The Paradox of Statistical Discrimination

Deborah Weiss

Abstract

A growing body of research suggests a biological basis for some of the differences between men and women in the distribution of aptitudes. Such a biological basis, however, produces a paradox. As is often noted, it suggests that men and women would not be equally distributed among occupations even in a perfectly fair society. At the same time, a statistical aptitude difference will almost inevitably cause discrimination in professions that make use of that aptitude. The larger the difference, the more likely discrimination is. This paradoxical result is supported by empirical evidence about the continued existence of discrimination.

Aptitude 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. Affirmative action policies are problematic: target goals are impracticable to set and, even if realistic, would need to become a permanent institution. Closer scrutiny of employer practices holds some promise as an alternative to numerically based theories.

Preliminary draft 5.3. Comments welcome; not for quotation or citation without permission.

Occupational segregation by sex is widespread. Most engineers are men, while most social workers are women.

A growing body of research suggests that the distribution of certain aptitudes differs somewhat between men and women. This research has generated a great deal of controversy. Some observers believe that the weight of the evidence now supports the hypothesis of statistical sex differences while others vigorously dispute this.

Even those who agree that some statistical differences exist are sharply divided about the size and implications of these differences. Some believe that existing occupational segregation is primarily the result of such differences, and should not be a major policy concern. Others have argued that any differences are small and cannot explain any significant amount of occupational segregation, which instead results from discrimination.

In this Article I will argue that research showing some biological basis for aptitude differences is too strong to ignore. Aptitude distributions between males and females overlap greatly, but they are not identical. These differences cannot entirely be explained by enculturation.

A biological basis for different aptitude distributions produces a paradox. It suggests that men and women would not be equally distributed among occupations even in a perfectly fair society. At the same time, a statistical aptitude difference between men and women will almost inevitably cause discrimination in professions that make use of that aptitude. The larger the difference, the more likely that discrimination is. This paradoxical result is supported by empirical evidence about the continued existence of discrimination. To further complicate matters, the existence of differing aptitude distributions implies that sex discrimination is not a problem that can be eliminated once and for all. Rather, it is a permanent feature of labor markets.

Aptitude distribution differences and the resulting permanent state of statistical discrimination creates 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. Affirmative action policies are problematic: target goals are impracticable to set and, even if realistic, would need to become a permanent institution, a result that many find undesirable. 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 a realistic view of the forces that underlie discrimination.

I. OCCUPATIONAL SEGREGATION IN INDUSTRIAL ECONOMIES

The sexual division of labor is as old as human history, and some degree of 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.

Studying the nature of occupational segregation is beset by several methodological problems. Existing scholarship comes from very several different social sciences, including anthropology, sociology, psychology and economics. Each of these fields has its own analytic framework and techniques, creating some problems in synthesizing findings from various fields. Among other problems, no single number can fully summarize a society's occupational segregation, and each discipline has favored various measures. Nonetheless, several general patterns can be discerned, as well as several areas of variability.

In studies of industrialized societies, the so-called dissimilarity index is number most commonly used to measure occupational segregation and provides a general picture.[1] This index, roughly speaking, 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 (roughly 66%), and perhaps even rose slightly.[2]

---

[1] O. D. Duncan & B. Duncan, *A Methodological Analysis of Segregation Indexes*, 20 AMERICAN SOCIOLOGICAL REVIEW 210-217 (1955).

[2] Jerry A. Jacobs, *Long-Term Trends in Occupational Segregation by Sex*, 95 AMERICAN JOURNAL OF SOCIOLOGY 160 (1989). See also E. Gross, *Plus Ca Change...? The Sexual Structure of Occupations over Time*, 16 SOCIAL PROBLEMS 198-208 (1968).

	Blau, Simpson and Anderson(1998)	Jacobs (1989) Method I	Jacobs (1989) Method II	Jacobs (1999)
1900		66.9-64.9*		
1910			63-69*	
1950		67.3		
1970	67.68		67.6	
1980	59.25	0	59.8	
1986			57.3	
1990	52.98			56.4
1997				53.9

*Indicates range

TABLE II
TRENDS IN U.S. INDEX OF DISSIMILARITY

These studies suggest that the index dropped at a rate of about .8% per year during the 1970's,[3] .42% per year during the 1980s,[4] and .36% per year during the early 1990s.[5] 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.[6]

Claudia Goldin's study of subsectors of the economy suggests that the numbers in Table I are a lower bound for the index of dissimilarity, especially in the earliest reported period. Goldin finds that the overall trend suggested by workforce studies is probably correct, but there may have been some desegregation between 1900 and 1970.[7]

Similar patterns and trends in occupational segregation have been obtained for Britain.[8] For other countries, fewer long term studies exist, but research on recent decades shows a roughly similar pattern, albeit with

---

[3] Francine D. Blau, Patricia Simpson and Deborah Anderson, *Continuing Progress? Trends in Occupational Segregation in the United States Over the 1970s and 1980s*, 4 Feminist Economics 29 (1998); Jacobs, *supra* note___; FRANCINE D. BLAU, MARIANNE A. FERBER & ANNE E. WINKLER, THE ECONOMICS OF WOMEN, MEN AND WORK (3rd ed. 1998).

[4]Jacobs, *supra* note___.

[5]Jerry A. Jacobs, *The sex segregation of occupations: Prospects for the 21st century*, *in* HANDBOOK OF GENDER AND WORK 125-144 (Gary N. Powell, ed., 1999).

