Introduction

Discrimination is sometimes the product of what the discriminator takes to be a fundamental distinction between classes of people. For example, some religions treat men and women differently – in how they shall dress, in who can serve in the clergy, and in where they should sit during religious events – typically in the belief (whether sound or unsound is not the point, at least not yet) that the gender-based distinction is fundamental, foundational, natural, or constitutive. Those who distinguish between humans and animals for any of a number of reasons and in any of a number of contexts ordinarily understand the distinction between human beings and other creatures to be in some way ontologically or morally foundational. And placing a lesser value in war on the life of an enemy than on a country’s own citizens is typically premised on a belief in a basic distinction between friend and foe, or between citizens and non-citizens.

By contrast, there are other forms of discrimination that appear even to the discriminator to be less fundamental. These forms of discrimination nevertheless exist, and that is because they are perceived to be instrumental to some other goal. For example, most countries discriminate against younger (than age eighteen, say) people for purposes of granting permission to drink or to drive, but such discrimination is not based on the belief that setting a minimum age to drink or drive tracks some sort of basic distinction. Rather, the discrimination against the young (at least in this context) is justified by the assumption that the distinction between the younger and the older is a reliable, even if imperfect, indicator of the more fundamental distinction between the responsible and the irresponsible. In much the same way, many universities discriminate in admission against those who achieve lower scores on standardized tests because the universities believe that scoring better on such tests is a plausible predictor of academic success. And employers often discriminate against those who have been convicted of crimes, typically in the belief that a past criminal conviction is predictive of future criminal conduct in the course of employment.

This chapter is focused on the latter form of discrimination, which is often labeled statistical discrimination. The discussion will deal with those forms of discrimination that are alleged by the discriminator to rest on a statistically valid empirical distinction – a correlation – between
two (or more) people or categories. Statistical discrimination is thus defended by reference to the instrumental benefits of using a distinction that appears to have indicative or predictive value in pursuing a legitimate end existing independent of the discrimination itself.

Two types of discrimination

To discriminate is to draw a distinction between two or more people, places, things, categories, etc. And although the censorious use of the word “discrimination” is common, it need not always be so. When we non-pejoratively describe a person as having discriminating taste, for example, we mean that the person so described can distinguish the good from the bad, the true from the false, the fine from the crude, and so on. Here the identification of someone as discriminating is typically by way of praise and not condemnation. But although these days the language of discrimination is more often disparaging than commendatory, the basic point is that any form of discrimination, whether good or bad, presupposes drawing a distinction, and accordingly puts one or more items on one side of some line and one or more on the other side.

To discriminate is to distinguish (cf. the Introduction to this book).

With this basic idea of discriminating as distinguishing in hand, and in order to focus on statistical discrimination, we can then draw a distinction between non-statistical and statistical discrimination (Hellman 1998). As noted in the Introduction, the former is that variety of discrimination purporting to rely on some non-instrumental distinction of importance between classes of, typically, people. Many forms of gender discrimination, for example, are based on some person’s or group’s belief that different genders have different roles to play, or simply that the sexes are fundamentally different. So too with many forms of racial discrimination, where those doing the discriminating often base their discriminatory acts on the belief that racial mixing is foundational misguided because, they believe, the races are in some natural or foundational way different, or even that some races are superior to others. And many (most?) of the people who justify discriminating against non-heterosexuals do so on the belief that homosexuality and those who practice or prefer it are simply wrong.

What these forms of discrimination share is their non-instrumentality. The belief in the fundamental importance of gender or racial separation, for example, or in the foundational wrongness of homosexual practices and those inclined to engage in them, is most commonly justified in much the same way that people (especially non-consequentialists) might justify the wrongness of torture or the rightness of altruism. For the non-instrumental discriminator, the basis for the discrimination lies at or near bedrock, and is not justified by reference to further or deeper goals. Let us call this foundational discrimination.

