The Annoyingly Indeterminate Effects of Sex Differences

Deborah M. Weiss
The Annoyingly Indeterminate Effect of Sex Differences

Deborah Weiss

Abstract

At present, men and women have different distributions of certain aptitudes and personality traits. A growing body of research suggests that some of these differences have some biological basis, although these distribution differences do not always conform neatly to traditional stereotypes, and cultural factors also contribute.

A heritable basis for difference, however, has frustratingly indeterminate implications. It suggests that some occupational segregation is not caused by discrimination. At the same time, a statistical aptitude or temperament difference will almost inevitably cause discrimination in professions that make use of that aptitude or temperament. This implication is supported by empirical evidence about the continued existence of discrimination from disparate sources, including labor market data and experimental laboratory studies in both economics and psychology.

The offsetting effects of difference and statistical discrimination make it extremely difficult to predict the degree of occupational segregation in a non-discriminatory world. To complicate matters further, there is no simple relation between the size of ability differences and the efficient degree of occupational segregation. The principle of comparative advantage suggests that even small ability differences might produce large degrees of occupational segregation. Conversely, high demand in a sector using skills that favor one sex may draw the other sex into sectors in which they have an absolute disadvantage.

Ability distribution differences create tremendous challenges for anti-discrimination policy. Liability rules based on current labor market representation are of minimal use, since the labor market reflects existing discrimination. Though perhaps useful in the short term, long-term numerical targets are problematic: target goals are extremely difficult to set, and would need to become a permanent institution. Moreover, numerical targets may create unintended problems for women. Men have historically dominated even fields in which women have an advantage. A target goal of equal representation would prevent women from attaining majority status in these fields. Closer scrutiny of employer practices holds promise as an alternative to numerically based theories.

Note: Earlier versions of this paper were circulated with the title “The Paradox of Statistical Discrimination.” This paper can be downloaded without charge from http://ssrn.com/abstract=1355585
http://works.bepress.com/deborah_weiss

1 Research Fellow, Center for Law, Business and Economics, University of Texas-Austin, deborahweiss@att.net.
Occupational segregation by sex is widespread. Most engineers are men, while most psychologists are women.\(^2\)

The observed distribution of certain abilities also differs somewhat between men and women. Some researchers believe that these differences are entirely environmental in origin. Others believe that some differences have a partially genetic basis, and in this Article I will argue that the weight of evidence supports this conclusion. Ability distributions for males and females overlap greatly, but they are not identical, and cannot entirely be explained by enculturation. To be sure, these differences are tremendously complex, and they do not always conform to sex role generalizations. For example, the much-discussed male advantage in math is concentrated in certain types of mathematical tasks, and in some tasks female have an advantage. Indeed, examination of ability distribution differences actually intensifies the puzzle of why women encounter a glass ceiling in their attempt to rise to top managerial jobs. In some important cases, though, the traditional division of labor between the sexes seems to be in part heritable.

Even those who agree that genetics plays some role are sharply divided. Some believe that genetic sex differences are large and account for much of occupational segregation, while others believe that the heritable component of differences is small and explains very little segregation, with the remainder resulting from discrimination.

The debate as thus framed, I will argue, misses an important consideration. A heritable component to ability distributions has frustratingly indeterminate implications. It obviously suggests that occupational segregation might not be caused by entirely by discrimination: men and women would not be evenly distributed among occupations even in a perfectly fair society. At the same time, an ability distribution difference between men and women will almost inevitably cause discrimination in professions that make use of that ability. Distribution differences tempt employers to engage in “statistical discrimination” – to use sex, consciously or otherwise, as a crude proxy for ability. Thus, a sex difference in the average level of a given aptitude will create incentives for discrimination in professions that make use of that aptitude.

Statistical discrimination creates a vicious cycle. Young people, aware of statistical discrimination, will choose educational and career paths that perpetuate occupational segregation and the underlying basis for statistical discrimination. Thus, even small genetic differences can be magnified by statistical discrimination into large and stable differences in labor market outcomes, with no easy way to separate the contributions of each factor. Evidence that genetic differences are large should not reduce concern about this problem but rather increase it. The larger the initial difference, the more likely statistical discrimination is. Evidence of statistical discrimination comes from disparate sources, including labor market data and experimental laboratory studies in both economics and psychology.

Current Title VII doctrine clearly prohibits statistical discrimination. The normative justification for this is open to debate but there are plausible grounds to continue the prohibition. Unfortunately, ability distribution differences and the resulting permanent state of statistical discrimination create tremendous challenges for anti-discrimination policy. Liability rules based on current labor market representation are of limited use, since the labor market reflects existing discrimination. To further complicate matters, the existence of differing ability distributions implies that sex discrimination is not a problem that can ever be eliminated once and for all. Rather, it is a permanent feature of labor markets.

Numerical targets may be useful in the short run, but are problematic in the longer run. To compute the nondiscriminatory equilibrium distribution of men and women among professions would require detailed information about the precise nature of sex differences; the abilities required in various occupations; and the demand for products in all output markets. Moreover, long-run numerical goals would need to be a permanent institution, which current law prohibits, and these targets might create unintended problems for women. Statistical generalizations about ability are not the only factor that has impeded women in the labor force. The view that women belong at home has meant that men have historically dominated even fields in which women have an advantage. A target goal of equal representation would prevent women from attaining majority status in these fields.

Closer scrutiny of employer practices holds some promise as an alternative to numerically based theories. More structured personnel practices have great potential to reduce statistical discrimination at relatively low private and social costs. Whatever path reform takes, however, must be grounded in empirical evidence of the forces that underlie discrimination.

I. THE INTERDEPENDENCE OF DISCRIMINATION AND DIFFERENCE

The sexual division of labor is as old as human history, and some occupational segregation seems to be a universal in all societies to date. The character and degree of this division, however, varies somewhat across time and place. The term “segregation” in this context is purely descriptive, referring only to the different distribution of men and women across occupations. It carries no suggestion that discrimination is the cause of that distribution.

In studies of industrialized societies, the so-called dissimilarity index is the number most commonly used to measure occupational segregation. This index indicates the proportion of women or men who would have to change occupations to produce a fully integrated market workforce. The number is symmetrical and does not depend on whether women or men are used as the reference point. A dissimilarity index of 100% indicates complete segregation, while an index of 0% corresponds to full integration.

Studies based on comprehensive workforce data have concluded that between 1900 and 1970, the U.S. dissimilarity index was more or less constant at about two-thirds, and perhaps even rose slightly. A sharp break with the past occurred in the 1970’s, evidenced not only by changes in the dissimilarity index but other measures, such as the labor force expectations of young women. The index dropped at a rate of about .8% per year during the 1970s, .42% per

year during the 1980s, and .36% per year during the early 1990s. The rate of occupational desegregation seems to be slowing, although the level of occupational segregation is still high. About fifty percent of all women (or of all men) would have to change jobs to integrate the workforce fully.

Whether occupational segregation is a reason for concern depends on its causes. Public discussion of this issue typically focuses on two possible causes, difference and discrimination. These categories can be usefully broken down further, and turn out to have a complex interdependence.

A. BIOLOGICAL DETERMINISM

Any discussion of sex differences raises questions of the relative role of genetics and environment. For some, any reference to genetic factors raises the specter of genetic determinism, the belief that genes alone determine human traits and behaviors. Such determinism has in fact been common in various times and places, often with ugly consequences. No scientist today propounds this view, and virtually all would agree that genes and environment interact in complex ways. None disputes that environment plays a role in observed sex differences and all agree on the existence of some heritable differences between males and females, if only anatomical ones. The debate concerns how many, if any, non-anatomical traits have a heritable component; how wide a range of outcomes can be produced by different environments; and the mechanisms that interact to produce the differences observed in various societies. This current consensus is sometimes misstated to say that there is no meaningful distinction between genetic and non-genetic factors. The distinction is meaningful, but it is complex and incompletely understood.

To stress the plasticity of gene-environment relationship, I will sometimes use the term “heritable” rather than “genetic”. In biology, “heritability” is the proportion of observed variation in a population that is attributable to genetic variation among individuals. Yet “heritable”, I hope, runs less risk of conjuring up concerns about biological determinism than “genetic”. It has little history of abuse by biological determinists and the suffix “able” conveys that it means only “capable of being inherited.” For reasons of accuracy as well as history, I will not equate the word “biological” with “genetic” or “heritable”: a physiological phenomenon, such as diabetes, can be called biological, although it is the result of both heritable and environmental forces.

B. THE NEOCLASSICAL FRAMEWORK

Neoclassical economics offers a helpful framework for analyzing possible causes of segregation, though its neat categories may blur somewhat on close inspection. The neoclassical framework suggests three possible sources for segregation: differences in ability, differences in tastes (also called preferences), and labor market discrimination.

---

7 Jacobs, Long Term, supra note ___.
The neoclassical paradigm generally treats tastes as fixed rather than acquired. In this setting, the terms “fixed” and “acquired” have a technical meaning, and the fixed-acquired distinction does not correspond to the genetic-environmental distinction. “Fixed” means simply “determined outside the economic system” and, most importantly, not determined by the choices of the individual in question. From a neoclassical perspective, fixed taste differences may be either heritable or the product of pre-market environmental factors, including what a non-economist would call discrimination. Social pressures during childhood may encourage members of one sex to develop tastes for activities deemed sex-role appropriate by that society. Although economists have developed a few models of how tastes are formed, none to date seems helpful in understanding the genesis of sex roles.

The economic literature is somewhat more helpful in understanding factors that contribute to ability. The neoclassical model assumes that ability, also called human capital, may be either heritable or acquired. Acquired ability is the product of choices by workers or their parents, such as the decision to acquire education. A great deal has been written on how expectations about the labor market affect ability acquisition decisions. Economists have had less to say about the diffuse social forces, including certain kinds of discrimination, that shape individual attitudes and expectations and thus influence job performance and ability acquisition decisions.

The final force producing occupational segregation is labor market discrimination, about which economics has a great deal to say. Discrimination in the labor market may result from irrational generalizations; the tastes of customers, coworkers or employers; or rational generalizations. This taxonomy differs slightly in its phrasing from that found in other discussions, though it is essentially the same in substance. The often-used term “stereotype” has different meanings in different disciplines and in common usage. For this reason I generally avoid it and also generally avoid the term “bias,” which creates similar confusion. Instead, I use the term “generalization,” and classify these by type. Generalizations about a group may be true or false, and rational or irrational. False generalizations are rational if they are based on the best cost-effective information available. False generalizations are irrational if they are not updated based on information that is either available or could be acquired at a cost justified by the benefits of the information. True generalizations will be treated as always rational.

Irrational discrimination is universally agreed to be undesirable, since it hurts not only those against whom it is directed, but also the discriminators. The normative analysis of taste-based discrimination depends how much deference should be given to individual tastes. The taste

---

10 A survey of these problems may be found in Gregory Mitchell & Philip E. Tetlock, Antidiscrimination Law and the Perils of Mindreading, 67 OHIO STATE LAW JOURNAL 1023 (2006). For a discussion of usages in legal settings, see Mary Anne Case, "The Very Stereotype the Law Condemns": Constitutional Sex Discrimination Law As a Quest for Perfect Proxies, 85 CORNELL LAW REVIEW 1447 (1999). In social science, the term has a less normative overtone. See, e.g., STEREOTYPE ACCURACY: TOWARD APPRECIATING GROUP DIFFERENCES, (Yueh-Ting Lee, Lee J. Jussim, & Clark R. McCauley eds., 1995).
12 True generalizations may in theory be either rational or irrational, but seldom have troublesome consequences that policy must confront. Moreover, it is hard enough to be right even when one is rational, so irrational true generalizations probably are not that common.
most often regarded as problematic is animus,\textsuperscript{14} which Title VII clearly regards as an impermissible motive. The law, in accord with many commentators, regards discrimination based on tastes other than animus with skepticism though not with complete condemnation.\textsuperscript{15}

The negative connotation of the term “discrimination” derives from the prevalence, for many years, of sex and race discrimination based on animus or irrational generalizations. The term “discrimination,” however, may also mean simply “distinction”, and it is this descriptive meaning which is intended in this article. Although much discrimination is undesirable, that judgment should not be presumed simply because a distinction is made.\textsuperscript{16}

Statistical discrimination is, like taste-based discrimination, a type of rational discrimination. In statistical discrimination, employers use group membership as a proxy for unobservable traits that affect work performance. Statistical discrimination based on sex or race is largely prohibited by employment law.\textsuperscript{17} Some scholars argue that such discrimination should be permissible because it is economically efficient.\textsuperscript{18} Others suggest either that it is not efficient, or that distributive concerns should override efficiency concerns.\textsuperscript{19}

Most researchers agree that all three factors – discrimination and the worker’s own tastes and abilities – play some role in occupational segregation. For the most part, these factors are treated as distinct and mutually exclusive. To the extent that the worker’s own tastes are shown to determine segregation, it is assumed that discrimination and ability are not playing a role: If at least 30% of segregation is determined by tastes, then no more than 70% can be determined by ability and discrimination.

Yet these three factors do interact. The most widely discussed interactions between are those between acquired ability and the labor market.\textsuperscript{20} Another interaction between causal forces has received little attention. Especially outside of economics, discussion of discrimination often focuses on animus or irrational generalizations.\textsuperscript{21} As a result, the role of statistical discrimination is often overlooked, and an important connection between such discrimination and differences in ability distributions has gone unnoted.

\textsuperscript{16} Perfect Proxies, \textit{supra} note___; \textit{An Economic Analysis of Sex Discrimination Laws}, \textit{supra} note___.
\textsuperscript{17} City of Los Angeles Department of Water and Power \textit{v.} Manhart, 435 U.S. 702 (1978); 29 C.F.R. 1604.2.
\textsuperscript{18} Forbidden Grounds, \textit{supra} note___. \textit{The Efficiency and the Efficacy of Title VII}, \textit{supra} note___. Both Epstein and Posner acknowledge the ambiguous efficiency of statistical discrimination. Epstein, however, clearly feels that the net efficiency effect is positive. Posner is more tentative but his textual discussions suggest skepticism about the wisdom of prohibiting it.
\textsuperscript{19} See Section I.B and VI.A.
\textsuperscript{20} See Section I.C.2.
\textsuperscript{21} Recent searches in the JSTOR database found that, of articles that included the terms “discrimination” and “race OR sex” the percent that also contained the terms “statistical discrimination OR rational discrimination” was 3.8% in economics, 2.3% in sociology, 1% in law, and .7% in women’s studies.
C. **The Indeterminate Effects of Sex Differences**

The existing scientific evidence suggests a heritable component to some sex differences in the distribution occupationally relevant abilities. The implications of a role for heredity, however, are less clear. Evidence that an ability is in part heritable is often offered to suggest that discrimination plays little role in the occupational segregation of fields that make use of that ability. Such evidence of heritable difference does tend to rebut, though it does not wholly disprove, inferences from segregation to discrimination based on tastes or irrational generalization. But even large differences in heritable ability distribution would not imply that discrimination plays no role in occupational segregation. If an employer has imperfect information about individual productivity, and the distribution of productivity is different between men and women, a rational employer, with no animus towards either sex, will use the evidence of average sex differences as a proxy for the ability of a particular individual.

1. **The Benchmark Model**

The basic mechanism of statistical discrimination was first described by Edmund Phelps and subsequently expanded by Dennis Aigner and Glen Cain. In this benchmark model, ability varies among workers, and employers care about the ability of the workers that they hire. Ability is assumed to be fixed, and the question of whether differences are heritable or environmental is not important. Employers cannot observe ability directly, but instead observe some simple objective measure usually referred to as a test score. The value of the test score predicts an applicant's ability with error. The employer's best estimate of an applicant's ability is a weighted average of the mean ability for all applicants and the applicant's score. The relative weight assigned to the test score and the group average depends on how well the test predicts ability. The more predictive the test, the more weight it receives relative to the group average.

\[ q = (1 - \gamma) \alpha + \gamma y \]

where \( \alpha \) is the mean quality for all applicants, and \( \gamma = \text{squared coefficient of correlation between } y \text{ and } q \).
Now suppose that the employer has one additional piece of information. Applicants are either female or male, and the two groups have different mean values of ability. The employer thus estimates the qualifications of each individual male and female by examining a weighted average of the group mean and the individual score. The employer will estimate a lower productivity for a low-ability group candidate with a given score than for a high-ability group candidate with the same score. The greater the difference in the group averages, the more discrimination will occur.