[6] Claudia Goldin, *From the Valley to the Summit: The Quiet Revolution that Transformed Women's Work*, NBER WORKING PAPER W10335 (2004).

[7]CLAUDIA GOLDIN, UNDERSTANDING THE GENDER GAP: AN ECONOMIC HISTORY OF AMERICAN WOMEN (1994).

[8]C. Hakim, *A century of change in occupational segregation 1891-1991*, 7 JOURNAL OF HISTORICAL SOCIOLOGY 435-454 (1994).

more occupational segregation than in the U.S.[9] Initially steep drops during the 1970s have been followed by steady but less rapid drops in subsequent decades.

This analysis of occupational segregation includes only market work and ignores non-market household production such as child raising. These activities are traditionally performed primarily and sometimes exclusively by women. Including these activities would probably increase observed occupational segregation.

The slowing rate of change raises important issues. 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. If a truly nondiscriminatory society would have no occupational segregation, then the diminishing rate of decline in occupational segregation is cause for concern. If such a society might exhibit occupational segregation, then we cannot assume that existing differences result from discrimination, though they might.

Researchers who have examined occupational segregation have offered roughly two explanations for observed patterns of sex segregation: differences in qualifications and discrimination. The next Section examines the evidence that biologically influenced differences cause occupational segregation.

II. DIFFERENCE

A. EVIDENCE OF APTITUDE DIFFERENCES

General intelligence, sometimes called *g*, is an important predictor of success in most occupations.[10] The measurement of general intelligence

---

[9] J. Dolado & F. Felgueroso, *Where do Women Work? Analysing Patterns in Occupational Segregation by Gender*, ANNALES D'ECONOMIE ET DE STATISTIQUE 293 (2003).

[10] J. E. Hunter & F. L. Schmidt, *Intelligence and job performance: Economic and social implications*, 2 PSYCHOLOGY, PUBLIC POLICY, AND LAW 447-472 (1996); F. L. Schmidt & J. Hunter, *General Mental Ability in the World of Work: Occupational Attainment and Job Performance*, 86 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 162-173 (2004); N. R. Kuncel, S. A. Hezlett & D. S. Ones, *Academic performance, career potential, creativity, and job performance: can one construct predict them all?*, 86 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 148-61 (2004). For a review including some misgivings about the use of *g*, see C. Viswesvaran & D. S. Ones, *Agreements and Disagreements on the Role of General Mental Ability (GMA) in Industrial, Work, and Organizational Psychology*, 15 HUMAN PERFORMANCE 211-231

is a complex matter, but most current research finds little or no difference in the mean level of general intelligence for men and women.[11]

In many occupations, specialized abilities are also valuable. In a number of measures of these abilities, men and, women also perform similarly, and it is probably fair to say that men and women are more similar than different in their cognitive skills.[12] However, men and women on average do perform differently on tests of some specific skills that might be relevant to occupational specialization.[13]

a) *Motor skills.*

One of the largest differential aptitude distributions is found in targeting skill, both throwing and intercepting. Men consistently outperform women on these tasks.[14] Women perform better than men on many fine motor tasks, beginning at an early age and even when finger size is not a factor.[15]

(2002). Even skeptics agree that g can have a significant predictive value in at least some situations. *Id.*

[11]Research finding no sex difference includes R. Colom et al., *Negligible Sex Differences in General Intelligence*, 28 INTELLIGENCE 57-68 (2000); A. R. JENSEN, THE G FACTOR: THE SCIENCE OF MENTAL ABILITY (1998); N. J. Mackintosh, *Sex differences and IQ.*, 28 JOURNAL OF BIOSOCIAL SCIENCE 559-71 (1996). Studies showing a difference include R. Lynn, *Sex differences in intelligence and brain sizes: a paradox resolved*, 17 PERSONALITY AND INDIVIDUAL DIFFERENCES 257-271 (1994); R. Lynn, *Sex differences in intelligence and brain size: a developmental theory*, 27 INTELLIGENCE 1-12 (1999).

[12] J. S. Hyde, *The Gender Similarities Hypothesis.*, 60 AMERICAN PSYCHOLOGIST 581-592 (2005).

[13] E. Govier & J. Feldman, *Occupational choice and patterns of cognitive abilities*, 90 BRITISH JOURNAL OF PSYCHOLOGY 99-108 (1999).

[14] Neil V. Watson & Doreen Kimura, *Nontrivial Sex Differences in Throwing and Intercepting: Relation to Psychometrically-defined Spatial Functions*, 12 PERSONALITY AND INDIVIDUAL DIFFERENCES 375-385 (1991).

