundrum that will be the central focus of much of the rest of this book.

CHAPTER 3

A Ride on the Blue Bus

Betty Smith and the Blue Bus Problem

On January 6, 1941, Betty Smith was driving her car from Dorchester to Winthrop, Massachusetts. Entering Winthrop at about one o'clock in the morning, she was crowded off the road by a bus, and collided with a parked car. Smith was injured in the accident, and sued Rapid Transit, Inc., in the Superior Court of the Commonwealth of Massachusetts.

Because the accident occurred in the middle of the night, because the bus that forced her off the road did not stop, and because she was preoccupied with trying to avoid the accident, Betty Smith did not see any of the identifying marks on the bus. At the trial she could testify only that the vehicle that forced her off the road was a bus, and that it was "a great, big, long wide affair." Smith was, however, able to prove that the Rapid Transit Company had been licensed by the City of Winthrop

to operate buses on the very route on which the accident had occurred; that Rapid Transit's buses left Winthrop Highlands on the thirty-minute trip to Maverick Square, Boston, at 12:10 A.M., 12:45 A.M., 1:15 A.M., and 2:15 A.M.; that this route included the Main Street location of the accident; and that no other company was licensed to operate its buses on this route.

Despite this evidence in support of the proposition that the bus that caused the accident was operated by Rapid Transit, the trial court refused to let the jury even consider the case. The judge ruled that Betty Smith could not, as a matter of law, recover against Rapid Transit, because there was no *direct* evidence that the bus that hit Smith was one of Rapid Transit's buses. This ruling was upheld by the Massachusetts Supreme Judicial Court, which noted that Rapid Transit's exclusive franchise "did not preclude private or chartered buses from using this street; the bus in question could very well have been operated by someone other than the defendant." The court acknowledged that this was unlikely and that "perhaps the mathematical chances somewhat favor the proposition that a bus of the defendant caused the accident." But "this was not enough," the court said, concluding that the mathematical probability that the bus in question was the defendant's bus was not the type of "direct" evidence that could lead a jury to have an "actual belief" in the proposition that this was one of Rapid Transit's buses.[1]

Had this been a criminal case, the ruling would strike us as unexceptionable. After all, the evidence that the bus belonged to Rapid Transit was hardly of the quality and quantity that would establish "beyond a reasonable doubt" that this was one of Rapid Transit's buses, especially given the small but hardly inconceivable possibility

that the accident was caused by a private bus or by a chartered bus. But Smith's lawsuit was a civil and not a criminal case. Accordingly, the required standard of proof was not that of "beyond a reasonable doubt." Instead, Betty Smith needed to establish her case only "by a preponderance of the evidence." And as the equivalent phrase in English law, "by a balance of the probabilities," indicates, we ordinarily understand the preponderance of the evidence to be the equivalent of just over a 50 percent likelihood that the proposition asserted is true.[2] Whatever the possibility might have been that a private or chartered bus caused the accident, no one claimed that the probability of such an occurrence was anything approaching 50 percent.[3] Thus there seemed to be no reasonable denial that the evidence presented by Smith established to a probability considerably greater than .5 that this was Rapid Transit's bus. But if that was the case, then why did Betty Smith not win?

Why not indeed? Smith's case is hardly unique,[4] and the Supreme Judicial Court's ruling is generally in line with the law as it was then, and as it is now.[5] Yet it still seems odd that if the plaintiff is required only to prove her case to a probability of .51 (to put it roughly), then statistical evidence that would do so is thought by itself to be insufficient, or so the courts routinely conclude. Indeed, it seems so odd to so many people that Smith's case has become a staple of academic teaching of evidence law in law schools, and the centerpiece of much of academic writing about what has come to be called the problem of "naked statistical evidence."[6] Commonly, the problem is made analytically crisper when presented as a hypothetical version of the *Smith* case known as the Blue Bus Problem: Suppose it is late at night, under one version of the problem, and an individual's car is hit by a

bus. This individual cannot identify the bus, but she can establish that it is a blue bus, and she can prove as well that 80 percent of the blue buses in the city are operated by the Blue Bus Company, that 20 percent are operated by the Red Bus Company, and that there are no buses in the vicinity except those operated by one of these two companies. Moreover, each of the other elements of the case—negligence, causation, and, especially, the fact and the extent of the injury—is either stipulated or established to a virtual certainty. In these circumstances can the plaintiff recover in civil litigation against the Blue Bus Company, or, if not (as the overwhelming majority of American courts would conclude), then why not? Or consider a variation of the Blue Bus Problem even closer to Betty Smith's case: The plaintiff's car is hit by a bus late at night and all she knows about the offending vehicle is that it was a bus. Eighty percent of the buses in town are operated by the Blue Bus Company. Can the plaintiff win a lawsuit againt the Blue Bus Company on that evidence alone, assuming, as in the previous examples, that there is nothing in dispute about the issues of causation, negligence, or injury?