To be contrasted with foundational discrimination is what we can label instrumental discrimination. Instrumental discrimination starts with some (presumably legitimate) goal that is itself independent of the discrimination at issue; but the discrimination is then based on the belief, assumption, or fact that drawing a distinction of a certain sort will, instrumentally, further the pursuit of that discrimination-independent goal. In other words, some discrimination is instrumental in the sense that possession of the trait that provides the basis for the discrimination is of significance to the discriminator only because of what it indicates about the likelihood that the person who is the subject of the discrimination possesses another trait. In a classic article, Joseph Tussman and Jacobus tenBroek (Tussman and tenBroek 1949) analyzed the same relationship under the rubric of an indicative instrumental connection between some “trait” and some “mischief” that the state sought to ameliorate, but, more broadly, the indicator need not be a natural human trait, and the goal need not be a mischief. Universities discriminating in favor of those with high test scores are an example, and there are many others. We might
believe, say, that it is better to have airline pilots who have good rather than poor vision. This goal itself is of course a form of discrimination, because it distinguishes good pilots from bad ones, and distinguishes pilots who see well from pilots who do not. But in most contexts these forms of discrimination are not taken even to be worth worrying about, precisely because the goals underlying the discrimination are so plainly legitimate.

Given the existence of a legitimate goal, and given the desire efficiently or effectively to pursue that goal, instrumental discrimination consists of the use of some further distinction that is believed to be statistically indicative of (and thus supportive of) the legitimate goal. The distinction serves as an imperfect proxy for something else, and it is the something else that is the real concern (Alexander and Cole 1997). Thus a relevant form of instrumental (or proxy) discrimination would be discrimination on the basis of age. To continue with the example of commercial airline pilots, those who support imposing a maximum age for commercial airline pilots (traditionally 55 or 60, but now increasing) believe that age is a reliable even if imperfect indicator of poor vision (as well as impaired hearing and slower reflexes) (Schauer 2003, pp. 108–30). Consequently, if we exclude older people from serving as commercial airline pilots we will wind up, ceteris paribus, with a commercial airline pilot cohort that has better vision, better hearing, and faster reflexes than we would without the exclusion.

Instrumental discrimination, which we can now re-label as statistical discrimination, is ubiquitous (cf. Chapters 27 and 28). Using age again as an example, but in a different context, the widespread exclusion of those under a certain age (16, 17, 18, or 21, most commonly) from buying alcoholic beverages or from driving is based on the belief that safe driving is better than unsafe driving. Again, the goal itself is a form of discrimination, but few people are concerned about discriminating against bad drivers or dangerous drinkers. Rather, the relevant discrimination here—the statistical discrimination—is distinguishing older from younger drivers, or drinkers, in the belief that youth is a statistically sound predictor of irresponsible behavior. If that belief is sound, then if we restrict driving to those, say, over sixteen years of age we will wind up with a population of drivers that is, again ceteris paribus, more responsible than if we did not impose the age-based filter.

Many other forms of discrimination, some desirable and some deplorable, fit this model (Alexander 1992). If one believes that it is good to have aggressive lawyers, for example, and if one believes that women are on average less aggressive than men, then excluding women from the legal profession may represent this kind of statistical discrimination, at least if we assume that the statistical discrimination is indeed an efficient way of pursuing the posited goal. Or if one believes that women are less able at mathematics than men, and if one believes that it is better to have accountants (or mathematics teachers) who are better rather than worse at mathematics, then excluding women would involve the same kind of statistical discrimination. So too with excluding people of certain races from certain professions in the belief that those excluded are less intelligent, or preferring people of certain races for certain professions or activities in the belief that race is a statistically valid predictor of some talent or ability thought necessary or useful for successful pursuit of that activity.

The empirical foundations of statistical discrimination

That some people believe that a statistical (or probabilistic) relationship exists between some proxy and what it is a proxy for does not mean that they are correct in so believing. Indeed, much of the history of pernicious discrimination is a history of beliefs about the existence of some supposedly valid statistical instrumental relationship that turns out to have no sound empirical basis whatsoever. Historically, for example, the cultural (and sometime official)
exclusion of women from serving as airline pilots was based on the belief that being female was a statistically reliable predictor of being worse at dealing with complex machinery, or being worse at reacting quickly and properly in case of emergency. We know now, however, that there is no basis for that belief, and that excluding women from the population of pilots would not produce a safer or more mechanically adept pilot population.