[15]D. Ingram, *Motor asymmetries in young children.*, 13 NEUROPSYCHOLOGIA 95-102 (1975); Karen G. Nicholson & Doreen Kimura, *Sex differences for speech and manual skill.*, 82 PERCEPTUAL AND MOTOR SKILLS 3-13 (1996); Doreen Kimura & C. H. Vanderwolf, *The relation between hand preference and the performance of individual finger movements by left and right hands.*, 93 BRAIN 769-74 (1970); Jonathan Lomas & Doreen Kimura, *Intrahemispheric interaction between speaking and sequential manual activity.*, 14 NEUROPSYCHOLOGIA 23-33 (1976); M. B. Denckla, *Development of motor co-ordination in normal children.*, 16 DEVELOPMENTAL MEDICINE AND CHILD NEUROLOGY 729-41 (1974).

b) *Spatial skills.*

Men are on average significantly better at mentally rotating objects.[16] The sex difference diminishes when meaningful rather than abstract objects are to be rotated.[17] Men have a moderate advantage in spatial perception,[18] and a have much smaller advantage in visualization.[19]

Women are on average better at remembering the location of objects in an array.[20] This difference may be associated with the fact that men and women on average take different approaches to navigation. Women are more likely to use relative directions (left and right) and landmarks.[21] Men are more likely to use absolute reference points such as north, south, east or west and their spatial reasoning is typically less field (i.e. context) dependent than women's.[22]

c) *Mathematics.*

In general, women score higher than men on tests of math computation, while men score higher on tests of math problem solving,

---

[16]D. Voyer, S. Voyer & M. P. Bryden, *Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables.*, 117 PSYCHOL BULL 250-70 (1995); Watson and Kimura, *supra* note___; D. W. Collins & Doreen Kimura, *A large sex difference on a two-dimensional mental rotation task.*, 111 BEHAVIORAL NEUROSCIENCE 845-9 (1997).

[17] G. M. Alexander & M. Evardone, *Blocks and bodies: Sex differences in a novel version of the Mental Rotations Test*, 53 HORMONES AND BEHAVIOR 177-184 (2008).

[18] Voyer, Voyer, and Bryden, *supra* note___.

[19] *Id.*

[20] M. Eals & I. Silverman, *The Hunter-Gatherer Theory of Spatial Sex Differences: Proximate Factors Mediating the Female Advantage in Recall of Object Arrays*, 15 ETHOLOGY AND SOCIOBIOLOGY 95-95 (1994); D. H. McBurney et al., *Superior spatial memory of women: Stronger evidence for the gathering hypothesis*, 18 EVOLUTION AND HUMAN BEHAVIOR 165-174 (1997).

[21]S. L. Ward, N. Newcombe & W. F. Overton, *Turn Left at the Church, Or Three Miles North: A Study of Direction Giving and Sex Differences*, 18 ENVIRONMENT AND BEHAVIOR 192 (1986); L. A. Galea & D. Kimura, *Sex differences in route-learning.*, 14 PERSONALITY AND INDIVIDUAL DIFFERENCES 53-65 (1993) (finding evidence of correlation between mental rotation task and navigational strategies)

[22] M. Robert & F. Harel , *The gender difference in orienting liquid surfaces and plumb-lines: its robustness, its correlates, and the associated knowledge of simple physics.*, 50 CANADIAN JOURNAL OF EXPERIMENTAL PSYCHOLOGY 280-314 (1996); H. A. Witkin, *A Cognitive-Style Approach to Cross-Cultural Research*, 2 INTERNATIONAL JOURNAL OF PSYCHOLOGY 233-250 (1967); R. Vasta, C. Lightfoot & B.D. Cox, *Understanding Gender Differences on the Water-Level Problem: The Role of Spatial Perception*, 39 MERRILL-PALMER QUARTERLY 391-414 (1993).

including word problems.[23] The difference appears to increase with the difficulty of the questions[24] and the selectivity of the sample population.[25] No sex difference appears in understanding of mathematical concepts[26] and, possibly, on tests of geometric proof writing.[27] The variance in male scores is greater, with males over-represented at both tails of the distribution.[28]

d) *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.[29]

[23] J. S. Hyde, E. Fennema & S. J. Lamon, *Gender differences in mathematics performance: A meta-analysis*, 107 PSYCHOLOGICAL BULLETIN 139-155 (1990). Since there is no significant male-female difference in general intelligence, general intelligence may be driving some of the convergence in scores. M. Brunner, S. Krauss & M. Kunter, *Gender differences in mathematics: Does the story need to be rewritten?*, INTELLIGENCE (2007); W. Johnson & T. J. Bouchard, *Sex differences in mental abilities: g masks the dimensions on which they lie*, 35 INTELLIGENCE 23-39 (2007).

[24] A. M. Penner, *International gender× item difficulty interactions in mathematics and science achievement tests*, 95 JOURNAL OF EDUCATIONAL PSYCHOLOGY 650–655 (2003);

[25] Hyde, Fennema, and Lamon, *supra* note___.

[26] *Id.*

[27] 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.

[28] 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); L. V. Hedges & A. Nowell, *Sex differences in mental test scores, variability, and numbers of high-scoring individuals*, 269 SCIENCE 41-45 (1995); A. Feingold, *Sex Differences in Variability in Intellectual Abilities: A New Look at an Old Controversy*, 62 REVIEW OF EDUCATIONAL RESEARCH 61 (1992)

[29]Elizabeth Hampson, Sari M. van Anders & Lucy I. Mullin, *A female advantage in the recognition of emotional facial expressions: test of an evolutionary hypothesis*, 27 EVOLUTION AND HUMAN BEHAVIOR 401-416 (2006); R. Buck et al., *Communication of affect through facial expressions in humans*, 23 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 362–371 (1972); H. L. Wagner, R. Buck & M. Winterbotham, *Communication of specific emotions: Gender differences in sending accuracy and communication measures*, 17 JOURNAL OF NONVERBAL BEHAVIOR 29-53 (1993); J. A. Hall, *Gender Effects in Decoding Nonverbal Cues.*, PSYCHOLOGICAL BULLETIN (1978); J. A. HALL, NONVERBAL SEX DIFFERENCES: COMMUNICATION ACCURACY AND EXPRESSIVE STYLE (1984); R. ROSENTHAL ET AL., SENSITIVITY TO NONVERBAL COMMUNICATION: THE PONS TEST (1979); S. Baron-Cohen et al., *The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism*, 42 THE JOURNAL OF CHILD

Women also show greater ability to identify socially inappropriate behavior.[30]

e) *Verbal ability.*

In the aggregate set of verbal tasks generally classified as "verbal intelligence", there is little if any difference between men and women.[31] However, men and women score differently on a variety of individual verbal tasks.