The Generality of Statistics and the Statistics of Generality

Scholars have been debating the Blue Bus Problem for decades, sometimes in the highly technical language of mathematical statistics, and sometimes in a more commonsense way.[7] A few scholars have defended the legal system's skepticism about statistical evidence. Laurence Tribe and, later, Charles Nesson, for example, have pointed to the way in which explicit acknowledgment of the probability of error might, even if accurate, undermine confidence in the legal system. Other scholars urge increased acceptance of statistical evidence, sometimes arguing that a legal rule should not be premised on keeping jurors and the public in the dark about the actual nature and consequences of legal decisions. Yet regardless of the correct outcome of this scholarly debate, it is important to draw attention to the way in which the debate about naked statistical evidence links more closely than the literature recognizes to the superficially different questions about the role of generality in decisionmaking.

Recall the discussion of pit bull regulation in chapter 2. The problem (to some) with using the generalization "pit bull," which gathered up all the individual pit bulls with their diverse individual characteristics under the single category of pit bulls, was that one attribute of the category—a tendency toward dangerous aggressiveness—was not necessarily an attribute of each member of the category. The generalization about the danger of pit bulls is not spurious—the evidence plainly establishes that dangerousness exists in the class of pit bulls to a greater degree than it does in the class of all dogs, and to a greater degree than it does in almost all the subclasses that we call breeds—but there is still no disputing that many pit bulls, quite possibly the vast majority of them, are not dangerous at all.

Similarly with the Blue Bus Company. If the relevant attribute is ownership of a particular bus, as in the second version of the Blue Bus Problem, and if the Blue Bus Company owns 80 percent of all the buses, then the Blue Bus Company possesses the attribute of ownership of this particular bus to a higher probability than does any another possible defendant, just as pit bulls possess the attribute of dangerousness to a higher probability than most other breeds and to a higher probability than the

class consisting of all dogs. Moreover, the Blue Bus Company possesses the attribute of ownership of the bus in question to a probability seemingly sufficient to justify liability in a civil lawsuit. If we were to hypothesize that 80 percent of the vicious dogs were pit bulls, then we could conclude, absent further information, that an attack by an otherwise unidentified vicious dog was 80 percent likely to have been an attack by a pit bull. Similarly, if 80 percent of the buses are owned by the Blue Bus Company, then we could conclude, absent further information, that an accident caused by an otherwise unidentified bus is 80 percent likely to have been an accident caused by a Blue Bus Company bus.

Casting the problem in this way brings to mind another famous hypothetical case, this one offered by the British philosopher L. Jonathan Cohen. In what he labels The Paradox of the Gatecrasher, Cohen hypothesizes a rodeo that charges for admission.[8] During the rodeo the organizers of the event count the spectators, and they discover that there are 1,000 in attendance. When at the end of the rodeo the organizers count the tickets collected at the ticket booth, however, it turns out that there are only 499 tickets in the ticket box. The mathematical corollary of this, of course, is that 501 of the 1,000 spectators at the rodeo were gatecrashers. So now suppose that the organizers of the rodeo bring a lawsuit against one—any one—of the 1,000 spectators for fraudulent entrance. No one saw this *particular* person enter fraudulently, and there is no other evidence connecting this particular defendant to a fraudulent entry. Yet still, absent any other evidence, there is a .501 probability that this person (or any of the other 999 spectators) was a gatecrasher. Why, then, cannot the statistical evidence by itself be sufficient to warrant a verdict, at least under the preponderance standard in civil litigation, in favor of the rodeo organizers? Cohen himself maintains that such a verdict would be profoundly unjust. For him the paradox consists not in the unwillingness of the courts to award a judgment to the rodeo organizers and against the alleged gatecrasher, because Cohen believes that this would be wrong. Rather, Cohen endorses the fact that courts would not award damages in such a case, but he finds it puzzling that courts continue to insist that the standard of proof in civil cases is a preponderance of the evidence, a standard that the .501 likelihood of gatecrashing by any one of the spectators appears to satisfy.

My aim is not to "solve" either the Paradox of the Gatecrasher or the Blue Bus Problem. It is, however, to show that both of these problems are best seen as variants on the larger problem of generality in decisionmaking. One way of doing this would be to start with seeing the problem of generality as an aspect of the problem of attempting to determine when we should and should not use statistically reliable but nonuniversal indicators. In this sense the problem of generality is "really" the problem of statistical inference, and thus the Blue Bus and Gatecrasher problems, which appear to be problems of statistical inference, resemble the problem of generality, because all are problems relating to the wisdom or justice of using nonuniversal but nonspurious statistical indicators.

Alternatively, and preferably, both the Blue Bus Problem and the Paradox of the Gatecrasher, which are typically presented as problems of statistical inference, are fundamentally problems about the use of generalizations. In each the issue is not simply a problem of statistics but instead a problem about the extent to which we can employ, at least for purposes of awarding damages in

civil litigation, generalizations about rodeo spectators (most but not all entered fraudulently) and a generalization about the Blue Bus Company (owns most of the buses in this city). When the issue is framed in this way, the problems of generality and generalization become primary, and the problem of statistical inference is seen as but another way of describing what is at its core an issue about generalization.