Such groundless correlations, which we can label as spurious, are widespread. Consider the empirical basis for astrology. Those who believe in astrology believe that being born under a certain sign is an indicator of certain abilities or characteristics. Capricorns, for example, are thought to be ambitious, and thus the hiring practices of a business looking to hire ambitious people might prefer (discriminate in favor of) Capricorns to those born under the other signs of the Zodiac. But although many people believe that astrological sign is indicative in this way, there is no evidence to support this belief, and much evidence to refute it. Being a Capricorn actually tells us nothing about ambition, and thus the statistical claim is spurious. So too with the nineteenth century belief in phrenology, the pseudo-science based on the principle that certain cranial shapes were indicative of certain attributes of intelligence, personality, and other non-physical characteristics. In form, phrenological or astrological discrimination in favor of (or against) people with certain astrological signs or head shapes is the same as discrimination in favor of or against those of a certain age with respect to drinking, driving, and piloting. The difference is that in the latter cases the statistical discrimination has a sound empirical basis, but in the former we now know that the empirical basis is totally absent. And when that empirical basis is absent, as with the examples just used or with beliefs about, say, the intelligence of people of certain races or the courage of homosexual men, the basis for justifiable statistical discrimination disappears.

The statistical relationship between being a Capricorn and being ambitious is spurious and therefore unreliable, but just what is it that makes a statistical relationship non-spurious, and what is it for such a non-spurious relationship to be reliable? Or, what is it to say that a non-spurious relationship is accurate? The easiest cases, of course, are those in which all or virtually all members of some class (or having some trait) have the further trait in which we are interested. If we were to prohibit those under the age of three from having driving licenses, for example, we could be confident (at least in this world) that all members of the class of those under the age of three possessed the relevant characteristic of being unable to drive safely. More realistically, if we were seeking people for some task or job that required total linguistic fluency in Mongolian, limiting the applicant pool to those born and educated in Mongolia would produce an applicant pool of those almost all of whom possessed the necessary ability. The correlation would not be perfect. It is possible to be born and educated in Mongolia and not be proficient in Mongolian, and even more possible to be born outside of Mongolia and have developed native fluency in the language. Nevertheless, being born and educated in Mongolia is a very strong albeit non-universal indicator of the relevant ability, and is thus, in a statistical sense, reliable.

Much the same applies to those instances in which a majority of members of the class possess the relevant attribute. To continue with the same example, most people born in Latin America are fluent in Spanish, but there are large numbers of people fluent in Spanish who were not born in Latin America, and an appreciable number of people born in Latin America who are not fluent in Spanish. Still, even if being born in Latin America is a less reliable indicator of Spanish language fluency than being born in Mongolia is of Mongolian fluency, it is safe to say that the majority of people born in Latin America are fluent Spanish speakers, and thus being born in Latin America is a more-reliable-than-random indicator of Spanish fluency, even if, again, the reliable indicator is far from perfect.
Less obviously, but more importantly, some traits may be statistically relevant (the word is important) indicators of some other attribute even if the attribute is possessed by less than a majority of those who possess the trait. Most Jewish women of Ashkenazi background do not possess the genetic makeup that makes them disproportionately prone to breast cancer, but that genetic makeup is more prevalent among Jewish women of Ashkenazi background than among the class of women generally, and thus being a woman of Jewish Ashkenazi background is a non-spurious indicator of susceptibility to breast cancer, even though the relevant genetic marker is possessed by far less than a majority of those in the designated class. Similarly, most pit bull dogs are neither aggressive nor dangerous, but the percentage of dangerously aggressive pit bulls is higher than the percentage of dangerously aggressive dogs. As a result, a dog being a pit bull is a non-spurious indicator of dangerous aggressiveness, even though, again, most of the members of the designated class do not possess the worrisome trait.

For the sake of clarity, we can designate the trait at issue as the indicator – Latin American, Ashkenazi woman, pit bull, Capricorn, etc. – and the attribute or action or behavior that the indicator is thought to indicate as the target – Spanish speaker, breast cancer susceptibility, dangerous aggressiveness, ambition, etc. And so an indicator is statistically reliable insofar as it in fact probabilistically indicates (or correlates with) the existence of the target. As long as there is a positive correlation between the indicator and the target – if the relationship is better than random – then we can say that the indicator is reliable or accurate, although the degree of reliability or accuracy will vary with the closeness of the correlation.