Women score significantly higher on tests that require writing samples[32] and slightly higher on tests of reading comprehension.[33] Women score higher than men on tests of grammar and spelling.[34] 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".[35] 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,[36] while men score higher for geometric categories.[37]

PSYCHOLOGY AND PSYCHIATRY AND ALLIED DISCIPLINES 241-251 (2001); S. Baron-Cohen et al., *Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger Syndrome*, 38 JOURNAL OF CHILD PSYCHOLOGY AND PSYCHIATRY 813-822 (1997); Erin B. McClure, *A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents.*, 126 PSYCHOLOGICAL BULLETIN 424-453 (2000).

[30] S. Baron-Cohen et al., *Recognition of Faux Pas by Normally Developing Children and Children with Asperger Syndrome or High-Functioning Autism*, 29 JOURNAL OF AUTISM AND DEVELOPMENTAL DISORDERS 407-418 (1999); F. G. Happe, *The role of age and verbal ability in the theory of mind task performance of subjects with autism.*, 66 CHILD DEV 843-55 (1995).

[31]J. S. Hyde & M. C. Linn, *Gender differences in verbal ability*, PSYCHOLOGICAL BULLETIN, 104 (1) 53-69 (1988).

[32] Hedges and Nowell, *supra* note___.

[33] *Id.*

[34]Hyde and Linn, *supra* note___*;* A. Feingold, *Cognitive sex differences are disappearing*, 43 AMERICAN PSYCHOLOGIST 95-103 (1988).

[35] DOREEN KIMURA, SEX AND COGNITION (1999).

[36]D. Kimura, *Body asymmetry and intellectual pattern.*, 17 PERSONALITY AND INDIVIDUAL DIFFERENCES 53-60 (1994).

[37] R. A. Harshman, E. Hampson & S. A. Berenbaum, *Individual differences in cognitive abilities and brain organization, Part I: Sex and handedness differences in ability.*, 37 CANADIAN JOURNAL OF PSYCHOLOGY 144-92 (1983).

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.[38]

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.[39]

f) *Perception.*

Women are on average better able than men to discriminate between tastes.[40] Women may also have a more discriminating sense of touch.[41] Men are able to see smaller objects within their visual field,[42] but women have larger visual fields.[43] Women appear to have better depth perception.[44] Women score higher on tasks of perceptual speed, which require subjects to match a given item with the identical item in an array of pictures.[45]

---

38 M. B. Denckla & R. Rudel, *Rapid" automatized" naming of pictured objects, colors, letters and numbers by normal children.*, 10 CORTEX 186-202 (1974); Doreen Kimura, D. Saucier, and R. Matuk, *Women name both colors and forms faster than men*, 22 SOCIETY FOR NEUROSCIENCE ABSTRACTS 1860 (1996).

39 D. McGuinness, A. Olson & J. Chapman, *Sex differences in incidental recall for words and pictures*, 2 LEARNING AND INDIVIDUAL DIFFERENCES 263-285 (1990); J. H. Kramer, D. C. Delis & M. H. Daniel, *Sex differences in verbal learning.*, 44 JOURNAL OF CLINICAL PSYCHOLOGY 907-915 (1988); M. L. Bleecker et al., *Age-related sex differences in verbal memory.*, 44 JOURNAL OF CLINICAL PSYCHOLOGY 403-11 (1988); K. Owen & R. Lynn, *Sex differences in primary cognitive abilities among blacks, Indians and whites in South Africa.*, 25 JOURNAL OF BIOSOCIAL SCIENCE 557-60 (1993); V. A. Mann et al., *Sex differences in cognitive abilities: a cross-cultural perspective.*, 28 NEUROPSYCHOLOGIA 1063-77 (1990).

40 W. Velle, *Sex differences in sensory functions.*, 30 PERSPECTIVE IN BIOLOGY AND MEDICINE 490-522 (1987).

41 L. Ghent, *Developmental changes in tactual thresholds on dominant and nondominant sides.*, 54 JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY 670-3 (1961).

42 A. Burg, *Visual acuity as measured by dynamic and static tests: a comparative evaluation.*, 50 JOURNAL OF APPLIED PSYCHOLOGY 460-6 (1966).

43 A. Burg, *Lateral visual field as related to age and sex.*, 52 JOURNAL OF APPLIED PSYCHOLOGY 10-5 (1968) .

44 Kimura, *supra* note___ at 34-37.

45 R.L. Majeres, *Sex differences in symbol-digit substitution and speeded matching*, 7 INTELLIGENCE 313-327 (1983).

g) *Emerging Research.*

Certain types of intelligence are very difficult to measure, but in recent years studies show promise in achieving such measurement. The results of new research in these areas must be viewed as preliminary. One study showed a female advantage in inductive reasoning.[46] Studies of creativity in general show no sex differences.[47] However, there is very tentative evidence suggesting that creativity might be positively correlated with other kinds of intelligence in females, and negatively correlated in males.[48]

h) *Time trends.*

Some (though not all) studies have found that test score gaps between men and women have been falling over the years.[49] These results are hard to interpret, however, because test writers have been deliberately restructuring tests to make them gender neutral by dropping questions on which one sex performs better or by adjusting the weights given to different types of questions so that any sex differences neutralize each other.[50]

B. NATURE AND NURTURE

Observed differential aptitude distributions between the sexes do not follow a neat or easily defined pattern. Disentangling the cultural and biological aspects of these observed differences is no simple matter.