The desirability of framing the issue as one fundamentally about generalization becomes even clearer once we understand that what the Massachusetts Supreme Judicial Court in *Smith* saw as the problem was not a problem of statistics at all. Rather, the court, although it did utilize the potentially confusing language of "mathematical chances" and "probability," was primarily focused on what it saw as the difference between so-called direct or actual evidence, on the one hand, and the kind of evidence that is based on the characteristics of the class of which the alleged perpetrator is a member, on the other.[9] This is even clearer in the Paradox of the Gatecrasher. Again, the fact that "statistics," in the numerical sense of that word, might have been part of the hypothetical rodeo organizers' case is largely beside the point. When the organizers bring a case against a particular individual, they base their claim on the attribution of nonspurious class characteristics—nonpayment of the admission charge—to an individual member of the class. In doing so, the organizers rely on the same process of generalization that Plato's training master relied on in attributing the characteristics of "the herd" to each of its members, that pit bull ordinances rely on in attributing the characteristics of the class of pit bulls to each individual pit bull, that insurance companies rely on in attributing the characteristics of the class of teenage male drivers to

each teenage male driver, and that many of us rely on in attributing the honesty of the class of dealers in used automobiles to each dealer in used automobiles. In all these cases, the process, in the final analysis, is the process of basing decisions for all members of a class on nonspurious but nonuniversal characteristics of the class taken as a whole. This is the process of generalization, and this is the process of which the problem of so-called statistical evidence is but one component.

The conclusion of the immediately preceding paragraph notwithstanding, it may not be overly important whether it is statistical inference that is primary and generalization secondary, or generalization that is primary and statistical inference secondary. What is important is that we can appreciate that the seemingly disconnected issues of generality and statistical evidence are in fact remarkably similar, and that the resources that enable us to understand and negotiate the problem of generality are the same resources that can be used to understand and negotiate problems about the use of statistical evidence in civil and criminal trials. And once we understand this, there remains more to be said about these problems of statistical evidence and the light they shed on the issue of generality.

Probabilistic Inference in an All-or-Nothing World

The Blue Bus and Gatecrasher problems are in an important way artifacts of the all-or-nothing manner in which most aspects of most modern legal systems operate. In much of nonlegal life people can act on their uncertainty by making decisions in accordance with the principle of expected value. In doing so, they value an uncertain outcome by multiplying the value of some set

of consequences by the probability that those consequences will come to pass. The product is the expected value of those consequences. If you have a 50 percent chance of winning ten dollars, the value to you is five dollars. Just as a wager of eight dollars is therefore a good one if you are betting on a 10 percent chance of winning one hundred dollars, so too do we act in similar ways in much of our daily life. We invest less in risky investments than in more certain ones, we make shorter commitments when we are unsure of the value of what we are committing to than when we have greater confidence, we calculate how much insurance to buy on the basis of expected value (just as the insurance company does in determining how much to charge us), we plan travel times by factoring in the probability of delays, and we calculate expected fines in deciding whether it is worthwhile committing minor illegalities such as overtime parking, all of these decisions and more making the expression to "hedge one's bets" applicable in much more of our lives than the occasional trip to the racetrack. In these cases, and many more, an imprecise but serviceable conception of expected value guides many of our daily decisions.

Because statisticians understand and use the principle of expected value, to the statistician the Paradox of the Gatecrasher may be no paradox at all. If there is a .51 probability that any given spectator entered fraudulently, and if the purchase price of a ticket is $1.00, then the statistician sees the easy solution: the rodeo organizers recover 51 cents against each of the 1,000 spectators. In this way the rodeo organizers recover only their fair share of the proceeds, and each spectator is liable only to the extent of the likelihood that he or she entered without purchasing a ticket. And so too with the Blue Bus Problem. If there is a .80 chance that the bus that plainly negligently caused an indisputable $1,000 worth of damages to, say, Betty Smith, is a bus owned and operated by the Blue Bus Company, then the principle of expected value would indicate that Smith should recover $800 against the Blue Bus Company.

The law, however, does not operate this way. Perhaps oddly to the statistician, the law would give Smith all of her damages if she proved her case to a .51 probability, and nothing if she proved it to a .49 probability. And it would give her not a dollar more if she proved her case to a .90 probability than if she proved it to a .51 probability. With rare exceptions, the expected value of a plaintiff's claim, by which the extent of the plaintiff's proof would be multiplied by the extent of the plaintiff's damages, is not a principle of advanced legal systems.[10] These systems, we see throughout the world, are all-or-nothing affairs.[11]

In the context of a criminal case, our intuitions confirm the approach of the law. If there is a .70 chance that the defendant is the one who committed an aggravated assault, and if the penalty for aggravated assault is ten years' imprisonment, few of us, and not even the statisticians, would be comfortable imposing a sentence of seven years based on the principle of expected value. And perhaps that is so because of the strength of the maxim, first offered by William Blackstone, that "it is better that ten guilty persons escape, than that one innocent suffer."[12] The value we place on liberty, and thus the graveness of the error of denying liberty to the innocent, makes us uncomfortable with imprisoning those who are .30 likely to have done nothing wrong, and thus the principle of expected value is properly a stranger to the criminal law.