It is worth emphasizing that the important relationship is an indicative or correlative one, and that the soundness of a statistical indicator is independent of the existence (or not) of a causal relationship between the indicator and the target, or vice versa. In some context we might well be interested in the existence or non-existence of a causal relationship, but indicators can be statistically sound even if there is no causal relationship. To take a shopworn example, air conditioner use does not cause people to eat ice cream, and eating ice cream does not cause people to use air conditioners. There is a common cause (heat) of both that produces the correlation, but there is no causation between the two effects of the common cause. Still, if we were interested in predicting the level of ice cream use, knowing the level of air conditioner use would be helpful – indicative – for this task, despite the absence of a causal relationship. And thus the statistical soundness of an indicator in predicting the existence or level of some target is independent of the existence of any causal connection between them.

In the language of conditional probability, an indicator is thus statistically sound, statistically reliable, or statistically accurate if the existence of the target is more likely given the indicator than it is without the indicator. In the law of evidence, such a relationship is described as logical relevance, and evidence is deemed logically relevant, to use the language of the American Federal Rules of Evidence, if “it has a tendency to make a fact more or less probable than it would be without the evidence” (Federal Rules of Evidence, Rule 401). So too with statistical discrimination, and an indicator is non-spurious if knowledge of it (or evidence of it) makes the existence of the target more or less probable than it would be without knowledge of the indicator.

Accuracy and reliability are of course matters of degree. There are strong and weak correlations, and thus strong and weak indicators, and thus indicators that are more or less reliable, or more or less accurate. Depending on the costs and consequences of using a particular non-universal indicator for some social or policy purpose, the degree of accuracy of the indicator will be important, and so although it is valuable to recognize the historical and contemporary significance of spurious indicators, the degree of non-spuriousness – the degree of accuracy of an indicator – remains a crucial consideration in the determination of which indicators should be used and in which contexts.
The conception of statistical relevance just sketched bears an affinity with modern work in the philosophy of language focusing on generics (Leslie 2008; Leslie 2015). As we know, a common feature of language is the use of descriptive statements about a class that are not universally true, even though the non-universality of the statement is not literally apparent from the statement itself. When we say that mosquitoes cause disease, that Volvos are reliable, that mathematicians are clever, or that Swiss cheese has holes, we do so knowing that many mosquitoes do not carry disease, that there are unreliable Volvos and dim mathematicians, and that some Swiss cheese does not have holes. Nevertheless, such generic statements are part of our language and our conceptual apparatus, and, to oversimplify, a generic statement is true if it bears the same non-spurious statistical relationship to facts about the world that we have just been discussing. Generic statements may not on the surface be explicitly comparative, but the truth of a generic statement presupposes some background or reference class with which the generic is being implicitly compared. “Volvos are reliable” is true if the class of Volvos is more reliable than the class of all cars, but false if being a Volvo is no more predictive of being reliable than simply being a car. Likewise, “pit bulls are dangerous” is true if a dog being a pit bull makes it more likely to be dangerous than just being a dog (or a dog of another breed), but if pit bulls are no more likely to be dangerous than any other type of dog then the statement that pit bulls are dangerous contains an implication that turns out to be false.

The virtues and vices of statistical discrimination

Within the broad category of statistical discrimination, the soundness of a statistical relationship between the indicator and the target will ordinarily be a necessary condition for its legitimate usability, but it is far from a sufficient condition (Hellman 2008, pp. 114–37). Of course we should not underestimate the importance of testing the empirical soundness of an alleged relationship between an indicator and a target, in large part because much of the history of unfortunate discrimination is a history of the erroneous belief in statistical relationships that turn out to have no basis in fact. It is simply untrue that homosexual men have less courage than heterosexual men, despite generations of belief to the contrary, and it is no more true that women inherently lack aggressiveness than that non-Capricorns lack ambition, again despite generations of belief to the contrary. So although the identification and rejection of spurious relationships between indicator and target may seem straightforward as a philosophical or conceptual matter, identifying such spurious relationships remains morally, politically, and practically important.