Sex differences thus imply two opposing forces. One implies that occupational segregation is caused by the differences themselves, while the other implies that they are caused by discrimination. Such evidence does tend to reduce the likelihood that occupational segregation results from tastes or irrational generalization, but it does not prove that discrimination plays no role in occupational segregation. Rather, it implies that employers have an incentive to engage in at least some degree of statistical discrimination.

2. The Dynamics of Statistical Discrimination

The discrimination literature has long debated whether markets will eventually drive out discriminating firms. Market forces do tend to reduce certain types of discrimination, such as irrational discrimination. Irrational firms are less efficient and will over time tend to fail. Market forces do not, however, tend to drive out other types of discriminators, such as those who cater to discriminatory customer preferences, since these firms effectively maximize profits.

The simple benchmark model of statistical discrimination makes, implicitly or explicitly, a long set of assumptions, including fixed ability, fixed employer information, and fixed labor supply. These assumptions preclude both efficiency effects and change over time. The consequence of prohibiting or allowing statistical discrimination is entirely distributional. A lower mean score for one group causes all members of the group to receive a lower wage than they would with complete information. The model has no dynamic tendency to move to a new equilibrium, but it is also efficient in its initial one.

The dynamic analysis of statistical discrimination drops the benchmark assumption that ability levels are fixed, and instead assumes that these levels have a heritable and an acquired component. Workers choose to acquire ability based on their future job market prospects. Efficiency generally requires that people with higher heritable ability also acquire more ability.

\[
\hat{q}^f = (1-\gamma) \alpha^f + \gamma y \\
\hat{q}^m = (1-\gamma) \alpha^m + \gamma y
\]

where superscripts \(f\) indicate values for females, and \(m\) for males.

26 For this reason, some commentators do not believe that the undesirability of irrational discrimination justifies prohibiting it. Forbidden Grounds, supra note___; The Efficiency and the Efficacy of Title VII, supra note___. But see John J. Donohue III, Is Title VII Efficient?, 134 UNIVERSITY OF PENNSYLVANIA LAW REVIEW 1411 (1986).


With acquired ability, statistical discrimination can occur even if both groups have identical heritable ability distributions. The long-run market equilibrium depends on the acquired ability distribution when the market starts. If members of two groups start with efficient levels of acquired ability, future members of both groups will continue to acquire efficient levels of ability, yielding a stable and efficient market equilibrium.

This equilibrium, though, is not unique. Suppose instead that for historical reasons, members of one group initially have inefficiently low acquired ability levels. This initial deviation from efficient ability levels perpetuates itself: even employers without any animus correctly observe that members of the disadvantaged group have less ability and the employers therefore statistically discriminate against them. This discrimination in turn discourages members of the group from acquiring ability, perpetuating the true generalization about ability. This result may hold even when employers use sophisticated strategies for updating their assessment of ability distributions. These equilibria will generally be stable, showing no tendency to converge to an efficient allocation.

Anti-discrimination policies can sometimes move the economy to the non-discriminatory equilibrium. These policies change incentives for acquiring ability, the initially disadvantaged group obtains more education, and differences diminish. Eventually the optimal stable equilibrium is reached, and there is no further need for anti-discrimination rules. The assumption of identical heritable ability distributions is a helpful way of thinking about types of discrimination, such as race, where current ability levels diverge between groups only because of historical forces. The assumption is not appropriate in analyzing sex discrimination if sex differences have a heritable component. Some ability distribution differences will be attributable to differences in acquired ability, and anti-discrimination policy may reduce these differences by changing the incentives of both groups to acquire ability. Yet even if antidiscrimination policy moves the labor market towards an efficient equilibrium in acquired ability, that equilibrium

Subsequent analysis was based on the model developed in Stephen Coate & Glenn Loury, Will Affirmative-Action Policies Eliminate Negative Stereotypes?, 83 AMERICAN ECONOMIC REVIEW 1220-1240 (1993).  


never becomes self-sustaining. Since heritable differences never disappear, employers have a
continued incentive to discriminate statistically. At best, policy can permanently eliminate
certain historically contingent elements of discrimination, leaving untouched the statistical
discrimination associated with the difference in heritable ability distributions. With heritable
differences, the indeterminate effect of of statistical discrimination is an immutable fact in the
labor market. Perpetual difference implies both nondiscriminatory segregation and permanent
discrimination.

3. **Personality Traits: Taste or Ability?**

Most models of statistical discrimination examine discrimination based on ability
distributions. Statistical discrimination may also be based on differences in worker tastes, such as
a taste for certain kinds of work over others, but the extension from abilities to tastes is not
straightforward, since both the informational conditions and the party’s incentives are different. Because of these complications, in this paper I examine only traits that can be classified
unambiguously as abilities. This precludes examination of statistical discrimination based on
personality traits, such as risk-taking and competitiveness, even though these have a significant
impact on labor market outcomes such as wages. Personality traits are often not easily
categorized as either preferences or abilities. Sometimes, as in the case of risk, the theoretical
line is clear but hard to apply in practice. In other cases, the line between tastes and abilities may
not be as sharp as neoclassical analysis suggests. One source of fuzziness results from the
complex association between what people enjoy and what they are good at. For example, people
who are good at competition often enjoy it. Although possible sex differences in personality are
critical to the understanding of labor market outcomes, they are beyond the scope of this paper.

II. **Evidence of Difference**

Some observed abilities are distributed differently in the current populations of males and
of females. Evidence suggests that these differences are in part heritable and in part
environmental, although the relative contribution is hard to ascertain.

A. **Potential Explanations For Ability Distribution Differences**

Explanations for currently observed differences in ability distribution can be classified
along a continuum from purely heritable to purely environmental. For expositional purposes I
will compare theories that place relatively heavy weight on the environment with those that place
relatively heavy weight on heritable factors, with the understanding that all scientists today reject
simple extreme theories.

---

34 See, e.g., Gerritt Mueller & Erik Plug, *Estimating the Effect of Personality on Male and Female
35 For a more extensive examination see Bedrooms and Boardrooms, *supra* note 3.
1. **Environment**

The simplest environmental theory of sex roles proposes that current sex role patterns result from chance events that occurred independently in different societies. Another theory that assigns a large role to chance is the common ancestor hypothesis, which suggests that chance determined sex roles in a single prehistoric society from which all modern societies have descended.\(^{36}\)

In contrast, functional theories attempt to identify functions that sex roles serve, or at least might have served at some time. Recent functional environmental theories, notably by Alice Eagly and Wendy Wood,\(^{37}\) begin by taking as given a small set of heritable sex differences, the most important of which are the different physiological roles of males and females in reproduction. Sometimes greater male physical strength is also taken as given.\(^{38}\) In foraging societies these few heritable differences constrained the roles that male and females are best suited to play, and made non-heritable behavioral adaptations valuable. Because of lactation, forager females were inevitably more involved than males in the raising of young. These females could only engage in certain kinds of productive activities, such as gathering, that could be done while watching over children. Only males could afford the travel and risk involved in hunting. Men and women therefore were socialized to specialize in subsistence activities that were suited to their respective reproductive roles.

All environmental functionalism acknowledges the role of minimal sex differences in reproductive roles, but focuses on the role of cultural forces in exaggerating these differences and perpetuating sex roles beyond the circumstances that made them useful. For most of our history, human beings have been foragers, and sex roles are adapted to this simple technology. The ten thousand years since the Neolithic agricultural revolution have witnessed almost incomprehensibly rapid technological change, leaving little time for cultural adjustment to new circumstances. Sex role socialization may encourage each sex to develop abilities that were valuable in the early human environment, but the function of this socialization might now be obsolete.

Many environmental theorists believe that patriarchal institutions also contribute to observed sex differences, which can be explained as the result of socialization intended to subordinate women. A variety of theories have been offered to explain how patriarchy, though pervasive, is not inevitable. Some suggest that physiology has relegated women to a subordinate role from the time of human origins, perhaps because of the burden that reproduction places on females\(^{39}\) or the mechanics of heterosexual intercourse.\(^{40}\) Like the division of labor, these theorists suggest, patriarchy has been rendered obsolete by new technology, especially reproductive technology. Other environmental functionalists have argued that men and women initially lived in a state of separate but relatively equal spheres, with the institution of patriarchy arising relatively recently, at more or less the time of the emergence of settled communities. The rise of patriarchy has been attributed by some to the institution of inheritance and consequent

---


\(^{38}\) *Id.*


increased importance of paternity certainty, and by others to the rise of quasi-market institutions and the commodification of women through exchange.

Regardless of the origin of sex roles, their persistence may be reinforced by the coordination problem described in an important paper by Gillian Hadfield. Most individuals will choose to mate with an individual of the opposite sex. If there are gains to specialization within a household, each member of a couple should invest in different abilities. Ability investments usually take place before individuals or their families identify mates, creating an incentive to choose the ability investment which will maximize potential mates. This coordination problem will tend to produce a highly stable equilibrium in which each sex specializes in one task in order to ensure abilities that are complementary to those of any future mate. If neither sex has any heritable comparative advantage at either task, chance determines the initial allocation and coordination makes that allocation stable. Heritable differences, even small ones, will be magnified by the coordination problem.

The persistence of sex roles may also result from boundedly rational cognitive processes. A generalization may be ‘irrational’ in the sense that it is not fully updatable, and yet still be the result of a heuristic that serves some function. One such heuristic might be a tendency to essentialize sex roles by viewing them not merely as a complex set of different trait distributions but as a neat package of universal and inextricably linked traits. Miranda McGowan has explored the cognitive psychology of the essentializing impulse, noting that the experimental literature indicates that environmental conditions can dampen or eliminate the tendency to make race distinctions but not sex distinctions.

2. **Heritability**

The simplest genetic account of sex roles looks for heritable sex differences in physiology and attempts to relate them to observed behavioral differences without explaining how such differences came about. Yet some explanation is called for, since chance genetic variation can occur in separated populations, but seems highly unlikely in two populations, male and female, which by definition interbreed each generation.

The most important functional genetic explanation for sex roles is the adaptationist approach of evolutionary psychologists. Like functional environmental theories, adaptationism assigns a crucial role to the reproductive division of labor between the sexes: pre-modern women were compelled by their mammalian physiology to care for young children, and thus to undertake subsistence tasks like gathering that were compatible with child care. Unlike environmental theorists, however, evolutionary psychologists suggest that different reproductive

---

45 Miranda O. McGowan, Engendered Differences, SSRN eLIBRARY 1361196 (2009). Regrettably, this essentializing impulse is shared by academics familiar with the relevant research, who should know better. Id.
roles caused selection pressures that in turn produced heritable differences between men and women.

Evolutionary psychology also assigns a major role to selection pressures favoring organisms that successfully reproduce, a factor only occasionally discussed in environmental theories. Male competition for female reproductive resources favors traits that are useful in eliminating male competitors or in persuading or coercing females to mate. Male parental investment is valuable but human paternity is uncertain, providing human males with a motivation to exert oppressive control over the behavior of females with whom they mate. Evolutionary psychology thus suggests a distinct theory of patriarchy.  

Extreme adaptationists insist that all heritable behavior be explained in adaptive terms. This view is increasingly rare, and more pluralistic versions of evolutionary psychology consider explanations besides adaptation. Some heritable sex differences might themselves fulfill no adaptive function, but rather be a by-product of some mechanism that produced a functional difference.

The existence of heritable predispositions is perfectly consistent with cultural mechanisms for shaping behavior. Two possible interactions between environment and heredity are especially interesting in the analysis of occupational segregation. First, the size of small heritable differences might be magnified by the coordination problem. Second, much of current evolutionary psychology suggests that even among foragers, optimal reproductive strategies varied among different environments. This variation should have favored the inheritance of flexible behavior strategies that adapted well to at least some range of changing conditions. These adaptations would in turn be transmitted culturally.

**B. PERFORMANCE EVIDENCE OF DIFFERENCES**

Sex differences in ability distribution turn out, on closer inspection, to be considerably more complex than the simple dichotomies found in popular culture. Patterns among these differences shed light on theories of the origin of these differences and support roles for genes and environment, function and by-product.

**1. GENERAL INTELLIGENCE**

Psychologists disagree about the relative importance of generalized versus specialized ability. Many believe that general intelligence, sometimes called $g$, is an important predictor of success in most occupations. The measurement of general intelligence is a complex matter, but

---

51 Frank L. Schmidt & John E. Hunter, *General Mental Ability in the World of Work: Occupational Attainment and Job Performance*, 86 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 162-173 (2004). For a review including some misgivings, though no complete skepticism, about the use of $g$, see Chockalingam
most current research finds little or no difference in the mean level of general intelligence for men and women.\textsuperscript{52}

Other psychologists emphasize the importance of specialized abilities in occupational success. Even those who believe in the importance of \( g \) agree that in many occupations, specialized abilities are also valuable. Some of these abilities are unambiguously cognitive; some are physical; and some have elements of both. In a number of measures of these abilities, men and women perform similarly, and it is probably fair to say that men and women are more similar than different in their cognitive abilities.\textsuperscript{53} However, men and women on average do perform differently on tests of some specific abilities, both physical and cognitive, that might be relevant to occupational specialization.

2. \textit{Some Non-Evidence and a Cautionary Note}

Males have accounted for a vastly disproportionate amount of the scientific and cultural output of the human race. From this some observers have inferred lower ability levels in women. No such inference can be drawn, since the history of women’s subordination could clearly account for much or even all of the discrepancy. The English-speaking world has been among the first to reduce obstacles to women’s advancement, but legal impediments existed well into the twentieth century. In both the United Kingdom and the United States, higher education for women did not begin until the mid-nineteenth century, primarily at single sex institutions. Many elite institutions did not admit women until the 1970’s.\textsuperscript{54} Married women were, with narrow exceptions, not permitted to own property until the nineteenth century.\textsuperscript{55} Women did not obtain the right to vote in most American states until 1920 or in the United Kingdom until 1928. No discussion of the history of women’s occupational status would be complete without the famous words of the Supreme Court in \textit{Bradwell v. Illinois}:

\begin{quote}
[T]he family organization, which is founded in the divine ordinance, as well as in the nature of things, indicates the domestic sphere as that which properly belongs to the domain and functions of womanhood. The harmony, not to say identity, of interests and views which belongs, or should belong, to the family institution is repugnant to the idea of a woman adopting a distinct and independent career from that of her husband.\textsuperscript{56}
\end{quote}

The relative achievements of men and women in earlier generations are of no value in determining the relative abilities of the two groups. The past is relevant only because its effects must be disentangled from present day evidence of ability distribution differences.

---

\textsuperscript{52} Viswesvaran & Deniz S. Ones, \textit{Agreements and Disagreements on the Role of General Mental Ability (GMA) in Industrial, Work, and Organizational Psychology}, 15 Human Performance 211-231 (2002).


\textsuperscript{56} Bradwell v. Illinois, 83 U.S. 130 (1872).
3. **Motor Skills**

A motor skill is the ability to complete a voluntary movement to execute a prescribed task. One of the largest differential ability distributions is found in targeting skill, both throwing and intercepting. Men consistently outperform women on these tasks. Women perform better than men on many fine motor tasks, beginning at an early age and even when finger size is not a factor.

Interview studies suggest that adult differences do not appear to be related to environmental factors such as participation in sports in which these skills are useful. Some differential ability distributions appear at relatively early ages. Differences in targeting ability are evident in three to five year olds. This early emergence tends to support a heritable component.

4. **Spatial Abilities**

Psychologists have not reached a consensus on the best method of defining and categorizing spatial abilities. Sex differences can be discerned in individual tasks, but the characterization of a pattern in overall differences is still in dispute.