---

[46] R. Colom & O. Garcia-López, *Sex differences in fluid intelligence among high school graduates*, 32 PERSONALITY AND INDIVIDUAL DIFFERENCES 445-451 (2002).

[47] J. Baer, *Gender differences*, 1 *in* ENCYCLOPEDIA OF CREATIVITY 753-758 (1999)(slight female advantage). M. P. Matud, C. Rodríguez & J. Grande, *Gender differences in creative thinking*, 43 PERSONALITY AND INDIVIDUAL DIFFERENCES 1137-1147 (2007)(slight male advantage on two of seven measures); E. E. MACCOBY & C. N. JACKLIN, THE PSYCHOLOGY OF SEX DIFFERENCES (1978).

[48] Matud, Rodríguez, and Grande, *supra* note___.

[49] Feingold, *supra* note___. But see Hedges and Nowell, *supra* note___.

[50] J. C. Stanley et al., *Gender differences on eighty-six nationally standardized aptitude and achievement tests*, *in* TALENT DEVELOPMENT: PROCEEDINGS FROM THE 41–48 (1991).

1. CHANCE

The simplest environmental theory of differential aptitude distributions would propose that observed differences result simply from random variations in sex roles between societies. These sex roles would in turn give rise to different socialization of and thus different observed aptitudes in men and women.

The degree of cross-cultural consistency in some observed differential aptitude distributions is hard to reconcile with the random environmental hypothesis. Males in a wide variety of cultures, including African, East Indian and Asian, score higher than females on tests of spatial rotation.[51] Females outscore males on tests of nonverbal communication in New Guinea, the U.S., Israel, and Australia. The male advantage in mathematical problem solving and female advantage in computation has been found in a number of different societies including the U.S., Thailand, Taiwan, and Japan.[52] On the other hand, the cross-cultural evidence is found less consistently for other skills. A male advantage in field dependent tasks has been confirmed in African societies but not in Inuit Eskimo groups.[53]

The assumption that sex roles vary widely in traditional societies is also not entirely consistent with the available evidence. The classic study of occupational segregation in preindustrial societies is George Murdock and Caterina Prevost's analysis of data from the Standard Cross Cultural Sample. Murdoch and Prevost calculated the extent to which various

[51] I. Silverman, J. Choi & M. Peters, *The Hunter-Gatherer Theory of Sex Differences in Spatial Abilities: Data from 40 Countries*, 36 ARCHIVES OF SEXUAL BEHAVIOR 261-268 (2007); J. T. Mayes, G. Jahoda & I. Neilson, *Patterns of visual-spatial performance and'spatial ability': dissociation of ethnic and sex differences.*, 79 BRITISH JOURNAL OF PSYCHOLOGY 105-19 (1988); Owen and Lynn, *supra* note___; Mann et al., *supra* note___; M. W. Steinkamp et al., *Cross-national gender differences in mathematics attitude and achievement among 13-year-olds*, 4 THE JOURNAL OF MATHEMATICAL BEHAVIOR 259–277 (1985).

[52] George Engelhard Jr., *Gender differences in performance on mathematics items: evidence from the United States and Thailand*, 15 CONTEMPORARY EDUCATIONAL PSYCHOLOGY 13-26 (1990); Penner, *supra* note___;M. Lummis & H. W. Stevenson, *Gender Differences in Beliefs and Achievement: A Cross-Cultural Study.*, 26 DEVELOPMENTAL PSYCHOLOGY 254-63 (1990); S.P. Marshall & J.D. Smith, *Sex differences in learning mathematics: a longitudinal study with item and error analyses*, 79 JOURNAL OF EDUCATIONAL PSYCHOLOGY 372-383 (1987).

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

activities were predominantly preformed by members of one sex or the other. Table II summarizes the extent to which various activities were performed by males:

Predominantly Masculine activities	**(%)**	**Quasi-masculine activities**	**(%)**	**Swing activities**	**(%)**	**Quasi-feminine activities**	**Index (%)**
Hunting large aquatic fauna	100	Butchering	92.3	Generation of fire	62.3	Fuel gathering	27.2
Smelting of ores	100	Collection of wild honey	91.7	Bodily mutilation	60.8	Preparation of drinks	22.2
Metalworking	99.8	Land clearance	90.5	Preparation of skins	54.6	Gathering of wild vegetal foods	19.7
Lumbering	99.4	Fishing	86.7	Gathering small land fauna	54.5	Dairy production	14.3
Hunting large land fauna	99.3	Tending large animals	82.4	Crop planting	54.4	Spinning	13.6
Work in wood	98.8	Housebuilding	77.4	Manufacture of leather products	53.2	Laundering	13.0
Fowling	98.3	Soil preparation	73.1	Harvesting	45.0	Water fetching	8.6
Making musical instruments	97.6	Netmaking	71.2	Crop tending	44.6	Cooking	8.3
Trapping	97.5	Making rope and cordage	69.9	Milking	43.8	Preparation of vegetal food	5.7
Boatbuilding	96.6			Basketmaking	42.5		
Stoneworking	95.9			Burden carrying	39.3		
Work in bone, horn, shell	94.6			Matmaking	37.6		
Mining and quarrying	93.7			Care of small animals	35.9		
Bonesetting	92.7			Preservation of meat or fish	32.9		
				Loom weaving	32.5		
				Gathering small aquatic fauna	31.1		
				Manufacture of clothing	22.4		
				Potterymaking	21.1		