That said, however, the crucial question at the vortex of the conceptual, moral, and policy dimensions of statistical discrimination is the question of how and when, if at all, to employ non-spurious but non-universal indicators for the purpose of making decisions that have real consequences for individuals. Given the conception of statistical soundness (or reliability or accuracy) just explicated, the question remains as to how and when, if at all, such statistically accurate but non-universal indicators should be used as the basis for official or personal discrimination.

Initially, it is valuable to recognize that some statistically sound indicators exist as the residue of previous spurious ones. Within the class of non-spurious indicators there is a sub-class comprised of those indicators whose statistical soundness is itself a product of previous and non-statistically justified discrimination. Consider, for example, the case of gender. Suppose, as the American state of Idaho supposed until 1971 (Reed v. Reed 1971), that the government is interested in the qualifications of those who administer the estates of deceased persons. And suppose as well, as Idaho supposed, that financial and business training and experience were
among the important qualifications for that task. Assuming, plausibly, that preferring those with
financial and business expertise was a legitimate governmental goal, then how should we
evaluate Idaho’s decision to use gender as a proxy for financial and business expertise, and thus
to prefer men to women to be administrators of estates?

Bearing in mind that this was 1971 (and earlier), it is difficult to deny that gender was then
a statistically sound indicator of financial training and experience. Then (and still, but to a much
lesser extent), few women possessed university training in business, economics, and finance, and
even fewer benefited from such training at the graduate level. Few women were certified as
accountants, and few held major positions in law or in the financial services industry. And so on
for all of the other plausible markers of financial or business acumen. And thus at the time the
policy was adopted, gender was in fact a non-spurious indicator of the skills reasonably thought
desirable in the administrator of an estate.

Yet although gender was a correlated and thus (somewhat) reliable indicator of the legitimate
target of financial and business expertise, it is almost certainly the case that this 1971 state of
affairs was the product of earlier and persistent discrimination based on the erroneous belief that
women were in some way genetically or naturally less able at financial tasks. Women were
steered away from certain skills and profession (law and business, most relevantly here) and into
others (grade school teaching and librarianship, for example) based on these beliefs. And it is
thus highly likely that the 1971 non-spurious relationship between indicator and target was
itself the product of an earlier and persistent belief in a relationship thought to be statistically
accurate but which was in reality spurious, or was instead the product of a now properly-
discredited belief in a foundational difference between men and women in the allocation of
social and professional roles.

Whether a non-spurious relationship that is the byproduct of an earlier belief in a relationship
that was in reality spurious (or was the product of a mistaken form of foundational discrimination)
should now be employed presents difficult questions. It seems morally attractive to believe that
we should not perpetuate the consequences of past wrongs, including past unwarranted
discrimination, but the question remains one of allocating the costs of remedying past wrongs.
If gender is now (or was in 1971) a non-spurious indicator of financial competence, as in the
Idaho example, then refraining from its use entails the cost of excluding a factor that might now
provide, *ceteris paribus*, useful information in pursuing a legitimate goal. And thus we can
consider the question of just who it is that should bear this increased cost. In this instance the
cost might turn out to be borne by the beneficiaries of less optimally managed estates, or, more
plausibly and more desirably, by the society as a whole in paying for a more costly individualized
system of assessment. Such allocations, especially the latter, seem appropriate in this and similar
instances, but it is nevertheless important to recognize that the unwillingness to employ the
most reliable (and efficient) indicator is not cost-free, and that the costs of that unwillingness
must be borne somewhere and by someone.