Males in a wide variety of cultures, including African, East Indian and Asian, score higher than females on tests of spatial rotation. The sex difference diminishes when meaningful rather than abstract objects are to be rotated. Spatial rotation abilities appear to be associated with the ability to reason about the behavior of moving objects. Studies in the industrialized world show that males have a moderate advantage in determining spatial relations despite

---


59 Nontrivial Sex Differences in Throwing and Intercepting, supra note 47.

60 Susan C. Levine et al., *Early Sex Differences in Spatial Skill*, 35 DEVELOPMENTAL PSYCHOLOGY 940-49 (1999); DOREEN KIMURA, SEX AND COGNITION 35 (1999).


distracting information. This male advantage has been confirmed in African societies but not in Inuit Eskimo groups. In what may be a related phenomenon, men and women on average take different approaches to navigation. Women are more likely to use relative directions (left and right) and landmarks while men are more likely to use absolute reference points such as north, south, east or west. If these differences were purely environmental, the less experienced sex should gain more from additional training, since training presumably has a declining marginal value. In spatial rotation tasks, studies suggest that both sexes benefit equally from additional training.

Women are on average better at remembering the location of objects in an array; at most tasks involving color recognition or discrimination and vividness of visual imagery; at depth perception; and at tasks of perceptual speed, which require subjects to match a given item with the identical item in an array of pictures. Studies on visualization have produced conflicting results, perhaps because they have examined different types of visualization.

Recent discoveries from brain imaging may have introduced some order into these confusing findings. Imaging studies show two distinct brain regions for visual image processing that correspond to two different visual processing styles. Object imagers construct colorful, pictorial, high-resolution images that provide more static detail. Spatial imagers schematically represent spatial relations among objects and strip away detail to perform spatial transformations. Object imagers excel at visual arts while spatial imagers excel at mechanical tasks. On average, females have an advantage in object imaging tasks and males at spatial imaging.

---


66 J. W. Berry, Temne and Eskimo Perceptual Skills, 1 INTERNATIONAL JOURNAL OF PSYCHOLOGY 207-229 (1966); Russell S. Macarthur, Sex Differences in Field-Dependence for the Eskimo. Replication of Berry’s Findings, 2 INTERNATIONAL JOURNAL OF PSYCHOLOGY 139-140 (1967).


71 Sex and Cognition, supra note __ at 34-37.

72 Larry V. Hedges & Amy Nowell, Sex Differences in Mental Test Scores, Variability, and Numbers of High-Scoring Individuals, 269 SCIENCE 41-45 (1995).

73 Magnitude of sex differences in spatial abilities, supra note __.


75 Id.
5. **MATHEMATICS**

In general, women score higher than men on tests of math computation, while men score higher on tests of math problem solving, including word problems. The difference appears to increase with the difficulty of the questions and the selectivity of the sample population. No sex difference appears in understanding of mathematical concepts and, possibly, on tests of geometric proof writing. The variance in male scores is greater, with males over-represented at both tails of the distribution.

Cross-cultural evidence indicates both heritable and environmental forces at work. The male advantage in mathematical problem solving and the female advantage in computation has been found in a number of different societies including the U.S., Thailand, Taiwan, and Japan. However, the magnitude of these differences varies, and may be correlated with cultural attitudes that support or undermine women’s participation in mathematics. The male advantage in math problem solving does not appear until puberty, and this timing is consistent either with a socialization hypothesis or with a genetic effect triggered by pubertal hormone changes. A role for environmental factors is suggested by the phenomenon of stereotype threat. When a member of a negatively stereotyped group is reminded of the stereotype before being tested, the subject’s performance frequently drops. Studies have found stereotype threat impairs the performance of women taking math tests.

Studies examining time trends in math test score gaps are

---

**Notes:**


78 Gender differences in mathematics performance, *supra* note___.

79 *Id.*

80 Sharon Senk & Zalman Usiskin, *Geometry Proof Writing: A New View of Sex Differences in Mathematics Ability*, 91 AMERICAN JOURNAL OF EDUCATION 187-201 (1983). Surprisingly, no efforts seem to have been made to replicate this study.

81 Camilla P. Benbow, *Sex differences in mathematical reasoning ability in intellectually talented preadolescents: their nature, effects, and possible causes*, 11 BEHAVIORAL AND BRAIN SCIENCES 169-232 (1988); Sex Differences in Mental Test Scores, Variability, and Numbers of High-Scoring Individuals, *supra* note___.

82 George Engelhard Jr., *Gender differences in performance on mathematics items: evidence from the United States and Thailand*, 15 CONTEMPORARY EDUCATIONAL PSYCHOLOGY 13-26 (1990); International gender× item difficulty interactions in mathematics and science achievement tests, *supra* note___.


complicated by many factors and yield mixed results, with the most careful study to date finding a very small trend to convergence.  

6. **Nonverbal Communication and Social Intelligence**

Women consistently outscore men on tests that require the subject to interpret nonverbal signs of what another person is thinking or feeling.  

Evidence for environmental influences comes from studies finding that stereotype threat impairs the performance of men taking tests on sensitivity.  

The age pattern of differences in interpersonal abilities tends to support a role for genetics. Newborn girls show more interest in animate objects while newborn boys show more interest in mechanical objects.  

The female advantage at interpreting other people’s emotions appears to emerge in infancy.

7. **Verbal Ability**

Men and women score differently on a variety of individual verbal tasks. These differences do not seem susceptible to simple categorization. Women score significantly higher on tests that require writing samples and slightly higher on tests of reading comprehension. Women score higher than men on tests of grammar and spelling.  

Men score higher on tests of

---


91. A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents., supra note___.

92. Sex Differences in Mental Test Scores, Variability, and Numbers of High-Scoring Individuals, supra note___.

93. Id.

94. Id.: Janet Hyde & M. C. Linn, *Gender differences in verbal ability*, 104 PSYCHOLOGICAL BULLETIN 53-69 (1988); Cognitive sex differences are disappearing, supra note___.

17
analyses. Men and women show highly task specific differences on verbal fluency, which is defined as the ability to produce words that meet specified conditions. Women score higher than men on tests of phonemic fluency, or the ability to produce words meeting phonemic conditions, such as “starts with C”. On tests of ideational fluency - the ability to produce words that meet certain conceptual criteria - the relative performance of males and females depends on the criteria requested. For example, women score higher for color categories, while men score higher for geometric categories. On tasks requiring the subject to translate visual images into verbal labels, women show a strong advantage where the image is a color, and little to no advantage for other images. Women outscore men on tasks that require memory of verbal material, or of material that can be given verbal labels. No sex difference appears on tasks requiring memory of nonsense material.

8. **Non-Visual Perception**

In addition to differences in visual abilities discussed earlier, men and women appear to differ somewhat in their average levels of other perceptual abilities. Women are on average better able than men to discriminate between tastes. Women may also have a more discriminating sense of touch.

9. **Creativity and Inductive Reasoning**

Certain types of intelligence are very difficult to measure, but recent studies show promise in achieving such measurement. One study showed a female advantage in inductive reasoning. Studies of creativity usually show no sex differences in the distribution of creative ability. When differences are found, they tend to favor females. There is very tentative evidence suggesting that creativity might be positively correlated with other kinds of intelligence.

---

95 Gender differences in verbal ability, *supra* note___.
96 **Sex and Cognition, supra** note___.
in females, and negatively correlated in males. If this turns out to be true it would seem to imply a higher number of women at the highest levels of creative ability where creativity and intelligence are both required.

10. **Speed and Strength**

In developed societies, strength and speed play a diminishing role in the workplace. Nonetheless, the skilled trades and manual labor continues to rely in part on strength. Speed remains in important in only a few areas, notably law enforcement, firefighting, and the military. However, these areas are both politically important and the subject of frequent litigation.

Untrained women are estimated to have approximately half of the upper and two thirds of the lower body strength of men. These differences appear to be primarily due to differences in muscle mass, with a secondary role played by metabolic differences, especially in anaerobic activities.

Some performance differences might result from social forces that cause the average male and female to engage in different levels of physical activity. As late as the 1980’s, exercise books regularly reassured women that working out would not cause them to develop unsightly muscles. To control for the effects of socialization, some research has used a relatively random sample of male and female subjects and examined the effect of additional training on each group. If socialization has hampered female development, additional training should help them more than it helps males, who should initially be closer to their own peak levels. In studies of anaerobic resistance training, men showed a greater absolute increase in strength, suggesting that the male advantage would persist even in an equal environment, but the relative increase is often greater for women implying that environmental factors still reduce relative levels of female strength.

Another technique to control for socialization is to compare the performance of male and female elite athletes, who train at maximum levels from a relatively early age. These comparisons, though, have their own limitations. In most sports, organized competition began at a later date for women than men and the infrastructure supporting women’s athletics may still be

---

107 P. J. Maud & B. B. Shultz, *Gender comparisons in anaerobic power and anaerobic capacity tests*, 20 *British Journal of Sports Medicine* 51 (1986); Claire L. Weber, Michael Chia & Omri Inbar, *Gender Differences in Anaerobic Power of the Arms and Legs: A Scaling Issue*, 38 *Medicine & Science in Sports & Exercise* 129 (2006); Jorge Perez-Gomez et al., *Role of muscle mass on sprint performance: gender differences?*, 102 *European Journal of Applied Physiology* 685-694 (2008). The contraction of muscle cells during exertion is powered the breakdown of adenosine triphosphate (ATP), which in turn releases energy. The body converts nutrients to ATP either anaerobically, without oxygen, or aerobically, with oxygen. For the about first five minutes of exertion, all metabolism is anaerobic; for about another five minutes, both types occur; and after ten minutes only aerobic metabolism occurs.
108 JANE FONDA, JANE FONDA’S NEW WORKOUT AND WEIGHT LOSS PROGRAM 91 (1986).
less developed than that supporting men’s. The pool from which men’s competition draws is larger than the corresponding pool for women. In order to address these issues, studies must not only look at current gaps, but must make a sophisticated analysis of time trends, to determine the likelihood that any sex gap will disappear.

The most studied area of physical performance is running, a sport with a long history of excluding females. Not until 1928 were women allowed in Olympic track and field events; no women’s 1500 meter competition was held until 1972, and no women’s marathon until 1984. As barriers fell in the 1970’s and 1980’s, female track records improved much faster than men’s. By the late 1990’s however, the female rate of improvement had slowed to roughly the male rate, and the difference between male and female records was about 10% in a wide variety of events. A similar gap between male and female swimming records stabilized at around 10% as early as the 1950’s, presumably because women have competed at high levels for a longer time.

One study even found that the sex difference began to increase again in the mid-1990’s. The authors speculate that improved drug testing had a greater effect on women than men, since women benefit more from dosing with testosterone and related substances. A number of investigators have begun to conclude that remaining sex differences represent heritable factors. Significant sex differences have been found in the basic biomechanics of movement, which, unlike endurance, are relatively invariant to training.

Although strength is relevant to more occupations than speed, sex differences in strength are hard to study through examination of elite performance. Virtually all athletic activities in some way make use of strength, but very few clearly isolate strength from other abilities. Weight-lifting is the principal purely strength-based sport, but differences between the rules of men’s and women’s events make comparisons difficult. One of the few studies to date finds a greater sex difference in weight-lifting than speed, about 35% as compared with 10%. This result is consistent with the general pattern of greater male advantage in anaerobic activities than in aerobic ones.

---

115 Id..
117 Elizabeth S. Chumanov, Cara Wall-Scheffler & Brian Heiderscheit, Gender differences in walking and running on level and inclined surfaces, CLINICAL BIOMECHANICS (2008).
118 Causes of sexual dimorphism in performance traits, supra note___
11. **EXPLAINING THE FACTS**

The relative performance of men and women in some tasks does correspond to certain features of the hunter-gatherer scenario that lies behind both evolutionary psychology and functional environmental theories. The male advantage in targeting ability might plausibly be explained by specialization in hunting. Spatial rotation is closely related to the ability to reason about moving objects, in turn a component of targeting skill. Speed and strength likewise seem more important in hunting than gathering and child-care, though carrying a thirty-pound child surely requires considerable strength.

The female advantage in fine motor skills might reasonably be connected to female specialization in gathering. Possibly the hunter-gatherer hypothesis explains sex differences in navigational strategies. The male use of absolute locational strategies might be more appropriate for a return home from a long and irregular journey. Reproducing the outbound route by landmarks would serve no purpose; rather the objective would be to find the shortest and safest route back. The female use of landmark strategies might be more appropriate for repeated navigation in a defined area where it is important to remember locations that are especially rich in foraging opportunities.

The hunter-gatherer hypothesis has been proposed to explain the female advantage in social abilities, but this connection seems less plausible. Women, it has been argued, need greater interpersonal abilities because of their child rearing responsibility. But the ability to read nonverbal cues of what other people are thinking would seem to be highly advantageous in the social interactions of males as well. Male reproductive success depends in part on the ability to perceive how receptive a female is to sexual overtures. Status-seeking among males would be helped by the ability to read the intentions of those other males. Conversely, it is unclear that a highly developed ability to read another person’s state of mind is important for parents. Being a parent is an extraordinarily difficult task that requires patience, time, and general intelligence. Reading the mental state of children, however, is usually remarkably easy. Babies have a very small set of simple needs, and once children talk they are all too willing to express, relentlessly, their needs or at least their demands.

A further limitation of functional explanations is that cognitive sex differences do not display simple patterns, such as verbal versus mathematical, or emotional versus logical. In every category of ability, each sex excels at some tasks, and neither excels at some. Men have an advantage at mathematical problem-solving; women at mathematical computation; and neither sex at mathematical concepts. Why should either nature or nurture make women better than men at math computation and men better than women at tests of math problem solving? Why is the pattern of verbal abilities so seemingly random? Why do females have an advantage at so many perceptual tasks? Strictly functional explanations seem implausible.

Sex roles might also be explained by theories of patriarchy in which the subordination of women is intentionally sought by males to control female sexuality. I believe the history of female subordination explains some sex differences in the distribution of personality traits, and is thus an important component of occupational segregation. However, I at least cannot see any connection between the objective of subordination and the strange observed pattern of ability

---

120 See Section II. A.
121 *Bedrooms and Boardrooms*, supra note___.
distributions. A group intent on domination should try to suppress general intelligence in the oppressed group, yet general intelligence shows no significant sex differences. Indeed, some of the abilities at which females excel, such as mind-reading and verbal communication, seems highly subversive to any system of subordination. Oppressors, I would expect, like their victims mute and clueless.

Perhaps a desire to dominate motivates males to essentialize sex roles by reinforcing already-existing differences. Suppose ability X does not inherently undermine patriarchy, but has a heritable male advantage. Rigidly-policed sex roles might well serve the goal of subordination, since any exception suggests that other exceptions might exist. Females who exhibit ability X undermine the essentialization of sex roles, and jeopardize patriarchy even if nothing about ability X is especially threatening. Patriarchy thus provides a motive to exaggerate sex roles, but if ability X does not itself support the system of subordination, there must still be some pre-existing heritable sex distribution difference to trigger this mechanism.

If functional explanation is incomplete, can its gaps be filled by any non-functional alternative? Perhaps some non-functional sex differences are a by-product of other functional forces, a possibility that will be addressed in the next section.