Table II
AVERAGE PERCENTAGE OF MALE PARTICIPATION IN ACTIVITIES[54]

Some activities are clearly strongly associated with one sex, while others are not. Women dominate gathering and are relatively important in

[54] This table is taken from W. Wood & A. H. Eagly, *A Cross-Cultural Analysis of the Behavior of Women and Men: Implications for the Origins of Sex Differences*, 128 PSYCHOLOGICAL BULLETIN 699-727 (2002), who in turn synthesized Tables 1–5 in G. P. Murdock & C. Provost, *Factors in the division of labor by sex: A cross-cultural analysis*, 12 ETHNOLOGY 203-25 (1973).

so-called extensive agriculture which uses low levels of capital, materials and labor relative to land area. The male role increases in importance with (in ascending order) animal husbandry, intensive agriculture, fishing and hunting.[55] Occupational segregation can also be seen in other activities. In all preindustrial societies, men play little or no role in the care of young children.

The sexual division of labor is imperfectly responsive to the economic value of sex-linked tasks. The relative contribution of men and women to subsistence varies with the dominant mode of subsistence production. In pre-industrial societies, female contribution is especially high in Africa and the Pacific Islands, where gathering and extensive agriculture are productive relative to hunting. In societies in which female dominated activities are relatively productive, males do not respond by switching their productive activity to characteristically female types of production. Instead, they engage in high levels of male-male violence. In all societies, industrial and pre-industrial, young men engage in homicide, especially same-sex homicide, at vastly higher rates than women; the discrepancy diminishes dramatically with age.[56] This may be part of a general pattern in which young males are willing to engage in risky behavior, perhaps because of a short time horizon.[57]

To reconcile these patterns with a purely cultural theory of differential aptitude distributions, Anne Fausto-Sterling suggests that chance determined sex roles in a single prehistoric society from which all modern societies have descended.[58] Such a scenario seems implausible. Tracing the movement of early humans within Africa is difficult, but to explain the observed patterns within Africa 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 it seems highly unlikely that a purely chance variation on such an important matter would have survived so robustly.

---

[55] Alice Schlegel & H.B. Barry III, *The Cultural Consequences of Female Contribution to Subsistence*, 88 AMERICAN ANTHROPOLOGIST 142-150 (1986).

[56] Martin Daly & Margo Wilson, *Crime and Conflict: Homicide in Evolutionary Psychological Perspective*, 22 CRIME AND JUSTICE: A REVIEW OF RESEARCH 51 (1997); Martin Daly & Margo Wilson, *Killing the competition*, 1 HUMAN NATURE 81-107 (1990).

[57] Martin Daly & Margo Wilson, *Evolutionary psychology of male violence*, *in* MALE VIOLENCE 253-288 (John Archer ed., 1994).

[58] A. FAUSTO-STERLING, MYTHS OF GENDER: BIOLOGICAL THEORIES ABOUT WOMEN AND MEN (1992).

2. FUNCTIONAL EXPLANATIONS: REPRODUCTIVE ROLE THEORIES

Perhaps the most plausible environmentalist account of differential aptitude distributions acknowledges the role of sex differences in another sphere, reproductive roles. One example of such a theory might run as follows. As in all mammalian species, female humans are inevitably more involved in the raising of young in simple societies. This fact constrains the ability of females to engage in certain kinds of productive activities, such as hunting, and makes other kinds, such as gathering, more attractive. Each sex was therefore socialized to excel in the skills that were most useful in the subsistence activity to which it was most suited. Freedom from reproductive constraints occurred quite late in human history, and so the cultural forces that encouraged the development of different aptitudes in each sex have not yet disappeared.

This modified environmental theory has much in common with the prevailing explanation of why sex differences in aptitude distributions might be biologically based.[59] Like the reproductive role environmental theory, evolutionary psychologists assign a crucial role to the reproductive division of labor between the sexes. But rather than arguing that different reproductive roles created different cultural roles, they believe that reproductive differences created different selection pressures for males and females. In particular, females developed aptitudes that were useful in pursuits that were complementary to child-raising while males developed other skills without the constraint of child care. In addition, however, evolutionary psychology proposes that other forces created differential aptitudes. Perhaps the most important such force is male competition for female reproductive resources.

The relative strength of men and women in some tasks does seem to correspond to the hunter-gatherer scenario that lies behind both the evolutionary psychology account and reproductive role environmentalism. Certainly the male advantage in targeting ability seems related to male specialization in hunting and the female advantage in fine motor skills to female specialization in gathering. Plausibly but less self-evidently, the hunter-gatherer hypothesis explains sex differences in navigational strategies. The male use of absolute locational strategies might more appropriate for returning from long and unpredictable journeys from home in which the return route is unimportant. The female use of landmark strategies might be more appropriate for repeated navigation in a defined

---

[59] Eals and Silverman, *supra* note___.

area where it is important to remember locations that are especially rich in foraging opportunities.