In civil cases, however, the aversion to expected-value verdicts seems less justifiable. After all, the plaintiff in a typical negligence case is claiming to have been injured through someone else's fault while doing nothing wrong. In such a case, it is not clear why erroneously denying recovery to a worthy plaintiff is any less harmful an error than erroneously awarding recovery against a nonnegligent defendant. To put it differently, the preponderance of the evidence standard presupposes that erroneous denials of liability and erroneous impositions of liability are equally regrettable.[13] The false positive is no worse than the false negative. And if this is so, if the Type I and Type II errors, to use the statistician's language, are equivalent, then it is by no means clear that the aversion to expected-value verdicts in criminal cases ought to be extended to civil cases.

The law, however, does not agree, and continues to be pervasively and perhaps perversely insistent on an all-or-nothing approach. The lessons of expected-value analysis notwithstanding, the law dismisses as too easy the statistician's solution to the Paradox of the Gatecrasher and the Blue Bus Problem. Most legal systems continue to resist expected-value outcomes, and as a result it is plausible to conclude that the difficulties presented by the Blue Bus Problem, the Paradox of the Gatecrasher, and other real and imagined examples are largely the products of the all-or-nothing character of most legal decisionmaking.

Once we see the relationship between the paradoxes of the law of evidence and the all-or-nothing nature of legal decision, however, we can understand the larger problem of generalization in a new light. For if the problem of statistical inference in the law of evidence is, as we have seen, little more than one instance of the problem

of generalization, then the problems created by an all-or-nothing legal system parallel the problems created by the all-or-nothing parts of many other dimensions of our decisional lives. In numerous instances in which we employ probabilistically sound but nonuniversal generalizations in ordinary decisionmaking, it is because the nature of the decision makes an expected-value decision impossible or, at the very least, impractical. If I am looking for a pet, it is not possible for me to have a pit bull for one day out of seven and a golden retriever for the other six. Similarly, tax officials rarely conduct partial audits (even though some audits are more thorough than others), customs officials rarely conduct partial inspections, police officers cannot conduct partial stops, airlines do not believe that they can deal with the problem of pilots 10 percent more likely to cause an accident by having them fly 10 percent fewer flights, and hockey referees who are 75 percent sure that a player has committed a high-sticking infraction do not have the option of sending the offender to the penalty box for ninety seconds rather than the designated two minutes for that offense, any more than a football official unsure of whether a defensive lineman was offside can penalize the defense three yards rather than five. These and many more examples suggest that life as well as law is often an all-or-nothing affair, and that what looks at first to be the special all-or-nothing quality of the legal system may actually be found in much of nonlegal decisionmaking. In more cases than we or the statisticians might suppose, nonlegal decisionmakers often understand themselves to be making all-or-nothing decisions (do I hire this person as a babysitter or not) in which the expected-value approach is just not available. The use of generalizations, therefore, appears to be not only an outgrowth of the frequent need to use

generalizations as a time- and effort-saving heuristic in circumstances in which individual determinations would probably be too costly or too prone to the errors of discretion; it is also a function of the way in which expected-value decisionmaking is considerably more of a stranger to everyday decisional life than we may at first have fully appreciated.

Individuality and Reliability

It is thus the nature of most of legal and more than we thought of nonlegal decisionmaking that requires us to engage in all-or-nothing decisionmaking. Consequently, if the nature of all-or-nothing decisionmaking pushes us toward what seem to many people to be unjust outcomes, then one way of understanding the instinct behind the *Smith* rule is as a desire to minimize the number of erroneous outcomes inevitably generated by all-or-nothing decision procedures. Perhaps the insistence on so-called direct or actual evidence, as the court in *Smith* naively put it, is explained by a reluctance to have the legal system forced into accepting the 20 percent error rate that giving Betty Smith 100 percent of her damages on an 80 percent chance of Rapid Transit's liability would entail.

Yet if this kind of error minimization is the goal, then it is hard to see how a supposed requirement of "direct" or "actual" evidence serves it. Initially, we can ask what the Massachusetts Supreme Judicial Court in *Smith* might have meant by the terms "direct" and "actual." Presumably the court had in mind evidence that comes from a perception of a witness, with that very witness then testifying to that perception in court.[14] Typically this would be a visual perception—an eyewitness—al-

though there can also be perceptions by any of the other senses—hearing, smelling, tasting, and touching. But apart from sensory perception testified to under oath by the perceiver in court, it is difficult to see what the court could have meant by the terms "direct" and "actual."