C. PHYSIOLOGY

1. STRUCTURAL DIFFERENCES

Male and female brains display non-trivial sexual dimorphism. Significant differences have consistently been found in the brain chemistry of men and women.\textsuperscript{122} Even adjusting for body size the female brain is smaller than the male,\textsuperscript{123} though it has more folds and fissures, creating more surface area that may compensate for smaller size.\textsuperscript{124} Males and females have different percentages of various types of brain tissue,\textsuperscript{125} and the lateralization of some tissue types is asymmetric in males but not females.\textsuperscript{126} However, researchers have sometimes been too quick to infer sex differences that are later explained by other factors. The widely-reported observation that women have a larger corpus callosum connecting the two hemispheres turns out to be artifact of smaller brain size: the corpus callosum grows less than proportionally with brain size in both males and females.\textsuperscript{127}

\textsuperscript{122} Recent surveys of this literature include Kelly P. Cosgrove, Carolyn M. Mazure & Julie K. Staley, Evolving Knowledge of Sex Differences in Brain Structure, Function, and Chemistry, 62 BIOLOGICAL PSYCHIATRY 847-855 (2007); Larry Cahill, Why sex matters for neuroscience, 7 NATURE REVIEWS- NEUROSCIENCE 477-484 (2006).
\textsuperscript{124} Eileen Lüders et al., Gender differences in cortical complexity, 7 NATURE NEUROSCIENCE 799-800 (2004).
\textsuperscript{125} Females have higher percentages of gray matter (nerve cell bodies, dendrites and small axons); males have higher percentages of white matter (larger axons and their myelin sheaths) and cerebrospinal fluid. Evolving Knowledge of Sex Differences in Brain Structure, Function, and Chemistry, supra note \textsuperscript{124}.
\textsuperscript{126} Ruben C. Gur et al., Sex Differences in Brain Gray and White Matter in Healthy Young Adults: Correlations with Cognitive Performance, 19 JOURNAL OF NEUROSCIENCE 4065-4072 (1999); Richard J. Haier et al., The neuroanatomy of general intelligence: sex matters, 25 NEUROIMAGE 320-327 (2005).
\textsuperscript{127} Christiana M. Leonard et al., Size Matters: Cerebral Volume Influences Sex Differences in Neuroanatomy, CEREBRAL CORTEX (2008).
If sexually dimorphic brain structures are linked to cognitive differences, the origin of these anatomical differences becomes critical to the debate over the heritable or environmental origin of cognitive differences. For many years, all genetic sexual dimorphism was thought to be caused indirectly through hormonal mechanisms. Recent research suggests that other genes also contribute to sexual dimorphism, though the exact nature of these mechanisms is not well understood. For the moment our knowledge about heritable differences in brain structure is limited to differences caused by hormonal mechanisms.

All embryos are initially female. Early in pregnancy, a single gene on the Y chromosome of genetic males triggers a series of events that masculinize the fetus. The testes are formed, and begin to secrete testosterone that sex-differentiate a number of cells, including some in the brain. These hormonal effects are organizational: they affect the development of structures that persist regardless of subsequent hormone levels. At later stages in development, especially puberty, new waves of internally produced (endogenous) hormones induce further organizational changes.

The organizational effects of hormones are probably more heritable than environmental, though not to the exclusion of all environmental factors. One environmental factor of particular importance is feedback from success or failure, which tends to depress testosterone. If girls are given less positive reinforcement than boys, this environmental factor could clearly influence hormone levels by the time of puberty, and might thus affect the organizational changes that occur at puberty. Organizational effects during fetal development are harder to explain by environmental factors. Although maternal hormone levels clearly influence fetal development, most of the factors affecting those levels, like stress to the mother, should be independent of the sex of the fetus. A connection between maternal hormone levels and fetal sex would seem to require that the mother’s body be aware of the sex of the fetus. This suggestion is not quite as preposterous as it sounds, but if mothers have such awareness, it cannot be very strong.

2. **Physiological Factors and Cognition**

The connection between brain anatomy and cognition is now studied by using imaging techniques such as fMRI to examine brain activation patterns while the subject conducts specific cognitive tasks. These studies indicate that sex differences in brain anatomy are not sufficient to demonstrate a heritable role in cognitive differences. Some anatomical sex differences apparently prevent behavioral differences by compensating for other anatomical differences.

---


130 Relative miscarriage rates of male fetuses rises with maternal stress, see, e.g., Ralph A. Catalano, *Sex ratios in the two Germanies: a test of the economic stress hypothesis*, 18 Human Reproduction 1972-1975 (2003), and some have speculated that the body of a stressed mother “prefers” female fetuses, Robert L. Trivers & Dan E. Willard, *Natural selection of parental ability to vary the sex ratio of offspring*, 179 Science 90 (1973). But see Sex ratios in the two Germanies, supra note ____. Any maternal ability to divine fetal sex cannot be very powerful, since an effective maternal ability to sex-select through miscarriage would obviate the need for sex-selective infanticide, a tragically common phenomenon. MARTHA C. NUSBAUM, SEX & SOCIAL JUSTICE (1999).

131 Evolving Knowledge of Sex Differences in Brain Structure, Function, and Chemistry, supra note ___.

132 A number of studies find males and females perform equally well on a given task, but show different brain activation patterns. Geert J. De Vries, *Minireview: Sex Differences in Adult and Developing Brains*:
This compensatory function may be seen in the anatomy of general intelligence. Men and women apparently achieve similar levels of general intelligence with different brain designs, including different brain sizes and different distributions of various types of brain tissue. Conversely, some sex differences in performance are not, at least yet, associated with sexually dimorphic brain structures. However, a number of studies indicate that brain dimorphisms, including differences in lateralization, may produce some of the cognitive differences between men and women. In both sexes, the left hemisphere specializes in language processing and the right in spatial and mathematical activities. This lateralization is more pronounced in males than in females, who tend to use both hemispheres for different tasks. One intriguing aspect of lateralization involves creative thinking. Men and women display very different patterns of brain activation during creative thinking. The female pattern of low lateralization has been associated with higher levels of creative thinking, and there is weak evidence that women may have a slight advantage in creativity. Interestingly, creativity is associated with higher levels of intelligence in men, and lower levels in women.

Before the development of imaging techniques, the heritable basis of cognitive difference was studied indirectly by examining whether hormonal variation at critical periods like gestation had organizational effects on later cognitive ability. These methods remain useful today. Studies on animals provide strong support for the hypothesis that hormones have organizational effects on cognitive behavior. Research on humans is less conclusive, but points to at least a few links. Prenatal testosterone affects various non-cognitive sex-linked behaviors such as play behavior and sexual orientation. Relatively strong evidence suggests that prenatal testosterone...
exposure affects targeting ability. Some research suggests that prenatal exposure to testosterone reduces interpersonal abilities and verbal ability. Studies on mental rotation yield mixed results.

3. **Activation Effects of Hormones**

The level of hormones circulating in the bloodstream at any given time does not have permanent structural effects but might have activational effects; that is, circulating hormones might influence behavior during the time they are present.

Many studies have examined whether circulating hormones affect cognitive functioning in tasks on which one sex shows an advantage. The most commonly examined link is between testosterone and spatial tasks. Animal research provides strong support for the role of circulating androgens, but the results of studies in humans are inconclusive. To the extent that studies

---


143 Studies finding no effect include J. Coolican & Michael Peters, Sexual dimorphism in the 2D/4D ratio and its relation to mental rotation performance, 24 EVOLUTION AND HUMAN BEHAVIOR 179-183 (2003); A preliminary investigation of the associations between personality, cognitive ability and digit ratio, supra note___. Studies finding an effect include Testosterone, supra note__ (also finding effect on figure disembedding task); John T. Manning & Rogan P. Taylor, Second to fourth digit ratio and male ability in sport: implications for sexual selection in humans, 22 EVOLUTION AND HUMAN BEHAVIOR 61-69 (2001); P. Kempel et al., Second-to-fourth digit length, testosterone and spatial ability, 33 INTELLIGENCE 215-230 (2005); Gina M. Grimshaw, Gabriel Sitarenos & Jo-Anne. K. Finegan, Mental Rotation at 7 Years-Relations with Prenatal Testosterone Levels and Spatial Play Experiences, 29 BRAIN AND COGNITION 85-100 (1995). Some studies appear consistent with the inverse U hypothesis. Id. (spatial ability raised in females, lowered males); Spatial abilities following prenatal androgen abnormality, supra note__ (no effect in females, negative correlation in males).


145 The interpretation of these studies is complicated many factors, including the fact that handedness may affect spatial ability; different measures of spatial ability show different degrees of sexual differentiation; the results vary with the chosen measure of testosterone; and some across-subject designs do not control for factors such as subject’s general intelligence. Compare Rozmin Halar et al., Sex differences and individual differences in cognitive performance and their relationship to endogenous gonadal hormones and gonadotropins., 119 BEHAVIORAL NEUROSCIENCE 104-17 (2005)(no between-subject effect of hormone levels. Normal subjects, no intelligence controls, no direct measure of free T); Lynn S. Liben et al., The Effects of Sex Steroids on Spatial Performance: A Review and an Experimental Clinical Investigation, 38 DEVELOPMENTAL PSYCHOLOGY 236-53 (2002)(steroids for delayed puberty; between-subjects analysis using within-subject benchmarks; no effect); Testosterone, supra note__ (no effect; between subjects, no control of general intelligence) with Andre Aleman et al., A single administration of testosterone improves visuospatial ability in young women, 29 PSYCHONEUROENDOCRINOLOGY 612-617 (2004); Albert Postma et al., Sex differences and menstrual cycle effects in human spatial memory, 24 PSYCHONEUROENDOCRINOLOGY 175-192 (1999)(exogenous testosterone correlated to spatial performance in normal females); Markus Hausmann et al., Sex Hormones Affect Spatial Abilities During the Menstrual Cycle, 114 BEHAVIORAL NEUROSCIENCE 1245-1250 (2000); Scott D. Moffat et al., Longitudinal
find a relationship, they tend to suggest that the optimal level of testosterone is in the low-normal male range, so that increasing testosterone in women should always increase spatial functioning, while in men it will have the reverse effect above a threshold level.\textsuperscript{146}

An activational effect of hormones is consistent with both heritable and environmental causes of dimorphism. On the one hand, circulating hormone levels are determined primarily by each individual’s heritable makeup. On the other hand, these levels can also be influenced by environmental factors.\textsuperscript{147} If girls receive less positive reinforcement than boys, the resulting drop in testosterone might contribute to differential performance on testosterone-linked tasks. Given current data, the activational effects of hormones on cognitive patterns might be determined entirely by the heritable component of hormones levels or entirely by the environmental component. Only more research can resolve this. At present, the simplest explanation for activational effects is that both environmental and heritable factors play a role in proportion to their importance in determining overall hormone levels, which would give a greater though not exclusive role to heritable factors.

4. **Strength and Speed**

Strength training appears to improve muscle associated with aerobic exercise in both young men and young women, but muscle associated with anaerobic exercise increases only in young men.\textsuperscript{148} Even when women gain relatively more strength than men, they may gain less relative muscle size.\textsuperscript{149} The sex difference in muscle mass is presumed to be mediated by higher testosterone levels in men, because of the well known effect of testosterone.\textsuperscript{150} Several studies have reported sex-related differences in the expression of genes that could account for the sex difference in muscle size.\textsuperscript{151}


\textsuperscript{147} Effects of ability-and chance-determined competition outcome on testosterone, \textit{supra} note\textsuperscript{___}.

\textsuperscript{148} Gregory F. Martel et al., \textit{Age and sex affect human muscle fibre adaptations to heavy-resistance strength training}, 91 \textit{Experimental Physiology} 457-464 (2006).

\textsuperscript{149} Effects of Strength Training and Detraining on Muscle Quality Age and Gender Comparisons, \textit{supra} note\textsuperscript{___}.


\textsuperscript{151} Priscilla M. Clarkson et al., \textit{ACTN3 genotype is associated with increases in muscle strength in response to resistance training in women}, 99 \textit{Journal of Applied Physiology} 154-163 (2005).
5. **THE PATTERN OF ABILITIES REVISITED: IS IT A BY-PRODUCT?**

The physiological evidence suggests some explanations for the odd pattern of sex-differentiated abilities. Many physiological sex differences can be traced to a few common mechanisms, including the testosterone surges prenatally and at puberty. Perhaps only few sex differentiated abilities serve any adaptive purpose, while the others are by-products of mechanisms that produce adaptive traits. This hypothesis is supported by evidence that there are several dimensions of intelligence. On each dimension, one end represents one ability, such as spatial rotation, while the other represents a wholly different ability, such as some aspect of verbal ability. Several of these dimensions exhibit a trade-off: individuals high in ability at one end of the dimension tend to be low in ability at the other end. Individuals high in spatial rotation abilities tend to be low in the cluster of verbal abilities in which females have an advantage. Thus, the adaptive forces that produce high male spatial ability seem, as a by-product, to reduce certain verbal abilities.

D. **OCCUPATIONAL SEGREGATION IN SIMPLE SOCIETIES**

Most theories of ability differences depend on empirical claims about early human society. The classic study of occupational segregation in preindustrial societies is George Murdock and Caterina Provost’s analysis of data from the Standard Cross Cultural Sample. Murdoch and Provost calculated the extent to which various activities were predominantly performed by members of one sex or the other. Males monopolize the hunting of large fauna and most wood and metal working related activities. Occupational segregation can also be seen in activities not included in the Murdock-Provost analysis. In all preindustrial societies, men play little or no role in the care of young children. In all societies, industrial and pre-industrial, young men engage in both individual and collective homicide against adults at vastly higher rates than women, although infanticide by women is not uncommon.

The Murdock-Provost data show other activities that are dominated though not monopolized by one sex. Women dominate gathering and cooking, while men dominate animal husbandry and fishing. Subsequent research shows female labor is relatively important in extensive agriculture that uses low levels of capital, materials, and labor relative to land area. Male labor is relatively important in intensive agriculture.

---

153. *Id.*
These strong patterns of occupational segregation seem unlikely to have resulted from numerous chance events in so many separate societies. The common ancestry hypothesis proposes that only one random allocation of sex roles occurred in one common ancestral society. However, the common society would have had to have existed well before fifty thousand years ago when the first humans left Africa and human cultures took clearly different courses. Human social organization has diverged in so many ways that the survival of a purely random variation on such an important matter requires further explanation.\textsuperscript{159} The coordination hypothesis provides reason to expect that the initial sexual division of labor might be more persistent than other cultural attributes and may explain the resilience of sex roles that existed fifty thousand years ago, such as female specialization in child care. However, at the time of separation and for the next thirty thousand years, all human societies relied on foraging for subsistence. Coordination and the common ancestor theory cannot explain why, as separated societies developed new technologies, they tended to segregate the sexes in similar ways. For example, extensive agriculture is typically female, while intensive agriculture is typically male. This pattern cannot be explained by a purely environmental functional theory, which would suggest a large female role in all agriculture, since it is close to home and not dangerous. Indeed, intensive agriculture cultivates a smaller area, and is in that respect more compatible with child-care. The dominance of males appears to result from their greater upper body strength, of particular importance in plow agriculture.

Functional explanations might account for some cross-cultural patterns. For example, females have a virtual monopoly on early child rearing, and their dominance in gathering could result from its complementarity to child rearing. These functional divisions appear to have both environmental and heritable components. The existence of swing activities is important evidence of some flexibility in sex roles.\textsuperscript{160} The sexual division of labor responds, though imperfectly, to different environmental circumstances. In societies in which female-dominated activities are relatively productive, males sometimes respond by switching some of their activity to characteristically female types of production. However, they also often reallocate their time from resource production to high levels of male-male violence, often for mate competition. Although this response provides some support for evolutionary psychology, it contradicts a version of genetic functionalism which posits that males dominate subsistence activities because female time is absorbed by child raising. Female contribution to subsistence is especially high in Africa and the Pacific Islands, where gathering and extensive agriculture are productive relative to hunting and intensive agriculture. In such societies, women remain the principal child care providers, but often make use of collective child care arrangements that allow female subsistence activities. The responsiveness of labor allocation supports a role for the environment, while the limits of this responsiveness support heritability.

\textsuperscript{159} For a similar argument, see Evolved sex differences and occupational segregation, supra note\textsuperscript{___}.  
\textsuperscript{160} Note that swing activities that may be assigned to either sex across cultures are often sex-segregated with a given culture. Factors in the division of labor by sex, supra note\textsuperscript{___}.  

28
III. EVIDENCE OF STATISTICAL DISCRIMINATION

The theory of statistical discrimination suggests that employers will use sex as a proxy for ability in hiring for positions in which sex-linked abilities play a role. Empirical support for this requires evidence that discrimination exists and that it can be characterized as statistical in nature.