The hunter-gatherer hypothesis has been proposed for various other observed differentiated aptitude distributions, but seems somewhat more of stretch. Women, it has been argued, need greater interpersonal skills because of their child rearing responsibility.[60] 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. For example, male reproductive success depends in part on the ability to perceive how receptive a female is to sexual overtures.

A further limitation of functional explanations is that cognitive sex differences do not display a simple pattern, such as verbal versus mathematical, or emotional versus logical. In each category, 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?

Thus, neither environmental nor biological explanations seem to explain many sex differences in aptitude distributions through functional explanations based on sex roles, such as female mothering and male aggression. One potential explanation for these patterns might be the existence of tradeoffs between levels of various skills. Perhaps not every skill difference has a functional explanation - some might be merely by-products of mechanisms that produce other skill differences that are functional. Some evidence supports this possibility.[61]

The most plausible account of sex differences, then, combines an element of functional explanation based on reproductive differences with constraints that require tradeoffs between different specialized types of skills. That these constraints appear to be physiological and genetic in nature sheds little light on whether cognitive sex differences are genetic as

---

[60] W. A. Babchuk, R. B. Hames & R. A. Thompson, *Sex differences in the recognition of infant facial expressions of emotion: The primary caretaker hypothesis*, 6 ETHOLOGY & SOCIOBIOLOGY 89-101 (1985).

[61]W. Johnson & T. J. Bouchard, *Sex differences in mental ability: A proposed means to link them to brain structure and function*, 35 INTELLIGENCE 197-209 (2007); John R. Kirby, *Collaborative and competitive effects of verbal and spatial processes*, 3 LEARNING AND INSTRUCTION 201-214; S. Wheelwright et al., *Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ)*, 1079 BRAIN RESEARCH 47-56 (2006).

well. Different environmental pressures acting on the two sexes might simply lead to different optimal points given an identical set of tradeoffs.

3. Evidence of Nature and Nurture

Several different types of evidence can illuminate the role of environment and biology in the odd assortment of sex-differentiated aptitudes. This evidence supports a role for both biology and environment.

a) *Stereotype threat.*

A strong piece of evidence for environmental factors comes from 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 and men taking tests on sensitivity.[62]

b) *Cross-cultural comparisons.*

As noted earlier, the relative advantages of both sexes appear cross-culturally. However, the magnitude of these differences varies, and may be correlated with cultural attitudes that support or undermine women's participation in mathematics.[63] These patterns suggest a role for both environment and genetics.

---

[62] S. J. Spencer, C. M. Steele & D. M. Quinn, *Stereotype Threat and Women's Math Performance*, 35 JOURNAL OF EXPERIMENTAL SOCIAL PSYCHOLOGY 4-28 (1999); A. M. Koenig & A. H. Eagly, *Stereotype Threat in Men on a Test of Social Sensitivity*, 52 SEX ROLES 489-496 (2005). Concerns about stereotype threat are sometimes offered as a reason for caution in discussing evidence that some sex differences may have a genetic basis, and experimental evidence shows some support for this concern. I. Dar-Nimrod & S. J. Heine, *Exposure to scientific theories affects women's math performance.*, 314 SCIENCE 435 (2006). However, other evidence suggests that stereotype threat may reduced or eliminated without restraining free discussion of an important scientific and policy matter simply by informing subjects that stereotype threat might interfere with their performance. Michael Johns, Toni Schmader & Andy Martens, *Knowing Is Half the Battle*, 16 PSYCHOLOGICAL SCIENCE 175-179 (2005).

[63] Delwyn L. Harnisch et al., *Cross-national differences in mathematics attitude and achievement among seventeen-year-olds*, 6 INTERNATIONAL JOURNAL OF EDUCATIONAL DEVELOPMENT 233-244 (1986).

c) *Age.*

Some differential aptitude distributions appear at relatively early ages. Differences in targeting skill are evident in three to five year olds.[64] Newborn girls show more interest in animate objects while newborn boys show more interest in mechanical objects.[65] The female advantage at interpreting other people's emotions appears to emerge in infancy.[66] Early emergence tends to support biological theories of difference. However, the male advantage in math problem solving does not appear until puberty. This timing is consistent either with a socialization hypothesis or with a biological effect triggered by pubertal hormone changes.

d) *Experience and Training.*

One environmental explanation for sex differences might be that social norms provide one sex with more experience in such tasks. However, 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.[67] Other evidence comes from examining the effects of additional training on performance. Training presumably has a declining marginal value, and so the less experienced sex should gain more from additional training. In spatial tasks, studies suggest that both sexes benefit equally from additional training.[68]

e) *Brain anatomy.*

The male and female brains have many anatomical differences. Such differences are not sufficient to demonstrate cognitive differences. Some anatomical sex differences apparently prevent behavioral differences by compensating for other anatomical differences.[69] This

---

[64] S. C. Levine et al., *Early Sex Differences in Spatial Skill.*, 35 DEVELOPMENTAL PSYCHOLOGY 940-49 (1999); Kimura, *supra* note___ at 35;

[65] Jennifer Connellan et al., *Sex differences in human neonatal social perception*, 23 INFANT BEHAVIOR AND DEVELOPMENT 113-118 (2000).

[66] McClure, *supra* note___.

[67] Watson and Kimura, *supra* note___ (finding one standard deviation difference between men and women in dart throwing task).