Somewhat less conceptually or morally problematic, although again unfortunately frequent,
are those reliable indicators that are employed as a pretext for some other form of discrimination.
Gender, again, is a common example. Although it is true that men have, on average, greater
upper body strength than women, and although it is hard to imagine (evolutionary explanations
going back thousands of years aside) that such a state of affairs is a product of prior discrimination,
employing this indicator in a hiring decision for a job in which upper body strength is not in
fact valuable would be just such a pretext. The underlying basis for using such a pretext might
be a belief in a relationship that is in fact spurious, as with using strength as a qualification for a
job in the computer industry, where the real motivation was a belief that women were not
mathematically adept enough to work in the industry. Or the real basis might be a form of non-
statistical discrimination. When the Virginia Military Institute attempted before the Supreme Court of the United States to justify its male-only policy by claiming that women were less willing to endure and benefit from the particular form of “adversative” training common in the military (United States v. Virginia 1996), it is more than plausible to suppose that the underlying basis for the decision was a non-statistical belief that women simply do not belong in the military, rather than the claimed (and arguably true at the time it was claimed) lesser willingness of women to accept a particular form of education and training (Case 2000).

“Pure” statistical discrimination and the argument for individuation

Moving on from statistical discrimination based on spurious correlations, and moving on as well from statistical discrimination that is the product of previous spurious correlations or a pretext for something else, we then confront the “pure” case of statistical discrimination. How then should we evaluate statistical discrimination whose sole goal appears to be regulatory optimization, as with the age restrictions for driving, drinking, and piloting, or, far more controversially, highway or airport targeting (profiling) of those from certain countries, or with certain ethnic backgrounds, or of certain races, assuming, for the sake of argument, that there is a statistical basis for such actions, and that the motivation is nothing other than regulatory optimization? (Risse and Zeckhauser 2004).

When the question is framed in this way, it is apparent that it cannot be answered without presupposing a moral framework within which the analysis takes place. Under a broadly utilitarian or cost–benefit framework (Risse and Zeckhauser 2004; compare Lippert-Rasmussen 2006), the gains of engaging in statistical discrimination, gains that plainly vary with the degree of accuracy of the indicator and the frequency of mistaken indications, will be weighed against various costs, including the costs of mistaken indications. The analysis is a decision–theoretic one, where the expected costs of the errors (call them Type I errors) of mistaken indication are weighed against expected costs of the (Type II) errors of mistaken non-indication. Such a decision–theoretic analysis must, however, take into account the full range of consequences. For example, is the harm to a 66-year-old pilot discriminated against because she is, by virtue of her age, mistakenly believed to have poor vision the same as the harm to a Muslim pilot discriminated against for exactly the same job because he is mistakenly believed to be sympathetic to terrorists? Even assuming, counterfactually, that the degree of accuracy of both indicators is the same, the harms to the individual mistakenly excluded will be vastly different given the different social meanings of the two forms of discrimination. All of this, and more, must be incorporated into a complete utilitarian (or more broadly consequentialist) analysis (Lippert-Rasmussen 2014).

Indeed, the full utilitarian or cost–benefit analysis will also take into account the likelihood of error for different forms of discrimination, and the spillover effects of different forms of discrimination. For example, will even statistically justified discrimination against some group be overused because of background and non-statistically justified assumptions by those tasked with enforcement? Even if, say, there are relevant physical differences between men and women with respect to some task for which the physical attribute is relevant, will background beliefs lead those making decisions to exaggerate both the physical differences and their relevance? If so, mandatory underuse may be necessary to prevent mistaken overuse. Similarly, will the salience and literal visibility of race and gender lead decision-makers to exaggerate the importance of even relevant indicators compared to other indicators that may be more accurate and more relevant, even if less visible? None of these questions can be answered acontextually, but they show that even under a utilitarian framework the analysis will be complex, and must
account for a wide range of positive and negative consequences in evaluating the use of even a statistically sound indicator.

Considering the question under various non-utilitarian frameworks is more complex yet, especially within frameworks that highlight the importance of equality as a pervasive good independent of consequences. Aristotle notwithstanding, equality is not about treating likes alike (Schauer 2003, pp. 199–207; Winston 1974). Rather, equality is an independent goal that requires treating unalikes alike. Just as the independent value of democracy will grant to all citizens the right to vote even in the face of knowing that some will vote more responsibly and knowledgeably than others, so too will the independent value of equality mandate treating people similarly even in the face of relevant differences. Equal access to the courts may require that those with weak claims for small amounts be treated, at least initially, the same as those with strong claims for larger amounts; equal access to public facilities typically requires that people be treated the same even if their qualifications to use the facility vary; and equal citizenship gives everyone a voice even if the quality of what is said is highly unequal. And thus when we evaluate statistical discrimination from a non-utilitarian or non-consequentialist framework in which values of equality, community, dignity, and respect, among others, have non-instrumental value, we find that treating everyone the same way in some contexts, or treating everyone with equivalent respect, dignity, and so on, are values that mandate ignoring even statistically relevant differences.