Two broad classes of evidence are helpful in understanding these questions. Evidence from real labor markets is illuminating primarily on the question of whether discrimination still exists in practice, while experimental laboratory evidence is helpful in identifying both the existence of and the nature of discrimination.

A. DISCRIMINATION IN GENERAL

1. STUDIES OF LABOR MARKET DATA

Economists traditionally measured discrimination by estimating wage regressions on observable proxies for productivity such as education. Discrimination is then measured in various ways, such as the unexplained residual or the coefficients on dummy variables for race or sex. Such studies tend to support the view that discrimination, though declining, continues to contribute to the wage gap.

Occupational segregation complicates the analysis of the wage gap. Several studies have confirmed that occupational and industry segregation plays a significant role in women’s lower pay, with the balance apparently attributable to discrimination. If occupational segregation itself is in part a result of discrimination, wage regressions underestimate the effects of discrimination.

Conversely, wage regressions may overstate discrimination by omitting hard to measure productivity variables. Economists have used various techniques to address this concern. One study approached the productivity question indirectly by examining the effects when an employer switched to sex-blind hiring procedures. It found that such procedures indeed increased the rate at which females were hired, suggesting that discrimination had influenced the non-sex blind decisions. Another major study was able to match employee and employer data and thus estimate worker productivity. The results provided evidence of wage discrimination.

---

161 Ronald Oaxaca, Male-Female Wage Differentials in Urban Labor Markets, 14 INTERNATIONAL ECONOMIC REVIEW 693-709 (1973) and Alan S. Blinder, Wage Discrimination: Reduced Form and Structural Variables, 8 JOURNAL OF HUMAN RESOURCES 436-65 (1973).
164 New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data, supra note__.
166 New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data, supra note__.
Other studies begin with datasets that have especially detailed information that can be used as controls for productivity. One study examined the labor market outcomes for graduates of the University of Michigan Law School.\textsuperscript{167} Right after graduation, the gap in pay between women and men was relatively small but after fifteen years, the average earnings of women graduates were sixty percent of those of men. Even controlling for current hours worked, family status, race, location, law school grades, years of law practice, months of part-time work, and type and size of employer, a male earnings advantage of 13 percent remained. Another study examined wage differences among recent college graduates. Even controlling for college major, college grade point average, and specific educational institution attended it found an unexplained pay gap of 10 to 15 percent between men and women.\textsuperscript{168}

Another study addressed concerns about unmeasured productivity variables by the so-called audit method, in which matched pairs of individuals are sent to apply for advertised job openings. These individuals were of different sexes but had similar education and experience, and researchers attempted to match less tangible qualities as well. The study found that females were less likely to be hired than comparable males.\textsuperscript{169} Field studies find that each sex receives better performance ratings in performance measures that are sex-role congruent, but these studies cannot fully control for possible differences in actual performance.\textsuperscript{170}

Ascertaining the effect of household production is perhaps the most difficult aspect of assessing the causes of the gender gap. The literature on this cannot briefly be summarized, but in a careful and comprehensive survey, Joni Hersch concludes that the evidence as whole indicates that turnover and family duties, in conjunction with other observables, cannot account for all of the gender gap, leaving a significant unexplained proportion that may be caused by discrimination.\textsuperscript{171}

2. \textit{Experimental Evidence}

Other studies, primarily by psychologists, have addressed the problem of limited productivity data by constructing experimental laboratory settings in which subjects are asked to make employment-related decisions about fictitious job candidates whose productivity characteristics can be held constant while their sex is manipulated by the experimenter. Sometimes the subject meets actual “job candidates” who are in fact actors, and sometimes the subject receives only information about the candidate, such as a resume. Many different employment-related decisions have been simulated, the most common being a hiring decision or a performance evaluation.

Two early meta-analyses of sex discrimination studies suggested that sex discrimination accounted for a share, though probably a small one, of the wage difference between men and women.\textsuperscript{172} The most recent meta-analysis by Davison and Burke found a much larger effect\textsuperscript{173}

\textsuperscript{168} Id..
\textsuperscript{171} JONI HERSCH, \textit{SEX DISCRIMINATION IN THE LABOR MARKET} (2007).
\textsuperscript{172} Judy D. Olian, Donald P. Schwab & Yitchak Haberfeld, \textit{The impact of applicant gender compared to qualifications on hiring recommendations: A meta-analysis of experimental studies}, 41 ORGANIZATIONAL
that appears to be attributable to several factors. Davison and Burke’s sample contained over twice as many studies as the earlier samples. This allowed them to examine the possibility that discrimination was not solely directed against women but rather against whichever sex was not traditionally associated with the job in question. Their study strongly supports this sex-role congruence theory of discrimination, and is consistent with observation that in real labor markets, men in female-dominated professions earn less than women in those professions.

B. TYPE OF DISCRIMINATION

Sex discrimination may take several forms, all of which will tend to produce occupational segregation. The strongest reason to believe that discrimination is statistical would come from evidence that employers update their decisions based on new information. Irrational generalizations are, almost by definition, resistant to updating: if an irrational generalization were updated, it would eventually become a rational estimate. Tastes likewise cannot change too readily. They may evolve over long periods of time, but the entire construct of neoclassical economics assumes that they are relatively stable in the short run.

In contrast, statistical generalizations are rational inferences from the best cost-effective available information, and should incorporate any new information. A rational employer should be willing to override a generalization given sufficiently detailed information about an individual candidate. An employer with a false but rational generalization should be willing to update that generalization based on new information, such as an increase in the proportion of qualified women in the applicant pool.

Labor market data tends to support the hypothesis of sex discrimination but sheds little light on whether its source is statistical. Laboratory evidence, in contrast, permits examination of the source of discriminatory behavior. Meta-analysis of this evidence suggests that a subject’s

---


The first meta-analyses did not address the sex role congruence issue, The effects of the type and amount of information in sex discrimination research, supra note___, while the second was unable to reach a conclusion because of the sample size. The impact of applicant gender compared to qualifications on hiring recommendations, supra note___. Tosi and Einbinder also made the mistake of dichotomizing all studies as either significant or not, rather considering their relative significance levels, a mistake corrected in both Olian, Schwab, and Haberfeld and Davison and Burke.

reliance on the applicant’s sex decreases when the subject is given more information about applicant productivity, suggesting that discrimination is at least in part rational and updatable.

Other studies do indicate the some discriminatory outcomes are motivated by irrational generalizations. A subject may rationally use statistical generalizations when asked to predict the future performance of those workers. However, a rational subject should not allow statistical generalizations to color the process of evaluating past performance when provided with all relevant work product. In fact, subjects given two otherwise identical work samples with male and female names attached will rate them differently.

IV. SOME METHODOLOGICAL ISSUES

The role of empiricism in legal policy-making is a subject of perennial contention. A thorough review of this debate is well beyond the scope of this paper, but a few methodological notes are in order.

A. EMPIRICISM GENERALLY

Some observers are skeptical of basing policy on empirical studies some of which, almost inevitably, will not withstand the test of time. The scientific consensus of one era may be debunked and discarded in the next. However, factual assumptions are all but unavoidable in employment discrimination law. If those assumptions are based on scientific evidence, subsequent researchers can challenge that evidence with new studies, and a line of inquiry will move slowly towards accurate description of the facts. Policy can co-evolve with a gradually improving understanding of the facts.

The alternative to the use of research is to base empirical assumptions on common sense. Common sense, however, is not always held in common by different people. Some people believe that discrimination self-evidently continues to be a problem, while others believe it self-

176 Sex Discrimination in Simulated Employment Contexts, supra note___. At the extreme, less information can actually reduce discrimination. With only information about group membership, subjects often protest that they cannot reach a decision based on stereotype. Subjects appear to need just enough performance information to convince themselves that they are not basing their evaluation on generalizations alone. John M. Darley & Paget H. Gross, A hypothesis-confirmation bias in labeling effects, 44 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 20-33 (1983).


evidently does not. Some believe that the heritable origin of many sex differences is obvious while others believe with equal firmness that the environmental origin of these differences is obvious. Common sense offers no method for resolving these differences. Scientific inquiry does. Indeed, it is not even clear that appeals to common sense actually eliminate the influence of academic research. As John Maynard Keynes famously observed, “Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist.”

B. **EXTERNAL VALIDITY**

Much of the evidence on both difference and discrimination raises the problem known as external validity. Even if a study supports the researcher’s inferences under the specific conditions studied, those results might not generalize to real-life situations. Perhaps the most commonly expressed concern is that experimenter expectations may influence the behavior of the subjects, and studies of both discrimination and difference have been criticized on this basis.

Some have argued that experimenters who wish to find difference in studies of cognitive ability subtly encourage participants to display sex role-appropriate abilities. This argument rests on the assumption that people are encouraged by positive feedback and discouraged by negative feedback, which is consistent with my own casual observation of human behavior. Studies that find difference thus may exaggerate those differences.

Others have argued that participants in discrimination studies know that the researcher wishes to find discrimination, and attempt to cooperate by obtaining this result. In resume studies, this effect would occur only if subjects were so eager to please the experimenter that they would humiliate themselves by engaging in behavior, discrimination, that they know is regarded as morally wrong by the experimenter, and that they themselves probably believe to be wrong. This seems unlikely to me, though not impossible. Indeed, I would not be surprised to find a different problem. Knowing that they were being carefully observed, subjects might be careful not to make discriminatory decisions that the experimenter would regard as wrong. Resume studies would thus underestimate the degree of discrimination found in real world settings.

Experimenter effects in audit studies seem somewhat more plausible. On the one hand, subjects would want to be perceived as competent and would perceive rejection as embarrassing, motivating them to try to receive a job offer. On the other hand, if a female or non-white subject believed that the experimenter would regard failure as evidence of discrimination rather than incompetence, this belief might compensate for embarrassment factor. Another (and to my mind more likely) problem might arise if these subjects sensed the experimenter expected failure, and

---

were discouraged by this expectation, thus lowering performance. In this case, audit studies would support stereotype threat, but not employer discrimination.

Experimenter effects should not be dismissed in any of these studies. In rare instances, a study design may be so patently defective that it clearly conveys the experimenter’s desired outcome to subjects, and it can be dismissed on grounds of an experimenter effect. More commonly, however, experimenter expectations would be conveyed more subtly, probably unconsciously, and could not be inferred from the reported research results. In the normal course of scientific method, experimenter effects can be caught when other researchers with different priors attempt to replicate the study and fail to do so. Until such studies are done, however, the burden of proof is generally on those who assert an experimenter effect to demonstrate it. At present, the available evidence, though not entirely satisfactory, supports the existence of both discrimination and actual difference.

Some commentators have suggested laboratory decisions of personnel decisions lack external validity because they provide subjects with far less information than what is available to real world employers. However, real world hiring procedures heavily screen out applicants based on resumes. That screening is almost perfectly replicated in lab situations. Even non-discriminatory follow-up procedures would be unlikely to undo discrimination in the initial screening, and real world follow-ups generally take the form of unstructured interviews, which may be influenced by discriminatory generalizations.

In promotion and discharge decisions, concerns about external validity are more plausible. After all, the employer has the work history of the employee to consider. An employer that uses a structured method for aggregating this information might indeed succeed in reducing the role of discriminatory generalizations, but such methods are far from universal. Even if the errors resulting from the absence of structural methods are small, these errors cumulatively can produce discrimination of large magnitude. The larger information set available to real employers will tend to require more data aggregation decisions, more cumulation of errors, and thus employer decisions will be even more, not less, discriminatory than laboratory experiments indicate.

Consideration of the full range of methodological concerns is beyond the scope of this paper, but one intriguing approach to resolving these debates has been proposed. In so-called adversarial collaboration, researchers with opposing views jointly design empirical studies. Each side stipulates, before the study is carried out, what results would force it to reconsider its views. Such collaborations would greatly improve the quality of the debate on these issues.

---

184 See Section VII.B.
185 See Section VII.B.
187 Philip E. Tetlock & Gregory Mitchell, Unconscious prejudice and accountability systems: What must organizations do to prevent discrimination?, 30 RESEARCH IN ORGANIZATIONAL BEHAVIOR.
V. CURRENT OCCUPATIONAL SEGREGATION

The most plausible account of sex differences in abilities combines both environmental and genetic functional explanations with physiological constraints that require tradeoffs between different types of abilities. To the extent that genetics plays a role in ability difference, how much can it explain of current patterns of occupational segregation?

A. COMPARATIVE ADVANTAGE

For many observed sex differences, evidence suggests some heritable basis but is far from conclusive about the size of the heritable component. Some observers argue that these magnitudes are relatively small and therefore cannot account for the observed degree of occupational segregation.\(^{188}\)

This reasoning, however, is mistaken. Even small ability distribution differences might produce large degrees of occupational segregation. Suppose that there are two economically valuable tasks, A and B. Suppose that general abilities, such as intelligence, are the main predictors of success at both tasks, and that the distribution of general abilities shows enormous intrasex variation and no intersex differences. However, suppose that there are also small task-specific abilities and that each woman is better by epsilon at task A than she is at task B, while each man is better by epsilon at task B than he is at task A. Each woman has a comparative advantage at A compared with each man, even women who have an absolute advantage at B over most of the male population.\(^{189}\) This scenario will tend to produce significant occupational segregation. The exact degree of segregation will depend on factors such as the relative value of tasks A and B, which in turn depends on the demand for products of the two industries. If the products of both activities are perfect substitutes for each other, a competitive market with complete information would produce complete occupational segregation even though the variation within sexes is huge and the variation between sexes small.\(^{190}\)

Those who believe that differences in ability distributions explain most observed segregation should not take too much comfort from this simple example. Suppose there are three sectors, A, B, and C. All females are more productive than all males in sectors A and B, although the female advantage is higher in A than B. All males are more productive than all females in sector C. At first glance this might seem to imply that all males should be employed in sector C, while all females should be employed in sectors A and B. However, few conclusions can be drawn about the efficient allocation of workers without more information about the demand for the products of each sector. Consider a society near subsistence level in which sector A produces something critical, like food, while sectors B and C produce something of less value. All female workers may end up in A, leaving only male workers in B, despite the absolute female advantage in B. By varying the demand for the products of the three sectors, an almost infinite number of efficient allocations are possible. Some generalizations are possible: each sex will be allocated to

\(^{188}\) MELISSA HINES, BRAIN GENDER (2004).

\(^{189}\) This hypothetical is intended only to illustrate how the size of sex distribution differences can be unrelated to the magnitude of occupational segregation. It seems highly unlikely that there is any task at which all members of one sex have a comparative advantage over all members of the other.

\(^{190}\) An analogous point is made by Gary Becker, who points out that small differences can create extreme occupational segregation in the household division of labor. GARY S. BECKER, A TREATISE ON THE FAMILY (1991).
sectors where the demand for sex-specific abilities are highest. However, in less extreme cases, the fact that one sex has an absolute advantage in a sector does not guarantee that that sex should predominate in that sector.

The principle of comparative advantage thus makes difficult any inference from the absolute advantages of each sex to the efficient degree of occupational segregation. Small differences in absolute advantage may be magnified; conversely, one sex might end up in a sector in which it has an absolute disadvantage. All that can be said is that each sex will tend to gravitate towards the fields which make the most intensive use of abilities in which that sex has a comparative advantage.

B. JOB ANALYSIS

Using sex differences in abilities to explain occupational segregation fully would require extensive study of the correlation between performance on specific cognitive tests and success at various occupations. Studies of test validation, however, have overwhelmingly focused on the predictive value of general intelligence, which is generally found to be high. Relatively little research has examined the association between specific abilities and success in specific occupations, or even between specific abilities and general occupational success.¹⁹¹

The job-relatedness of physical abilities has been successfully validated in the military,¹⁹² firefighting, and, to a lesser extent, law enforcement. Unfortunately, there appears to be no research on the extent to which physical strength is still of importance in areas such as the skilled trades, which remain dominated by males.