If "direct" and "actual" refer to perceptual evidence testified to by the perceiver, then we must consider the reliability of this evidence as compared to the allegedly indirect or "nonactual" evidence offered in *Smith* and similar cases. Consider, therefore, another hypothetical variation of Betty Smith's case. Suppose Betty Smith testified that she saw what looked like the words "Rapid Transit" written in red letters on the side of the blue bus that hit her. But then suppose that on cross-examination the accuracy of her account is called into question by Rapid Transit's lawyer. Betty Smith, let us suppose, is forced to admit that it was foggy and rainy that night, that the eyeglasses she always wears were knocked from her head by the impact of the accident, that she first reported her observation of the words "Rapid Transit" not to the police officer who came upon the scene of the accident but only later after having consulted with an attorney, and that she saw the words only as the bus was heading away from her, at an angle to her direct vision, at a speed of no less than thirty miles per hour, and at a distance of no less than 200 feet. Yet despite all these reasons to doubt the accuracy of the hypothetical Smith's observation of the words "Rapid Transit," and despite the fact that it might be reasonable to place the probable accuracy of her observation of the words "Rapid Transit" at well less than .80, the very court that refused to let the real case go to the jury, even on a probability well above .80 that the bus in question was a Rapid Transit bus, would almost certainly have let the "fuzzy observation"

case go to the jury on a probability well below .80 that the bus in question was a Rapid Transit bus. In this hypothetical case, the court would in all likelihood have said that these issues of credibility are for the jury and for the jury alone to determine.

Part of this anomaly of excluding more-reliable statistical evidence and admitting less-reliable personal testimony is explained by the widespread but empirically unsupported faith in eyewitness identification. Although there persists an aura of credibility historically attached to eyewitness accounts, a raft of serious psychological research has established that much of this historical faith in eyewitness testimony lacks a sound empirical foundation. People often see what they want to see, or see what they think they are expected to see, or see what they are positively reinforced in seeing. To put it slightly differently, people's perceptions are somewhere between usually and always filtered through their own biases, prejudices, and preconceptions; they simply forget or misremember what they saw; and they are afflicted with a host of other cognitive deficiencies that make eyewitness testimony much less reliable than the conventional wisdom would suppose.[15] If the preference for direct or actual evidence is based on a preference for perception over inference, then almost all of what we know about the deficiencies of human perception cast doubt on such a preference.

These doubts about perceptual abilities are exacerbated by the tendency of people not only to overweight perception as an empirical matter, but also to ignore what statisticians and psychologists call "base rates," thus leading people to make logical as well as empirical errors.[16] Consider an example made famous by Amos Tversky and Daniel Kahneman, an example that bears a close resemblance to the Blue Bus Problem.[17] Suppose

that the Green Cab Company owns and operates taxis that are green in color, and the Blue Cab Company owns and operates blue taxis. Eighty-five percent of the taxis in town are the green taxis of the Green Cab Company, and the other 15 percent are the blue taxis of the Blue Cab Company. As in the *Smith* case, suppose a car is sideswiped or run off the road by a taxi, and a witness is 100 percent certain, presumably from the light on top of the cab, that the "guilty" car is a taxi, and is confident, but not certain, that the guilty taxi was blue. Suppose that the witness is 80 percent confident that the taxi was blue and thus that it was a taxi of the Blue Cab Company. On this basis, is it more likely that the taxi was a green taxi of the Green Cab Company or a blue taxi of the Blue Cab Company?

On these facts, most people would conclude, with the witness, that it is more likely that it was a blue taxi than that it was a green taxi and that the Blue Cab Company should therefore be held liable. But this conclusion gets it exactly wrong. The conclusion that the taxi was probably blue because the witness said so with a moderately high degree of confidence ignores the base-rate distribution of taxis. For most people, what they perceive as "evidence" overwhelms the underlying base rate. In other words, the conclusion that the taxi was blue ignores the fact that the witness's .2 likelihood of error must be applied to the actual distribution between blue and green taxis and not to a presumed 50–50 distribution when the distribution is not in fact 50–50. Thus the number of cases, on these probabilities, in which a witness said the cab was blue when it was green turns out to be somewhat higher than the number of cases in which a witness said the cab was green when it was blue. On these probabilities, in fact, the probability that the cab was green is .59

despite the fact that the witness was .80 certain that it was blue.[18]

The prevalence of ignoring the base rate, combined with the prevalence of overestimating the reliability of eyewitness testimony (which may be a contributing factor in people's willingness to ignore the base rate in cases like these), makes the legal system's prevailing skepticism about statistical evidence even more puzzling. As the above examples of fuzzy or otherwise uncertain observations are designed to illustrate, the kind of evidence commonly thought to be direct or nonstatistical is often far less reliable than the kind of evidence often thought to be indirect or statistical. Or, to translate this into the language of generality, it may frequently be the case that the inferences to be drawn from nonspurious but nonuniversal generalizations are empirically stronger than the inferences to be drawn from decisionmaking approaches that seemingly do not rely on generalizations or at least rely on smaller rather than larger ones.