Some of the foregoing becomes more concrete in the context of contemporary debates about racial and ethnic profiling. Assuming that being a member of a certain group is at times a non-spurious indicator of a legitimate law enforcement or other governmental goal, and assuming as well that a pure cost–benefit analysis supports the discriminatory practice of focusing on members of those groups, it may be more important not to single out an already (or historically) discriminated-against group, for either consequentialist or non-consequentialist reasons, than to achieve some law enforcement or regulatory goal with maximum efficiency.

To repeat what was said above, however, refraining from statistically reliable discrimination entails a cost. Once we recognize that there is a cost to relinquishing the instrumental value of statistically non-spurious indicators, the question is how to allocate that cost. Will the cost of the comparative inefficiency of abstaining from using a reliable indicator be imposed on, and distributed among, everyone, as would be the case if we substituted targeted law enforcement screening with random searches and intensive examination of everyone? Will it be imposed on the best off or the worst off, as would be the case if government were to make available for a fee an exemption from intrusive screening? Obviously there are other possibilities, but the point remains that the avoidance of an instrumentally efficient and statistically justified discrimination comes at a cost that must be paid or borne by someone.

Typically, the chief alternatives to statistical discrimination have been understood to be individuation and randomization (Harcourt 2007). Individuation, or particularization, looks at each individual separately, attempting to determine whether that individual possesses the attributes that are the object of interest. An individuating approach would, for example, test every actual or aspiring pilot for vision, hearing, and reflexes, rather than relying on age as an indicator of decreased faculties on these dimensions. Such an approach would examine every dog for dangerous aggressiveness, instead of assuming that pit bulls were more dangerously aggressive than other breeds. And it would subject every airline passenger to the same scrutiny, rather than deploying more intensive scrutiny for those with certain ethnic, national, or physical characteristics.

In reality, individuation faces two obstacles. First, pure individuation is in the final analysis impossible. As the above discussion of generics in language shows, our language and our conceptual apparatus would be incomprehensible without generalizations, many of which are
probabilistic and not universal. And even those attributes that appear particular are in some way generalizations. I observe a brown cow, and identify it as such, but that identification is based on the fact that the visual bombardments of sense data that I now interpret to indicate “brown cow” are arrays of sense data that in the past have reliably indicated “brown” and “cow” and so I rely on them now, even though those past indications might not be true on this occasion. More realistically, consider the idea of engaging in a particularized evaluation of pilot vision, without employing the proxy of age. We could test each pilot, but those tests would be based on the generalizations – proxies – that one’s vision today is a reliable indicator of one’s vision tomorrow or next week or next month, and that one’s vision in the laboratory is a reliable indicator of one’s vision in the cockpit. We could eliminate the use of standardized tests to predict academic performance, but we would still wind up predicting academic performance in the future based on academic performance in the past. And if we did not use race or ethnicity or national origin in airport screening, we would use something else that was in itself based on some generalization.

Even more serious than the virtual impossibility of complete or even maximum individuation is the reality of the necessity of using proxies in a world of limited resources of time, personnel, and money. We could eliminate the usual and suspect proxies in airport screening by scrutinizing everyone more carefully, at considerable cost and with imperfect success, and could accommodate for the possible mistakes by tripling the number of security personnel aboard each airplane, but the costs of such an approach are obvious. And the police, rather than focusing on the “usual suspects,” itself a process of statistical discrimination, could investigate everyone whenever a crime was committed. More plausibly and more generally, every substitution of more rather than less individuation increases the costs of scrutiny, and also increases the possibility of error. Individuation requires individuators, and often the errors consequent on the use of imperfect proxies will, in some contexts, be less than the errors made by imperfect human beings attempting to take everything into account.