The limited research that exists suggests that specific cognitive abilities have some predictive value in specific occupations. One laboratory study has found that both males and females in male-typical occupations perform better on a spatial test than both males and females in female-typical occupations. Conversely, males and females in female-typical occupations perform better on a female advantage verbal-fluency test than males or females in male-typical occupations.¹⁹³ Another study found that high spatial scores in high school predict choice of engineering as a major and an occupation, even controlling for mathematical ability.¹⁹⁴

C. WHAT CAN BE EXPLAINED?

A daunting number of problems complicate any effort to separate the effects of discrimination and difference in occupational segregation. Either statistical discrimination or comparative advantage can magnify small differences. Males might dominate a female advantage field because of discrimination or because females are needed even more elsewhere. The current state of research on sex differences, discrimination, and job requirements is nowhere close to what would be needed to resolve these issues.

¹⁹⁴ Lloyd G. Humphreys & Grace Yao, Prediction of graduate major from cognitive and self-report test scores obtained during the high school years, 90 PSYCHOLOGICAL REPORTS 3 (2002).
All that can be said at this point is that there is some rough correspondence between evidence of difference and at least some patterns of segregation. The implications of sex differences in physical abilities have been studied primarily in the military, and there is support for the view that sex discrepancies are in part due to heritable physical sex differences. \footnote{Women at War: An Evolutionary Perspective, supra note \textit{___}.} For reasons I will not belabor, academic discussions of occupational segregation often focus on academic jobs. Discrimination against women continues to be a problem, \footnote{See Section III.} and probably contributes to the lower proportion of women in scientific fields than in humanistic ones. Discrimination does not, however, seem to account fully for the pattern of women within the sciences. In 2005, women received 38\% of all Ph.D’s in science and engineering. \footnote{National Science Foundation, S & E Doctorate Awards: 2005, Table 2, Doctorates awarded to women, by field of study: 1996–2005 \url{http://www.nsf.gov/statistics/nsf07305/content.cfm?pub_id=3757&id=2}.} Women were better represented in biological than in other sciences, receiving 49\% of doctorates. This discrepancy might be explained in part by non-statistical discrimination. Because of the association of biology with medicine and other caring professions, women may be more accepted in these fields. The pattern within math and physical science, however, defies easy explanation in terms of irrational generalization. \footnote{A similar point is made in Women in science, supra note \textit{___}.} Women receive 34\% of doctorates in chemistry, 27\% in mathematics, 24\% in chemical engineering, but only 15\% in physics and 12\% in mechanical engineering. Similar patterns are found in the distribution of women in the science and engineering workforce. \footnote{National Science Foundation, Table H-5. Employed scientists and engineers, by occupation, highest degree level, and sex: 2006, \url{http://www.nsf.gov/statistics/wmpd/employ.cfm}; Census 2000 Special EEO File, \url{http://www.eeoc.gov/stats/census/index.html}.} The distinguishing feature of physics and mechanical engineering would appear to be their high reliance on three-dimensional spatial reasoning rather than on mathematical and general scientific ability. Though I am aware of no study directly addressing the issue, my strong impression is that negative views about women in the physical sciences and math cuts across the divide between spatial and other abilities. This divide, however, comports with the results of cognitive ability testing which shows by far the strongest male advantage in three-dimensional spatial ability, especially where moving objects are involved, as they would be in mechanical engineering.

Other types of occupational segregation, though, are not readily explained by cognitive differences. In fields in which women have cognitive advantages, they are usually better represented than in male-advantage fields. However, they are often still a minority, and regardless of their numbers often do not rise to the highest levels. This discrepancy obviously occurs in areas such as art, design, writing and psychology. More subtly, cognitive considerations intensify the puzzle of the glass ceiling that seems to keep women from the critically important managerial and executive positions at the summit of prestige and pay scales. Women have an advantage in the communication abilities and social perception that appear to be the cognitive traits most relevant to managerial ability.

Several possibilities might explain the apparent anomaly of continued male dominance in fields in which females have a cognitive advantage. Discrimination may be based on factors other than true statistical generalizations. Sex differences in personality might affect both worker...
productivity and worker preference for jobs. Finally, women might have a comparative advantage in nonmarket activities such as child-rearing. These issues are beyond the scope of this paper, though I explore them elsewhere.

VI. NORMATIVE CONSIDERATIONS

A. EFFICIENCY

Statistical discrimination can sometimes be economically efficient, and perhaps any gains in equity from prohibiting it must be weighed against accompanying efficiency losses. Perhaps the prohibition on statistical generalization should be rescinded altogether. Even those who would prohibit statistical discrimination may find efficiency considerations relevant to how vigorously the prohibition on statistical discrimination should be pursued.

1. FORCES MAKING STATISTICAL DISCRIMINATION EFFICIENT

The benchmark analysis made a long list of assumptions that eliminated efficiency effects. It assumed a single type of ability, but in real labor markets, different abilities are useful in different jobs, and individuals have higher output when matched with an appropriate job. Some of the qualities that determine optimal job match are observable, while others are not. Employers in a competitive economy will use all observable information to match employees with jobs, and for these purposes, there is no meaningful distinction between observables that are causally connected with productivity, such as education, and those that are merely correlated, such as sex. Disallowing the use of any observable information will decrease the quality of job match and therefore reduce efficiency.

The simple model assumed fixed labor supply. If labor supply can vary, efficiency requires that high ability workers work more than low ability workers, since the value of their work is greater. Allowing statistical discrimination means that high ability workers will receive higher wages and thus have the proper incentives to work more. Prohibiting statistical discrimination reduces wage dispersion, producing efficiency losses.

The benchmark model also assumed that each worker's output is independent of the quality of other workers. In real workplaces, lower ability workers may slow down higher ability workers, and production may be higher if workers are segregated by ability level. Statistical discrimination increases this beneficial segregation.

200 See, e.g., Kingsley R. Browne, Sex and Temperament in Modern Society: A Darwinian View of the Glass Ceiling and the Gender Gap, 37 ARIZONA LAW REVIEW 971 (1995); Forbidden Grounds, supra note__; Stereotype Threat, supra note__.
201 Bedrooms and Boardrooms, supra note__; Id.. See also Mary Anne Case, Of Richard Epstein and Other Radical Feminists, 18 HARVARD JOURNAL OF LAW & PUBLIC POLICY 369 (1994).
204 Id..
2. **Forces Making Statistical Discrimination Inefficient**

Relaxing other assumptions of the baseline model can make statistical discrimination inefficient. As we have already seen, statistical discrimination can cause inefficiency when workers make ability acquisition decisions based on their market opportunities. To the extent that differences are based on historical factors, a prohibition on statistical discrimination is desirable.

The benchmark model assumed a single sector whose identical employers all have the same imperfect evidence of ability. In an important 1986 paper, Stewart Schwab noted that self-employed people have better information about their own abilities than employers would. Market forces will thus tend to misallocate high ability workers to self-employment. With two groups, one of which has a higher mean ability than the other, statistical discrimination improves economic efficiency by drawing high ability members of the high average ability group into the standardized market, but reduces efficiency by pushing even more high ability members of the low ability group into self-employment. Whether efficiency requires allowing employers to use this information to discriminate statistically in hiring depends on which effect is larger.

In the benchmark model, employers had fixed information about employee abilities. However, job candidates will have an incentive to provide firms with additional information about their abilities through signals of that ability. Those signals must be credible, which requires that they be costly, and that those costs be lower for high ability workers than low ability workers. The canonical case of a signal is educational level, an example of education that can be misleading, since education (we hope) increases marginal product as well as providing a signal. Education is only a signal if its costs exceed the amount by which it raises marginal product. Signaling will raise wages to high ability workers, and is efficient only if that wage increase raises the labor supply of the more able enough to offset the signal costs. Higher levels of the signal will often lead to inefficient but stable equilibria.

The benchmark model also assumed fixed levels of employer investment in information. The usual market equilibrium level of such investment, however, is likely to be suboptimal. Other firms will have an incentive to poach the employees of an employer which invests resources in information about employee quality, and the law seldom offers compensation to the original employer. Firms collectively then will tend to under-invest in information, relying on employee generated signals such as education and any available proxies for ability such as sex. A ban on statistical discrimination may improve welfare by inducing firms to invest more in screening employees.

---

205 See I.B.2.


208 This point is made specifically in the context of incorrect sex stereotypes by An Economic Analysis of Sex Discrimination Laws, supra note ___ at 1319.

3. **IMPLICATIONS**

The strongest case against statistical discrimination can be made in simple short run models where the effects of allowing or banning such discrimination are purely distributional. In more complex models, a ban on statistical discrimination may cause inefficiency by reducing efficient sorting and, perhaps, by reducing wage dispersion. These losses must be balanced against gains from possible lower net distortion in labor supply, a reduction in inefficient signaling and improved overall information in the market.

In the longer run, when ability investment decisions are made, the case for intervention is stronger. When historical factors create misallocations in ability acquisition, these become self-fulfilling, and can be corrected only by intervention such as prohibition on statistical discrimination. Unfortunately, there are no clear criteria for determining when this adjustment is finished.

Economic models suggest that there is no general *a priori* efficiency case for or against statistical discrimination. The considerations are complex. Further work may make it possible to evaluate whether there is a net long run gain or loss from allowing statistical discrimination but no firm generalizations seem possible now.

**B. DISTRIBUTIVE CONCERNS**

For decades, most legal observers drew a sharp distinction between “true” anti-discrimination rules and accommodations. Accommodations imposed social costs, typically on employers, while “true” anti-discrimination rules, such as the prohibition of stereotype-based decisions did not. In an important article, Christine Jolls pointed out that the prohibition of much conduct classified as discrimination in fact constitutes in part an accommodation mandate. A policy that imposes accommodation costs makes sense only if those costs are balanced by gains in fairness.

If statistical discrimination is on balance efficient, its prohibition must be justified by appeal to distributive concerns. Most people, I suspect, would regard the principal case for prohibition to be meritocratic. It just doesn’t seem fair that two equally able people be treated differently simply because they are of different sexes. Such discrimination clearly places a burden on individuals who cannot engage in the activity at which they are most productive. Still, life is not always fair, and the legal system does not try to redress many inequities that are formally similar to statistical discrimination based on sex.

A more fundamental response to the meritocratic case against statistical discrimination rejects meritocracy altogether. Welfarist approaches, notably the social welfare functions of

---


211 For example, people believe that discrimination based on appearance is more common that race discrimination, Timur Kuran & Edward J. McCaffery, *Expanding Discrimination Research: Beyond Ethnicity and to the Web*, 85 Social Science Quarterly 713-730 (2004), yet there appears to be little social pressure to prohibit these forms of discrimination, see, e.g., Employment Law Alliance, NATIONAL POLL SHOWS PUBLIC OPINION SHARPLY DIVIDED ON REGULATING APPEARANCE - FROM WEIGHT TO TATTOOS - IN THE WORKPLACE, http://www.employmentlawalliance.com/en/node/1321 (last visited Jun 1, 2009).

Many observers have a tendency to assume that any distributive effect of the use of group characteristics is inherently invidious. For a thoughtful critique of this, see Colin S. Diver & Jane M. Cohen, *Genophobia: What Is Wrong with Genetic Discrimination*, 149 University of Pennsylvania Law Review 1439 (2000).
economics, assess outcomes by aggregating the well-being of all individuals in society. For welfarism, rewarding ability is of no importance as an end. The distribution of abilities is simply a fact to be worked with in maximizing welfare. Welfarist approaches tend to favor relatively equal distributions of income, subject to the constraints of efficiency losses. From this perspective, the distributive effects of statistical discrimination depend on many factors, including the response of employers. If employers do not respond to a ban on statistical discrimination by gathering more information, that ban will reduce the dispersion of income between men and women, which will tend to lower efficiency but be more equitable from a welfarist perspective.\textsuperscript{212} Suppose, however, that employers respond by gathering better information. This is not examined by the standard economic analysis, but I have suggested that it is both possible and efficiency-enhancing. However, the same information and consequent increased wage dispersion that improves efficiency will be distributionally undesirable.

Distributive analysis is also complicated by the fact that statistical discrimination operates symmetrically against individuals of both sexes in counter-typical occupations. If male and female occupations were in general of equal status and pay, statistical discrimination in labor markets might be no more problematic than other non-invidious types of statistical discrimination. Female-dominated occupations, however, are still typically lower paid than male-dominated occupations, and males tend to occupy the upper echelons of even female-majority careers. If women are, as group, less well off than men, welfarist considerations may support policies that redistribute to women, thus justifying prohibition of statistical discrimination.

Welfare economics can provide a tremendously clarifying starting point for analyzing policy, but it has been often criticized for omitting important concerns, such as moral entitlements.\textsuperscript{213} These concerns are particularly important in matters regarding the status of women.\textsuperscript{214} Even those of us who do not believe ability is generally relevant to distributive equity may feel that rewarding ability plays a role in efforts to raise the status of women. Recognizing merit in particular settings determines general social perceptions about the range of roles than women can play.

The normative analysis of statistical discrimination is riddled with uncertainty, and any policy prescription will unavoidably be made with reference to the prescriber’s own values. Many libertarians would argue that there should be a strong presumption against government action absent clear and compelling evidence of its necessity. Thus, they would respond to uncertainty by abstaining from intervention unless more facts become available.\textsuperscript{215} Those who place a high weight on equity may argue that even imperfect evidence of distributive problems creates a strong presumption in favor of intervention unless ancillary social losses are likely to be large.

My own inclination is to prefer in general the market to the state, but to find the current evidence of market failure sufficient to overcome a presumption against intervention. For this

\textsuperscript{212} Recall that in this situation, employers will estimate marginal product using a weighted average of the population mean (now including both men and women) and the individual test score. The use of a population mean rather than group means for men and women will reduce wage dispersion.

\textsuperscript{213} RICHARD S. MARKOVITS, TRUTH OR ECONOMICS, ON THE DEFINITION, PREDICTION AND RELEVANCE OF ECONOMIC EFFICIENCY (2008).


\textsuperscript{215} An Economic Analysis of Sex Discrimination Laws, supra note ___.
reason, I would continue Title VII’s prohibition on statistical discrimination, while reserving the option to re-evaluate that prohibition as social conditions change.

VII. DISCRIMINATION LAW AND POLICY

Any legal restriction on statistical sex discrimination must balance two competing concerns. Evidence of statistical sex differences cautions against embedding in the law any assumption that labor market outcomes would be the same for men and women in the absence of discrimination. At the same time, employers have an incentive to discriminate statistically, and continued evidence of discrimination suggests that the law should continue to scrutinize employment decisions carefully. This balancing act is further complicated by the highly tentative nature of evidence regarding both discrimination and difference. Policy towards discrimination must be made without any really clear picture of what a world without discrimination would look like.²¹⁶

A. THE USE OF NUMERICAL STANDARDS

1. STATISTICAL EVIDENCE IN THE PLAINTIFF’S CASE

In non-legal settings, discrimination is sometimes inferred when the representation of a group in a given job is not proportional to the composition of the population as a whole. Writing in the early 1990’s, critics including Kingsley Browne and Richard Epstein suggested that such inferences play an important role in the liability provisions of Title VII. These doctrines, they suggest, presume that men and women are inherently the same; cannot be reconciled with evidence of statistical sex differences; and create pressure on employers to hire according to population quotas regardless of merit.²¹⁷

Twenty years later, comparisons to population figures can seldom be the basis for the imposition of Title VII liability. Discrimination proof patterns distinguish between two theories of recovery, disparate treatment and disparate impact. Disparate treatment liability requires a showing of intent to discriminate, while disparate impact does not.²¹⁸ Each of these can be brought by either individual or class plaintiffs. In a class disparate treatment suit, the plaintiff must show that the defendant has an intentional policy or practice of discriminating against a protected group.²¹⁹ Class plaintiffs generally provide statistical evidence of defendant’s past

²¹⁶ This observation has been made by observers with a broad variety of perspectives. Doriane Lambelet Coleman, The Alchemy of a Moral Discourse about the Biology of Gender: Historical Sensitivity, Genetic Literacy, and the Will to Imagine a Different Equality, 11 CARDOZO WOMEN'S L.J. 543 (2004); Stereotype Threat, supra note___; Unconscious prejudice and accountability systems, supra note___; JOHN STUART MILL, THE SUBJECTION OF WOMEN (1869).