There is an interesting parallel between the legal system's (and the public's) traditional but misguided preference for eyewitness testimony and the traditional and often equally misguided preference of many psychologists and physicians for clinical as opposed to actuarial assessments. Suppose the issue, a very common one, is trying to predict which offenders if released on parole will commit further offenses. Or suppose it is the similar issue of which people, having been found not guilty of some crime by reason of insanity, can safely be released into the community. In these and related cases the traditional view has been that a thorough and face-to-face psychological examination—a clinical assessment—is the most reliable method of predicting dangerousness. Yet much of the modern research has shown that actuarial assessments turn out to be more reliable than clinical ones.[19] If instead of performing a clinical assessment the authorities were simply to look at a group of actuarially tested but easily identified indicators—nature of the offense; age of the defendant; number of previous offenses; and so on—they would have more reliable indicators of dangerousness than if they were to rely on clinical assessments; and this is so even if the clinical assessments take into account these very same factors along with any others that the clinician believes relevant in the particular case. This outcome may at first seem surprising, but it is much like the problem with eyewitness testimony. Clinicians, even well-trained ones, often have excess confidence in their own perceptions, are sometimes influenced by biases and agendas that they themselves do not fully appreciate, and are frequently resistant to the base rates of dangerousness for the population they are evaluating.[20] For these and other reasons, therefore, relying on actuarial generalizations typically turns out to be more reliable than relying on the direct perceptions and intuitions of even highly trained professionals.

Whether we are talking about evidence in court or assessments of dangerousness by psychologists, the frequent empirical superiority of decisionmaking by generalization over direct individual perception may not be all there is to the matter. As we will explore in subsequent chapters, people may think that there is a moral imperative in maximal individuation in decisionmaking even if the actual practices of such individuation are less reliable than the alternative.[21] But at the very least the preference for individuation, of which Betty Smith's case is but one example, cannot plausibly be seen as resting on some overall greater accuracy of nongeneralized decisionmaking.

The possibility that relying on generalizations known from the beginning to be imperfect might still be empirically superior to relying on allegedly direct or individualized assessments also replicates important aspects of the debate about the virtues and vices of rules and rule-based decisionmaking. As prescriptive generalizations, rules necessarily entail the possibility that their strict application will produce suboptimal outcomes in some cases, where suboptimality is measured by reference to the outcome that would have been produced by accurate application of the background justification lying behind the rule.[22] To take a hoary example from the world of legal philosophy, if in order to prevent noise in the park (the background justification) we prohibit all vehicles from entering the park (the rule), we then produce a suboptimal result whenever we exclude nonnoisy vehicles (bicycles and electric cars) and whenever we fail to exclude noisy nonvehicles (musical instruments and loud radios).[23]

The inevitable suboptimality of rules, however, is premised on a supposition about the accuracy of individualized decisionmaking. We know, however, that this accuracy often does not exist, and especially when there are reasons of bias and mistake, among others, to distrust the reliability of the individualized decision. If there were grounds to believe that enforcement officers would make numerous mistakes in trying to determine which instrumentalities were noisy and which not, then in practice the suboptimal rule could very well produce fewer errors than the theoretically optimal individualized assessment.

The same question arises in a recent and very real context. In the face of evidence that many (possibly as many as a thousand a year in the United States) fatal automobile accidents have been caused by inattentive drivers talking on their cell phones when they should have been watching the road, the state of New York enacted a law prohibiting people from using telephones while driving, and many other states and a number of countries outside the United States are now considering similar laws. But as with the bans on pit bulls, people complained that focusing only on cell-phone users was under- and overinclusive, and therefore unfair.[24] Just as many pit bulls are nonvicious and many other kinds of dogs can be vicious, the cell-phone users and the cell-phone industry argued that for many people, talking on the phone while driving is no more distracting than listening to the radio or conversing with a passenger, making the law overinclusive, and that there were many sources of distraction, such as billboards, not covered by the law, making the law underinclusive. As a consequence, some states, such as New Hampshire, rejected cell-phone-specific regulation, and instead enacted laws prohibiting not cell-phone use, but driving while distracted.

As with the debates about both vehicles in the park and clinical assessments of dangerousness, however, the cell-phone issue presents the debate between the virtues of admittedly under- and overinclusive regulation by easily identifiable indicators—either you are on the phone or you are not, and it is not that difficult for a police officer to make that determination—and the virtues of more sensitive assessment by determining in each individual case whether the driver was distracted or not. But of course the sensitive determination of the police officer about which drivers are distracted and which not, like the sensitive determination of the clinician about which offenders are still dangerous and which are not, is

also subject to mistakes. These are not the mistakes built into crude but simple actuarial measures; instead, they are the mistakes that come when police officers, like clinical psychologists, substitute the errors of misperception and bias, among others, for the errors that might be part of using a nonuniversal but statistically reliable and easily applied actuarial assessment, of which the actuarial assessment that cell-phone use is a common distraction for drivers is but one example.