Recognizing the impossibility or the prohibitive cost of maximum individuation, some have suggested using randomization as a replacement for statistical discrimination (Harcourt 2007). Rather than using statistical indicators to select which taxpayers to subject to intensive audit, for example, the tax authorities could simply (as they do now, albeit not exclusively) conduct random audits, and the police could do much the same in deciding which drivers to stop for close inspection, and so on.

Although randomization would eliminate some of the problems associated with the non-universality of statistical generalizations, it would again hardly do so without cost. Some of these costs would stem from the perception that the application of state power should simply not be based on chance (Duxbury 1999, pp. 131–41). But other costs are more straightforward. If using a statistically reliable indicator allows us to train our attention and resources on those people, places, and things most relevant to some object of our concern, then relinquishing the use of the indicator, even if in the service of fairness or equality, is again not free. And at this point we can again no longer avoid the issues of just who statistical discrimination is discriminating against, and with what problems, and at what cost, and it is to this we now turn.

**Conclusion: the inevitability of statistical discrimination**

Statistical discrimination has been around for so long, and has so many benign applications, that until relatively recently it has rarely been a topic for discussion, academic or otherwise. Insurance companies, after all, are built around statistical discrimination – we call it actuarial decision-making – and in most contexts we have accepted that insurance companies will discriminate
against smokers, drinkers, homeowners in floodplains, young drivers, and workers who are engaged in dangerous occupations (cf. Chapter 28). And although many forms of accepted statistical discrimination involve attributes voluntarily assumed – smoker, drinker, race-car driver – not all of them do. Statistical discrimination on account of advancing age, for example, is widely accepted, perhaps because the data on decreasing faculties is so strong, and perhaps because almost everyone is either elderly or hoping to get there. And statistical discrimination on account of youth is even less often questioned, again in part because youth is wisely understood to be a temporary status.

When the basis for the statistical discrimination is race, ethnicity, national origin, religion (sometimes), gender, or sexual orientation, however, the nature of the discussion changes, and the implicit presumption in favor of statistical discrimination in many domains shifts to a presumption in the opposite direction. There is nothing inherently problematic about this differential treatment. But what it reflects is an important aspect of the very idea of discrimination. Individuals, groups, associations, and governments discriminate constantly. And it could not be otherwise. When some people are hired and others not, when some are admitted to universities and others not, and when in countless other contexts distinctions are drawn, the entity doing the distinguishing is engaged in an act of discrimination. Sometimes these acts of discrimination are irrational, relying on spurious correlations between indicator and target. But often the relationship is non-spurious, and it is simply implausible to imagine that we could or should eliminate or even be skeptical about the use of non-universal instrumental relationships throughout our decisional or policy-making existence. Rather, it is essential to recognize that the aversion to statistical discrimination in domains such as race, ethnicity, citizenship, gender, sexual orientation, and the like is not an aversion to statistical discrimination qua statistical discrimination. Such aversions are specific to the grounds for the discrimination, and to the historical, moral, and psychological dimension of some forms of statistical discrimination. When the history, the morality, and the psychology suggest special concern about discrimination on certain grounds, it will be appropriate either to elevate the standard of statistical accuracy necessary to permit their use, or perhaps even to prohibit use entirely. But to suggest, as some of the popular discourse and some of the more academic literature (Harcourt 2007) suggests, that statistical discrimination can or should be eliminated entirely is neither possible nor desirable.

Notes

1 It is common to refer to statistical generalizations of this sort as stereotypes, but the word is misleadingly ambiguous (Bernstein 2013). For some, the word refers to all statistical generalizations, for others it refers to all statistical generalizations about classes of people, for others it refers only to those statistical generalizations that are inaccurate, and for others it refers only to those statistical generalizations that, whether because of their inaccuracy or because of their other consequences, are worthy of condemnation. Because of this radical divergence in meaning, it seems better to avoid using the term entirely.

2 If what appears to be statistical discrimination is not in fact less costly (in the broad sense of “cost”) than examining the trait of primary interest, this may of course indicate that some more foundational discrimination is occurring.

3 “Spurious” as designating a claimed correlation for which there is no empirical support is consistent with ordinary usage, but is admittedly different from the technical use by statisticians to refer to a genuine relationship that nevertheless provides no basis for a causal inference.
Bibliography