²¹⁸ Both theories were originally derived by courts from the core prohibition on discrimination contained in the original Civil Rights Act of 1964, although this prohibition does not clearly include either a requirement of intent or a distinction between intentional and non-intentional cases Civil Rights Act of 1964 § 703. (a), PL 88-352 88th Congress, H. R. 7152 as codified and amended 42 USC §2000e-2.

treatment of the protected group, but statistics alone suffice only if the discrepancies are extreme. In a disparate impact case, the plaintiff claims that one of defendant’s employment practices had a disparate impact on members of the protected class, even if that impact was unintentional. This claim is typically supported by statistical evidence of the effects of the practice on the protected group. If the plaintiff succeeds, the defendant may prevail by showing that the practice is job-related and consistent with business necessity. Disparate impact analysis spares the plaintiff the obligation of having to prove intent, while providing the defendant with the defense of business necessity, for which there is no analogue in disparate treatment theory.

Similar rules governing the use of statistics apply under both disparate treatment and disparate impact theory. In the 1977 hiring case Brotherhood of Teamsters v. United States, the Supreme Court did authorize the use of comparisons between the employer’s workforce and the general population or entire workforce. The Court, however, soon placed limits on the use of these comparisons, stating, in Hazelwood School District v. United States, that general population comparisons were not appropriate for jobs that required special skills, and that the proper benchmark was the “qualified …population in the relevant labor market.”

Initially, some lower courts did not apply the Hazelwood rule, and permitted general workforce comparisons for skilled jobs. In 1989, in Ward’s Cove Packing Co. v. Atonio, the Court strongly reiterated that statistical comparisons must be limited to the “ ‘qualified …population in the relevant labor market.’ ” At the time that Browne and Epstein wrote, their concern about the use of population-based standards had some foundation in the case law. Ward’s Cove was only a few years old, and it was too early to predict its impact. Twenty years after Ward’s Cove, it is now clear that courts adhere to the rule that general population comparisons can be used only in relatively low skill jobs. In hiring cases, courts look at the data on the relevant labor market for the job in question, which is now collected in great detail in the Census Special EEO File, a joint effort of the Census Bureau and several federal agencies.

---

220 Statistical evidence is not an absolute requirement in systemic disparate treatment cases, but virtually all plaintiffs attempt to provide such evidence. See, e.g., Pitre v. Western Elec. Co., 843 F.2d 1262 (10th Cir. 1988).
221 Hazelwood School District v. United States, 433 U.S. 299 (1977); Teamsters, supra note__.
222 The elements of a disparate impact case are now governed by statute. 42 USC § 2000e-2(k).
223 42 USC § 2000e-2(k).
225 Teamsters, supra note__.
226 Hazelwood School District v. United States, supra note__.
including the EEOC.\textsuperscript{230} Often, even relevant labor market data is reduced to include only actual applicants deemed qualified.\textsuperscript{231} Where employees challenge the employer’s promotions and other policies towards current employees, general labor market comparisons are not relevant, but courts implement the \textit{Hazelwood} principles by restricting comparison to the subset of eligible employees who meet relevant qualification standards.\textsuperscript{232}

The limits in current doctrine on the use of general population comparisons creates little risk that employers will be pressured into using population-based quotas in hiring. Indeed, the use of applicant pool and relevant workforce comparisons will tend to perpetuate the effects of any prior discrimination, whether statistical or otherwise, since that discrimination will be reflected in the applicant and workforce numbers.

\section{Voluntary Numerical Targets}

The imposition of Title VII liability does not embody any presumption that all groups be proportionately represented in every position. Such a presumption may, however, be implicit in the law governing numerical targets in voluntary affirmative action schemes.\textsuperscript{233}

In 1987, the Supreme Court addressed what numerical goals were permissible in a voluntary affirmative action plan for skilled positions.\textsuperscript{234} The plan promulgated by the Transportation Agency of Santa Clara County stated that its long-term goal was fill each job category by women and minorities in proportion to their presence in the total area labor force. For skilled positions, the plan did not attempt to achieve immediate parity but aimed instead at representation proportional to the qualified workforce. Following the earlier \textit{United Steelworkers of America v. Weber}.\textsuperscript{235} \textit{Johnson v. Santa Clara County} upheld the plan on the grounds that it had three characteristics of which are relevant here. First, the plan sought to remedy a “manifest imbalance” resulting from historical discrimination, even though there was no evidence of discrimination on the employer’s part. \textit{Johnson} elaborated the “manifest imbalance” requirement, holding that total labor market comparisons could be used in formulating long term goals, as long as narrower relevant labor market comparisons were use in formulating short term goals. Second, the plan was temporary, intended eliminate imbalance rather than to maintain balance.

These requirements of the \textit{Weber-Johnson} framework are strikingly similar to the requirements for economically efficient numerical targets to end statistically-based race discrimination.. As Steven Coate and Glenn Loury have pointed out, numerical targets may be more effective than costly monitoring of employers.\textsuperscript{236} With no heritable differences in ability

\begin{thebibliography}{99}
\bibitem{230} Census 2000 Special EEO File, \url{http://www.eeoc.gov/stats/census/index.html}.
\bibitem{231} EEOC v. Turtle Creek Mansion Corp., 70 FEP 899 (N.D. Tex. 1995), aff’d, 82 F.3d 414 (5th Cir. 1996); see generally \textsc{Ramona L. Paetzold \\& Steven L. Willborn}, \textsc{The Statistics of Discrimination} 4.3 (2008).
\bibitem{232} Morgan v. United Parcel Service of America, Inc., 380 F.3d 459 (2005); The statistics of discrimination, \textit{supra} note\textsuperscript{231} at 4.3 fn. 6.
\bibitem{233} Outside the doctrinal setting, policy analysis also seems to assume equal representation as norm. See, e.g., \textsc{Federal Glass Ceiling Commission, Bette Woody \\& Carol Weiss}, \textsc{Barriers to Work Place Advancement: The Experience of the White Female Work Force} (1993).
\bibitem{235} United Steelworkers Of America, AFL-CIO-CLC, 443 U.S. 193 (1979). That employer chose unskilled workers for a training program, and so the issue of relevant labor market did not arise.
\bibitem{236} Will Affirmative-Action Policies Eliminate Negative Stereotypes?, \textit{supra} note\textsuperscript{232}; Antidiscrimination enforcement and the problem of patronization, \textit{supra} note\textsuperscript{233}
\end{thebibliography}
distributions, long-run population targets are appropriate. A well-designed policy, they argue, aims for proportionate representation, but proceeds incrementally based on current ability levels to avoid reducing incentives for the disadvantaged to acquire abilities. As these policies change incentives for ability acquisition, differences diminish, the labor market reaches the optimal stable equilibrium, and antidiscrimination policies become superfluous.

When applied to sex, however, the Weber-Johnson framework for analyzing numerical targets becomes problematic. When discrimination clearly exists, short-run incremental goals based on the current qualified work force may reduce that discrimination. However, there is no simple way of knowing when sex discrimination has ended. General labor market targets assume the absence of heritable ability distributions, an assumption which in the case of sex is not supported by the available evidence. In occupations which should be disproportionately male or female, these targets would require employers to hire the opposite sex past the nondiscriminatory level.

Labor market or population-based numerical targets would harm women in fields that use abilities with a female cognitive advantage. Much of the discussion of discrimination within universities has focused on the minority status of women in the sciences. At least as troubling, however, is the failure of women to achieve majority status in fields like fine arts. In 1988, women made up 41.2% of post-secondary fine arts faculty. By 2005, this number had risen a mere 1.5% to 42.7%. Since virtually all abilities related to fine arts show a female advantage, cognitive considerations suggest that women should predominate in these departments. The humanities make use primarily of female-advantage skills like writing. Women now make up about half of all humanities faculty, but comparative advantage suggests that even this number might reflect discrimination if males with relevant talents are absorbed by strongly male-advantage fields like engineering. Any presumption that 50% representation is fair may disserve women in these fields.

More accurate long-run numerical targets, based on sex-differentiated ability distributions, could in theory reduce the problem of reverse discrimination against the more able sex. However, even a plan that used accurate long-run numerical targets could not, as Johnson and Weber require, be a means to attaining rather than maintaining a balanced workforce. With permanent difference and imperfect information, numerical targets must also be permanent or else the labor market will revert to a statistically discriminatory point. The purely legal aspect of this problem is not insuperable: The requirement that a plan be temporary could, of course, be dropped by judicial decision or statute. However, the Johnson-Weber requirement of a defined end point serves an important policy purpose. Absent other evidence of discrimination, even accurate labor market goals suffer from the Lake Wobegon problem. Not every firm can be above or even at the average. Firms that find themselves below the average might well want to examine their own procedures, and perhaps prophylactically reform those procedures even in the absence of evidence of discrimination. Nonetheless, in a world without discrimination it would

---

238 Humanities Resource Center Online, supra note__.
239 This process would not be prohibited by Ricci v. DeStefano, --- U.S. --- (2009), (last visited Jul 21, 2009), which was not argued as an affirmative action case. Instead, the employer claimed that abiding by its original procedures would have exposed it to Title VII liability.

45
not be merely superfluous but actually distortionary for them to engage in affirmative action using general labor market targets.

And formulating accurate numerical targets would be a daunting task. The efficient sex distribution in each occupation depends on the sex distribution of every other occupation. Calculating economy-wide efficient sex distributions would require analysis of the abilities required for all occupations; the true distribution of those abilities in a nondiscriminatory world; and the supply and demand in every output and labor market. This computation is simply not realistically possible.

B. STRUCTURAL APPROACHES

1. DISCRIMINATION TODAY

Statistical discrimination can in principle be addressed by the basic Title VII prohibition of discrimination. However, the efficacy of this ban has in recent years been eroded by changes in the form taken by discrimination. Current attempts to reduce discrimination are impeded by very success of Title VII in shaping public opinion. For some time after the passage of the Civil Rights Act of 1964, employers continued to make remarkably direct statements indicating an intent to discriminate. Eventually, they became more circumspect, even when consciously holding discriminatory views. Over time, many employers came genuinely to believe that discriminatory behavior was wrong, and this belief no doubt significantly reduced the level of discrimination in the workplace. Of course, reality often lags behind aspirations, and in an important article, Linda Hamilton Krieger documented evidence of discriminatory decision-making even by individuals who had a sincere desire to make non-discriminatory choices. Subsequent psychological research discussed by Krieger and others has provided additional evidence that the discriminatory decision-making that occurs today is principally unconscious.

The problem of unconscious discrimination has been compounded by changes in the structure of American business. At the time the Civil Rights Act of 1964 was passed, the typical American workplace was still influenced by the principles of scientific management pioneered

---

240 See Section V.
241 The Content of Our Categories, supra note ___. The importance of unconscious discrimination had earlier been discussed from a psychoanalytic point of view by Charles Lawrence. Charles R. Lawrence III, The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism, 39 STANFORD LAW REVIEW 317 (1986).

Some of this literature makes use of the Implicit Association Test (IAT), which attempts to measure an individual’s tendency to associate members of a group with certain generalizations. Research within the IAT framework has found only tentative evidence of a connection between the IAT and propensity to engage in discriminatory behavior. Antidiscrimination Law and the Perils of Mindreading, supra note ___. Banks, Eberhardt, and Ross, supra note ___. at 1188. However, substantial evidence of discriminatory behavior outside the IAT framework does exist. The Content of Our Categories, supra note ___. Section III supra. Once the case for the existence of discriminatory behavior has been made, IAT research becomes more important, suggesting that individuals who make discriminatory associations are often not aware that they do.
Job categories and hierarchies were clearly defined. Turnover was relatively low and compensation was based primarily on tenure. In the 1980's, this pattern began to change, and American workplaces became more fluid. Job descriptions captured less and less of the variation between the work actually done by different employees. Turnover increased, and pay was set more by supervisor assessments than under the old scientific management policies. Workplaces became less hierarchical, and fewer employees reported to a single supervisor.

Changes in workplace organization increased the importance of a phenomenon first described by Susan Sturm. In an important article, Sturm noted that much current discrimination does not fit the model of first generation discrimination, which often took the form of easily identified decisions in which discriminatory intent played an overt role. Today’s second generation discrimination results from the gradual accretion of small discriminatory choices made by a variety of actors, often leaving no written record. Contemporary discrimination, Sturm argued, is structural: it results from the newly opened-ended structure of workplace. Even conscious discrimination would be hard to monitor in a flexible workplace, and unconscious discrimination is still more difficult to control. The job standardization inherent in old-style scientific management facilitated the comparison of each employee’s work product. Career ladders and pay scales simplified compensation decisions. New and more loosely defined supervisory patterns, along with higher turnover by supervisors, make difficult the identification of the decision-maker whose intent is relevant. Since few human activities can be fully standardized, this second generation discrimination can arise in older Taylorist workplaces, but it is less common than in more flexible environments.

2. **Potential Changes in the Law**

A number of observers have proposed legal changes to address the problems caused by structural and unconscious discrimination. Often these proposals assume, implicitly or explicitly, that discrimination is irrational. But irrational and statistical discrimination are very different phenomena. To discriminate irrationally is to act on a negative generalization that is false. To discriminate statistically means to act on a generalization that is true but impermissible. Specific policies might correct one behavior and not the other.

Anti-discrimination policies can either insulate outcomes from discriminatory decisions or reduce the amount of discrimination in the first instance. Insulating strategies are typically useful regardless of the rationality or irrationality of the discriminatory belief. The most important insulating strategy is the simple legal prohibition of discrimination. The deterrent power of simple prohibition, however, depends on the employer’s awareness that it is discriminating. As Amy Wax has cautioned, even an explicit prohibition of unconscious or

---

246 The Law of Implicit Bias, *supra* note____.
structural discrimination may have no deterrent effect, since employers may have no way of preventing wrongs of which they themselves are not aware.\textsuperscript{247}

Numerical affirmative action can insulate employment outcomes from both unconscious and conscious discrimination,\textsuperscript{248} and, like most insulating strategies, is effective against both irrational and statistical discrimination. However, numerical targets are more useful in combating race discrimination than sex discrimination because of the existence of actual sex differences.