The debate about statistical evidence, therefore, is like the debates about clinical assessments and much like the debates about rules, whether in cases involving pit bulls, vehicles in the park, or mobile phones. Each of these debates turns out to be about the advantages and disadvantages of relying on nonspurious but nonuniversal generalizations, and each of these debates then turns out to compel a focus on the advantages and disadvantages of relying on generalizations compared to relying on seemingly more individualized assessments. And when we look at the evidence, it is often the case that the aversion to generalization rests on erroneous empirical foundations. An aversion to generalization is typically based on an unwillingness to accept the mistakes that decisionmaking by generalization necessarily entails. But it is less often recognized that an aversion to large-scale generalizing must assume that the actual human beings who make more individualized decisions would in practice make fewer mistakes than those made in relying on the generalization. As the comparison of the record of unreliability of eyewitness testimony with the greater reliability of at least some statistical generalizations shows, however, and as the studies comparing actuarial with clinical psychological assessments reinforce, this assumption is often simply false. If there is something that

is troublesome about relying on larger generalizations per se, it cannot be that there is good reason to believe that such reliance is necessarily or even typically likely to produce more errors than the alternative.[25]

The Nonindividual Nature of Individualized Evidence

The objection to preferring so-called direct or actual evidence to other sorts of evidence, however, is not only an empirical one. Rather, the objection rests as well on understanding that the avoidance of generalizations is, with few or no qualifications, simply not possible at all. Put differently, even those decisions that appear initially to be maximally individual, that appear to be "direct" or "actual," in the words of the Massachusetts Supreme Judicial Court in the *Smith* case, may turn out to rely more on generalizations than many people suppose. The inevitability of generalization was the conclusion of our analysis of pit bull regulation, and considering what might possibly be meant by "direct" as opposed to statistical evidence makes the point even clearer.

Because most readers of this book are not visually impaired, it may be easier to see the issue by considering another hypothetical example, here one involving direct but nonvisual perception. Suppose there was a totally blind passenger in Betty Smith's car. And suppose as well that the Blue Bus Company owns all the buses in the city, and indeed all the buses in the county and surrounding counties. Because the possibility of buses owned by others is so minuscule, the defendant Blue Bus Company is prepared to concede that if Betty Smith's car was crowded off the road by a bus then it was crowded off the road by one of the Blue Bus Company's buses.

That Betty Smith was crowded off the road by a bus rather than a car, truck, or piece of construction equipment, however, is something that the Blue Bus Company is not willing to concede. Taking the position that Betty Smith's alleged visual observation of a bus was a fabrication (the Blue Bus Company being wealthy and well insured), the Blue Bus Company attempts at trial to cast doubt on the part of her story maintaining that it was a bus that crowded her off the road. In order to counter the bus company's strategy, Betty Smith's lawyer calls to the witness stand Smith's blind passenger, Walter Wilson. Wilson then testifies that he heard the sound of the offending vehicle approaching Betty Smith's car, that the vehicle approached the car to a distance of no more than two feet, and that the vehicle was definitely a bus. On cross-examination by the Blue Bus Company's lawyer, Wilson testifies to his previous experience with perceiving the sounds of vehicles and inferring their size, nature, and distance from the sounds. Betty Smith's lawyer, in further support of Walter Wilson's testimony, then introduces two expert witnesses who bolster Wilson's account by reporting that laboratory experiments bear out the ability of blind people to determine the proximity and nature of vehicles on the basis of hearing alone, which is just what Wilson claimed to have done.

There is, of course, nothing more or less "direct" or "actual" or "real" about Wilson's primary aural sensory perceptions than about Smith's primary visual ones. Yet in considering what to make of Wilson's perceptions, we would naturally think that the validity of these perceptions depends on a process of generalization and noncertain inference. Wilson has perceived certain sounds in the past, and they have turned out to be buses. He has perceived distances in the past, and they have turned out

to be accurate. And so on. As a result, Wilson's inference from this sound to this conclusion (it is a bus at this distance) is an inference based on most but not necessarily all sounds of this type's having turned out in the past to be buses. This is a nonspurious but nonuniversal generalization—most but not all sounds like this are buses—that undergirds what appears to be a direct and thus individualized perception.

Though less obvious to those of us who are sighted, the process of making visual observations from what philosophers refer to as "sense-data" is conceptually no different in the case of visual observations than it is in the case of aural ones. And as the studies of the unreliability of eyewitness identification indicate, there may not be much of an empirical difference either, no matter how hard it is for those of us who are sighted to confront the possibility that, more often than we think, we should simply not believe our eyes. As a result, acknowledging the way in which seemingly direct observation involves a process of inference and generalization enables us to appreciate that even the processes that initially appear to us to be "direct," "actual," or individualized turn out to rely far more on generalizations from past experience than is often appreciated. Once we see that all evidence is in the final analysis probabilistic, the distinction between the probabilistic and the "direct," "actual," or "real" emerges as even more of an anomaly.