In contrast to insulating strategies, debiasing strategies address irrational discrimination by attempting to correct a false generalization. No such correction is possible in the case of statistical discrimination, which is by definition based on a true generalization. To prevent statistical discrimination, the decision-maker can either be provided with better information about the individual candidate, or induced to ignore the correct statistical generalization and unlink it from the specific decision. Both individual information strategies and unlinking strategies may sometimes be useful in correcting irrational generalizations as well. Structured decision making, a technique developed to control imperfectly rational decisions, can both unlink and improve information quality, and can thus be useful in controlling statistical discrimination.\textsuperscript{249}

3. \textbf{STRUCTURED DECISION MAKING}

A large body of research now supports the view that structured decision procedures are almost always as good as, and usually superior to, unstructured ones.\textsuperscript{250} Decision procedures have two components: the mode of data collection and the mode of combining data into a final decision. Structured decision procedures draw on a fundamental distinction between algorithmic and clinical modes of data collection and combination. Algorithmic judgments are based on a clearly specified procedure that contains no discretionary elements. Clinical judgments are made by a human observer, often an expert, without the guidance of an explicit procedure.\textsuperscript{251} Algorithmic judgments seem to be as good or superior for combining data, but both clinical and algorithmic judgments are useful for collecting data.\textsuperscript{252} Structured decision procedures make use of these findings by using algorithmic judgments for data aggregation but either clinical or algorithmic judgments for data collection. The superior performance of structured procedures has been strongly confirmed in hiring interviews and more tentatively confirmed in assessing job

\textsuperscript{247} Amy Wax, \textit{Discrimination as Accident}, 74 \textit{INDIANA LAW JOURNAL} 1129 (1999).
\textsuperscript{248} The Law of Implicit Bias, \textit{supra} note\_,
\textsuperscript{249} Amy Wax notes the success such mechanisms have had in medical diagnosis and their potential application to employment decisions. Discrimination as Accident, \textit{supra} note\_
\textsuperscript{251} The foundational analysis is Paul E. Meehl, \textit{Clinical Versus Statistical Prediction: A THEORETICAL ANALYSIS AND A REVIEW OF THE EVIDENCE} (1954). Meehl and others have used the terms “statistical”, “mechanical”, and “objective” equivalently to what I am calling “algorithmic” here. The term “clinical” is Meehl’s. Other researchers have referred to “holistic” or “subjective” judgments. This profusion of terminology has led to a great deal of confusion in the personnel literature, Rationally Ignorant Discrimination, \textit{supra} note\_.
performance.\textsuperscript{253} The weaker consensus for performance evaluation appears to result from the smaller number of studies rather than from less conclusive results in individual studies.\textsuperscript{254}

No study of structured procedures has explicitly examined whether structure reduces irrational or rational discrimination, but the effectiveness of structure does not obviously seem to be related to the rationality of the subject’s belief. Structure can increase the information the subject has available about the individual candidate, and this availability by itself should reduce the weight given to group generalizations, possibly even more for rational than for irrational generalizations. Structure takes data combination rules out of the subject’s control, and these structured rules can always be designed to eliminate consideration of undesired factors, whether rational and irrational. Structure, however, does not prevent subjects from importing generalizations into data collection, even though these generalizations are necessarily irrelevant. A conscious discriminator, rational or irrational, might well try to circumvent the structure by introducing generalizations into the data collection process. Unconscious discriminators seem more likely to attempt to focus on the task at hand without importing irrelevant generalizations.

Studies suggest that most people are less discriminatory in data collection than in data aggregation. In one series of experiments,\textsuperscript{255} subjects were asked to select individual candidates for a traditionally male job, chief of police. Subjects were given a written description of a candidate with a male or female name and one of two clusters of credentials relevant to the job. The streetwise applicant was tough, had worked in rough neighborhoods, and got along with other officers. However, this applicant was poorly educated and lacked administrative skills. The educated applicant was well schooled and experienced in administration but had little street experience and got along poorly with fellow officers. Subjects were asked which applicant should be hired. After this decision, subjects rated the strength of the applicant for a series of streetwise credentials(tough, risk taker, physically fit) and a series of educated credentials (has political connections, able to communicate with media) The applicant’s sex had no significant effect on how subjects perceived the strength of specific credentials. However, after-the-fact ranking of job qualifications varied so as to justify a preference for a male candidate. This inconsistency in candidate rankings greatly diminished in a second study when subjects were first asked to rank the importance of different job qualifications and then individual candidates.\textsuperscript{256}


\textsuperscript{254} \textit{Id.}. The smaller number of studies reflects the greater difficulty of studying performance decisions. \textit{Id.}.


\textsuperscript{256} When work product is complex, the tendency to introduce generalizations may be stronger, and such evaluations should probably also be structured. Rationally Ignorant Discrimination, \textit{supra} note___.
Structure can be introduced into personnel policies at any number of levels. Some of these may be expensive. To meet the highest standards of personnel professionals, the structuring process should begin with a formal job analysis that identifies specific behavior that constitutes good, average and poor performance.\textsuperscript{257} But formal job analysis is not the critical component of structure, and algorithms derived from clinical judgment may be quite effective. The crucial factor is the \textit{ex ante} determination of the structure of the evaluation to prevent the evaluator from introducing undesired considerations into the final decision.\textsuperscript{258}

C. STRUCTURE UNDER TITLE VII

1. \textsc{In General}

Nothing in Title VII authorizes courts to dictate the procedures that employers must use in personnel decisions. Still, various doctrinal issues invite courts to create incentives for employers to use certain employment practices. Courts have shown sympathy to the principle that excessively “subjective” procedures open the door to discrimination,\textsuperscript{259} but are understandably concerned about whether they are competent to evaluate those procedures, since subjectivity is a legitimate component of personnel decisions. In practice, courts often place defer to employers, and challenges to subjective processes often fail.\textsuperscript{260}

Rather than focusing on subjectivity in isolation, the law should examine the extent to which subjective judgments are constrained by a structured decision process. Emphasizing structure permits legitimate subjective judgments while providing a relatively clear standard to evaluate personnel procedures as a whole. Incentives to adopt structured practices would not place on employers a burden to learn some arcane set of human relations practices. These practices embody a principle most of us learned on the playground: rules are set before the game is played, and can’t be changed afterwards. Title VII does not require employers to use specific procedures when making personnel decisions. However, the purposes of Title VII are best served if courts create incentives for employers to use methods that do not resemble Calvinball.

\begin{itemize}
\item \textsuperscript{257} Michael A. Campion, David K. Palmer & James E. Campion, \textit{Structuring Employment Interviews to Improve Reliability, Validity, and Users’ Reactions}, 7 \textsc{Current Directions in Psychological Science} 77-82 (1998); A Review of Structure in the Selection Interview, \textit{supra} note.\textsuperscript{___}.
\item \textsuperscript{258} Rationally Ignorant Discrimination, \textit{supra} note.\textsuperscript{___}. Melissa Hart provides an excellent discussion of ways informal structure has been used to avoid discriminatory decisions. Melissa Hart, \textit{The Possibility of Avoiding Discrimination: Considering Compliance and Liability}, 39 \textsc{Connecticut Law Review} 1623 (2007).
\item \textsuperscript{259} Subjective Decisionmaking And Unconscious Discrimination, \textit{supra} note.\textsuperscript{___}.
\item \textsuperscript{260} \textit{Id.}.
\end{itemize}
Judges concerned about the limits of their institutional competence often look to academic research, primarily through the testimony of experts in individual cases. The views of academic psychologists have long played an important role in the law of employment discrimination and their testimony should suffice to delineate a simple distinction between structured and unstructured practices as source of concern.

A preference for structured procedures would serve two purposes. First, structured procedures seem less likely to lead to discriminatory decision-making. Second, unstructured procedures make judicial review of decisions difficult. The relevant information may not be available if the defendant makes personnel decisions in an unstructured way that leaves no paper trail. Employers therefore have a perverse incentive to adopt procedures most likely to produce discriminatory outcomes.

Whether the law should go beyond a simple skepticism about *ex post* explanations presents difficult questions that I consider in detail elsewhere. More detailed scrutiny of employment processes might place a burden on smaller firms who could not afford sophisticated legal or personnel advice. Complex rules also have the potential to create legal uncertainty, raising transactions costs. If such rules are to be adopted, they should be embodied in EEOC guidelines so as to give employers some advance notice of what is expected of them.

**2. Individual Disparate Treatment Cases**

Most discrimination suits are brought by individual plaintiffs on a theory of disparate treatment. The plaintiff bears the burden of proving both discriminatory action and intent. As existing case law recognizes, procedures that make no effort to reduce discriminatory decisions are probative, though not necessarily conclusive, of intent. Focusing on structure rather than subjectivity should ease judicial concerns about institutional competence, and the weight given to evidence of faulty procedures could be increased.

After proving intent, plaintiffs still bear the burden of proving discriminatory conduct, which they often attempt to meet by showing that comparable employees outside the protected class were treated better, or that the defendant’s mischaracterized the plaintiff’s performance. The use of unstructured procedures should be relevant to determining whether a discriminatory outcome has occurred. Where unstructured decision making is combined with ambiguous evidence of how comparable employees were treated, the employer should not benefit from having obscured the merits of its decisions.

**3. Class Actions**

A number of observers have suggested that systematic problems in decision-making procedures are better addressed though class actions that through individual litigation. Among

---

261 Rationally Ignorant Discrimination, *supra* note__.
263 Rationally Ignorant Discrimination, *supra* note__.
264 Subjective Decisionmaking And Unconscious Discrimination, *supra* note__.
other things, class litigation is better suited to producing injunctive relief in the form of revamped personnel procedures.

Class certification poses a serious obstacle to class challenges of unstructured practices. Members of any class action must have common interests and underlying complaints. However, despite contrary Supreme Court authority, many courts have concluded that the absence of practice is not itself always a practice. They regard determination of when excessive discretion is a common practice to which all class members were subject as a problem of the merits of the case so that determining class certification creates a vicious circle. The merits cannot be decided unless a class is certified, yet a class cannot be certified without determining the merits of whether the defendant has engaged in a cognizable practice. Shifting focus from subjectivity to structure should simplify the definition of a “practice”, and make courts more willing to permit class certification.

Once a class is certified, plaintiffs proceed on a class disparate treatment theory, a class disparate impact theory, or both. The use of unstructured procedures is relevant in slightly different ways to each theory. In a systemic disparate treatment suit, the plaintiff must show that the defendant intentionally discriminated against a protected group. Just as in individual disparate treatment cases, unstructured procedures should be, though they are not always now, regarded as relevant to the determination of intent, especially where the alleged intent is unconscious or structural. Again, as in individual cases, the use of unstructured procedures may be useful in determining the existence of a discriminatory outcome where statistical or comparative measures are not dispositive.

The plaintiff in a disparate impact suit claims that one of defendant’s employment practices had a disproportionately negative effect on members of the protected class, even if that effect was unintentional. If the plaintiff succeeds in this, the defendant has the opportunity to provide evidence that that practice is job-related and consistent with business necessity. Under Watson v. Fort Worth Bank and Trust, the employer’s delegation of subjective decision-making power to supervisors may be challenged as a “practice” and plaintiffs can make a prima facie case if the discretion tends to disfavor members of the protected group. Because of concerns about the institutional competence of courts, Watson imposed on an easier standard of business necessity for so-called subjective practices than objective ones. Again, shifting the focus from subjectivity to structure would improve judicial oversight of employer discretion. Structured personnel procedures permit the legitimate subjective judgment of supervisors about hard-to-quantify performance attributes, and subjective decisions that are part of structured procedures deserve the deference accorded by Watson. Highly unstructured decisions procedures, however,
warrant much closer scrutiny, and employers who use them should bear a heavy burden of showing business necessity.

D. BEYOND EMPLOYMENT DISCRIMINATION LAW

1. THE TAX SYSTEM

Statistical discrimination is a partially distributive problem resulting from the operation of the labor market. One of the great achievements of law and economics has been to emphasize that distributive problems created in a given market are not always best solved solely by direct intervention in that market. In particular, the tax code may be a useful tool for redistribution.274

Employers hire, they set pay, they promote, and they fire. If the legal system polices one of these decisions more stringently than others, employers may simply shift discrimination to the least monitored choice.275 Edward McCaffery has suggested that pay discrepancies may be the least pernicious of these, and proposes that we consider weakening the prohibition on pay discrimination while compensating women by taxing the market income of secondary earners at low or even negative rates.276

McCaffery’s proposal was addressed to statistical discrimination based on women’s greater likelihood to quit to care for children, which is currently increased by the tax code’s exclusion household production and high rates on wages of secondary earners at.277 A policy towards ability differences that was precisely analogous to his proposed policy towards workforce attachment would require the unrealistic policy of basing tax rates on the sex-role congruence of an individual’s job. Nonetheless, McCaffery’s original proposal to reduce taxes on married women’s labor income might have a greater impact on women in role-incongruent jobs, since these women may be more often discouraged from workforce participation. Though the proposal would not directly help single women, it might have beneficial indirect effects. Each workforce exit motivated by statistical discrimination is both problematic in itself, and has a negative externality. By reducing the number of qualified members of the relevant sex, it distorts both rational and irrational employer perceptions about the number of qualified members of that sex, and creates a vicious circle of exclusion.

274 Louis Kaplow & Steven M. Shavell, Why the legal system is less efficient than the income tax in redistributing income, JOURNAL OF LEGAL STUDIES 667-681 (1994). Though the tax system may not always be the most efficient tool for redistribution, Truth or Economics, On the Definition, Prediction and Relevance of Economic Efficiency, supra note___; Chris William Sanchirico, Taxes versus Legal Rules as Instruments for Equity: A More Equitable View, 29 JOURNAL OF LEGAL STUDIES 797-820 (2000), it may still complement intervention in the problematic market.


276 Slouching Towards Equality, supra note___.

2. EDUCATION AND CHILD-REARING

Statistical discrimination can occur in contexts other than employment. Fifty years ago, the educational system formally engaged in discrimination, prescribing different elementary and secondary courses for boys and girls and admissions quotas against women at the graduate level. Tremendous progress has been made, and formal discrimination has been eradicated. However, the educational system, like the workplace, contains many subjective decisions, and might benefit from the introduction of the same formalization of these decisions that I have advocated in employment settings.

All of the policies considered up until this point tacitly assume that girls and boys are raised in the same environment, permitting heritable differences to express themselves. A more interventionist approach would try to narrow the gap between boys and girls by providing each with training designed to compensate for their relative weaknesses, either through single-sex education, or through sex-blind education that attempted to develop more balanced abilities on an individual basis, addressing statistical sex differences incidentally. Considering this possibility is a valuable mental exercise that helps contain two critical errors, the naturalistic fallacy and essentialism.

VIII. CONCLUSION

The extent to which occupational segregation is due to discrimination or heritable difference is the subject of intense and not always civil debate. Both sides can muster considerable evidence in support of their position, and both discrimination and heritable difference seem to play some role. The coexistence of both forces, moreover, is not mere chance. Differences in heritable ability distribution create a strong incentive for imperfectly informed employers to use sex as a signal of ability, and thus to discriminate statistically.

At this time, the relative contribution of difference and discrimination to occupational segregation cannot be estimated with any confidence. Statistical discrimination and comparative advantage can both advantage magnify the effects of small differences. Our understanding of sex differences is still far from complete, and even the relevant cognitive categories are still at issue. Evidence of the skills required for specific occupations is surprisingly sparse.

Reasonable minds may differ about whether statistical discrimination should be prohibited. The evidence now available does not permit any firm conclusions about the net efficiency effects of statistical discrimination. However, ending statistical discrimination would tend to improve the general perception of women's competence, and would redistribute towards a group that is still economically less well-off. My own view is that, at the present time, the strength of the distributive case overcomes the indeterminacy of the efficiency case. However, as circumstances change, a different set of trade-offs may emerge.

The achievements of Title VII should not be understated, but its further effectiveness is now limited by its own effectiveness in stigmatizing intentional discrimination and the increasingly fluid structure of the American workplace. Numerical affirmative action may provide some short term amelioration of statistical sex discrimination, but it cannot be a long term answer. A goal of equal representation can hurt women, since in some fields, women do or

278 Engendered Differences, supra note ___
should have a numeric advantage over men. Setting accurate long term goals is infeasible because of the complexity of calculating what a discrimination-free workplace would look like.

Title VII does not empower courts to dictate personnel procedures to employers. However, discrimination law has long created incentives for and against the use of certain approaches. Employers might be encouraged to use more structured decision making to provide decision makers with better information and to eliminate consideration of discriminatory factors. Structured procedures are both less likely to lead to discriminatory decision-making and facilitate later review of decisions.

The problems surrounding occupational segregation may have their solution outside the labor market. Perhaps the tax system can be used to address the distributive problems that arise from statistical discrimination. And we need not take sex differences as given. Natural does not mean good, and perhaps both sexes would be better off if they were socialized to acquire the skills more often found in the opposite sex.

Those concerned about the status of women are understandably uncomfortable with any claims about differences in ability distributions between women and men. Such claims have long been used to prevent women from realizing their potential as full participants in human affairs. These concerns, however, must not obscure our evaluation of the growing body of evidence that some abilities are indeed differently distributed, on average, between men and women. These differences have often been overstated and over-generalized, but they are real.

To acknowledge that differences exist, however, does not mean accepting the status quo. Differences and discrimination not only can coexist, but almost inevitably do. The challenge we now face is to reformulate discrimination policy based on a deeper understanding of both forces.