Not only are individualized assessments still based on probabilities and generalizations, but such individualized assessments are also always only partially individualized, omitting numerous dimensions of the particular case that might under other circumstances or other rules be relevant. Let us return to the real *Smith* case, and assume that what the Supreme Judicial Court was looking

for was testimony by Betty Smith that she actually *saw* the words "Rapid Transit" on the side of the bus that crowded her off the road. But even if this evidence had been forthcoming, Smith would not have been permitted, under well-accepted principles of tort law and evidence law, to testify to how much she needed the money from a recovery against Rapid Transit, to how easily Rapid Transit or its insurer could have afforded to pay the judgment, to how exemplary a life she had lived in the past, to how many times Rapid Transit had been found liable for the negligence of one of its bus drivers, or to the positive effect that even a mistaken judgment for Smith would have on bus safety in the Town of Winthrop. Yet in a truly particularist account of the events, in which we are not applying a legal rule but are simply trying to reach the most just result or to achieve the result that will maximize utility, none of these genuinely "real" facts would be deemed irrelevant, and all of them would be components of a fully individualized consideration of all the equities of the case.

So what are we to make of the fact that Betty Smith would not have been allowed to testify to some number of facts that a fully individualized determination might have allowed into consideration? If we accept the inevitability and desirability of not allowing her to present evidence of her own need, the company's insurance, and the like, then we can see that most so-called individualized determinations are not as individualized as we suppose. Moreover, the exclusion of these facts is itself something that occurs by virtue of the operation of a rule (in this case all the combined rules of tort and evidence law), and that consequently operates by virtue of a generalization. We exclude evidence of the plaintiff's poverty, the defendant's wealth, the existence or terms of insurance coverage, and the defendant's past negligent acts, among others, because it has been determined at some earlier time that these facts would not *as a rule* promote justice, or increase utility, or whatever. But because these are rules, we exclude the evidence even in the face of a showing in the particular case that admission of this evidence might serve justice, or might increase utility, or might promote some other goal that can be seen as one of the background justifications lying behind the exclusionary rules.[26]

In many cases the parties on one side or the other will argue that the exclusionary rules should be overridden in a particular case, and the exclusion wrought by an exclusionary rule is best thought of in presumptive rather than absolute terms.[27] Nevertheless, every piece of unadmitted evidence is typically unadmitted, whether consciously or not, by virtue of a rule. The rule will itself be based on a generalization about the usual or probable, but not universal, irrelevance of the excluded fact, thus further underscoring the way in which decisionmaking in a totally individualized or particularistic manner is essentially impossible.

That all seemingly particular or individualized decisions turn out to have important dimensions of generality is not totally to deny the logical distinction between the particular and the general. Although pressing against this distinction has a distinguished philosophical provenance, there is no need for us to examine here the deepest questions of metaphysics and philosophical logic bearing on the nature and existence of the distinction between the particular and the general, or the relationship between particulars and universals.[28] For our purposes, the commonsense distinction between a thing and a group of things will suffice. The only point here, an

important one, is that many of the things we perceive as particular objects or particular observations turn out to depend on the kinds of generalizations that, even if not on the same metaphysical status as true universals, are much the stuff of ordinary reasoning. This is still not to deny that there are important differences in degree between the more and the less particular and the more and the less general. Nevertheless, once we understand that most of the ordinary differences between general and particular decisionmaking are differences of degree and not differences in kind, we become properly skeptical of a widespread but mistaken view that the particular has some sort of natural epistemological or moral primacy over the general.

It turns out, therefore, that the Supreme Judicial Court's unwillingness to allow a jury to consider Betty Smith's case against the Rapid Transit company is a product of two significant mistakes: an overconfidence in the empirical reliability and even the very directness of direct evidence, and an underappreciation of the essential continuity between so-called indirect or statistical evidence and evidence that on its face appears to be more individualized and thus less statistical. The Supreme Judicial Court's skepticism about a "mathematical" case, therefore, even if the court was correct that this was a mathematical case, is, as we have seen, not so much a skepticism about mathematical or statistical evidence but a skepticism about resting legal decisions on nonspurious but nonuniversal generalizations.

Seen in this way, the Supreme Judicial Court's skepticism is of a piece with the skepticism of Plato's Stranger and of Aristotle about relying too heavily on what they called "laws," and with the inflammatory slogans of the pit bull sympathizers. In all these cases, the

preference for particulars is seen as a moral imperative. But if particularism itself relies on generalizations, and if particularized decisions provide no guarantee of greater reliability, then the foundations for the preference for particularism are shakier than they often appear.