[68] M. Baenninger & N. Newcombe, *The role of experience in spatial test performance: A meta-analysis*, 20 SEX ROLES 327-344 (1989).

[69] G. J. De Vries, *Minireview: Sex Differences in Adult and Developing Brains: Compensation, Compensation, Compensation,* 145 ENDOCRINOLOGY 1063-1068 (2004)(survey); M. Piefke et al., *Gender differences in the functional neuroanatomy of*

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[70] and different distributions of various types of brain tissue.[71] Conversely, some sex differences in performance are not associated with sexually dimorphic brain structures.[72]

However, some sexually dimorphic features of the brain do in fact seem to be associated with sex differences in performance. In both sexes, the left hemisphere specializes in language processing and the right in spatial and mathematical activities.[73] This lateralization is more pronounced in males than in females, who tend to use both hemispheres for different tasks.[74] Though the issue is still open, many researchers

---

emotional episodic autobiographical memory, 24 HUMAN BRAIN MAPPING 313-324 (2005) (retrieval of emotionally charged personal memories accomplished equally well by male and female subjects, , but different regions of brain activated). T. J. Grabowski et al., *Effects of gender on blood flow correlates of naming concrete entities*, 20 NEUROIMAGE 940-954 (2003)(image naming task; men and women performed equally well but displayed different patterns of brain activity).

[70] Even adjusting for body size the female brain is smaller than the male, C. Davison Ankney, *Sex differences in relative brain size: the mismeasure of woman, too?*, 16 INTELLIGENCE 329-336 (1992), but female brains appear to have more folds and fissures, creating more surface area that may compensate for smaller size. E. Luders et al., *Gender differences in cortical complexity*, 7 NATURE NEUROSCIENCE 799-800 (2004)

[71] Gray matter consists primarily of nerve cell bodies, dendrites and small axons White matter is composed of larger axons that transport information away from nerve bodies and the myelin sheath of those axons. Females have higher percentages of gray matter; males have higher percentage of white matter and cerebrospinal fluid. In males the percentage of gray matter is higher in the right hemisphere; the percentage of white matter symmetric; and the amount of cerebrospinal fluid higher in the left. Women show no asymmetries. 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); R. J. Haier et al., *The neuroanatomy of general intelligence: sex matters*, 25 NEUROIMAGE 320-327 (2005)

[72] E. C. Bell et al., *Males and females differ in brain activation during cognitive tasks*, 30 NEUROIMAGE 529-538 (2006).

[73] N. GESCHWIND & A. M. GALABURDA, CEREBRAL LATERALIZATION: BIOLOGICAL MECHANISMS, ASSOCIATIONS, AND PATHOLOGY (1987); D. Kimura, *Functional asymmetry of the brain in dichotic listening*, 3 CORTEX 163-178 (1967). Some tasks, such as mental rotation, show more lateralization than others. S. L. Rilea, B. Roskos-Ewoldsen & D. Boles, *Sex differences in spatial ability: a lateralization of function approach.*, 56 BRAIN COGN 332-43 (2004); J. J. Vogel, C. A. Bowers & D. S. Vogel, *Cerebral lateralization of spatial abilities: A meta-analysis*, 52 BRAIN AND COGNITION 197-204 (2003).

[74]Geschwind and Galaburda, *supra* note___ J. McGlone, *Sex differences in human brain asymmetry: A critical survey*, 3 BEHAVIORAL AND BRAIN SCIENCES 215-263

believe that differences in lateralization may produce some of the cognitive differences between men and women.[75]

Even if lateralization is shown to be a cause of cognitive sex differences, it might result from both social and environmental factors. Studies have clearly shown that environmental factors can affect the physiology of the brain. However, some evidence, discussed below, does tentatively point to a biological basis for at least part of sexually dimorphic lateralization.

f) *Genetic influences*

In many species, including all mammals, most genetic influences on brain sex differences arise indirectly through hormonal secretions. 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 newly-formed testes begin to secrete testosterone that sex-differentiate many cells, including many in the brain. The effects of these hormones on neural and cognitive sexual differences will be discussed in the next section.

For many years, these internally produced (endogenous) hormones were believed to be the sole source of genetic sex differentiation. Recent research has shown that other genes also contribute to sexual dimorphism in the brain. This research, however, is relatively new, and has been conducted almost exclusively in non-human mammals, and thus the precise nature of other genes' contribution to sex differences in the brain and in cognition is not yet well understood. [76]

(1980). N. Jausovec & K. Jausovec, *Spatial rotation and recognizing emotions: Gender related differences in brain activity*, 36 INTELLIGENCE 383-393 (2008); A. C. Neubauer et al., *Intelligence and neural efficiency: Further evidence of the influence of task content and sex on the brain–IQ relationship*, 25 COGNITIVE BRAIN RESEARCH 217-225 (2005).

One intriguing aspect of lateralization involves creative thinking. Men and women display very different patterns of brain activation during creative thinking O. M. Razumnikova, *Gender differences in hemispheric organization during divergent thinking: an EEG investigation in human subjects*, 362 NEUROSCIENCE LETTERS 193-195 (2004) 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.

[75] Geschwind and Galaburda, *supra* note___.

[76] A. P. Arnold, *Sex chromosomes and brain gender*, 5 NATURE REVIEWS NEUROSCIENCE 701-708 (2004).

g) *Hormonal influences*

(i) *Organizational effects*.

Hormones can have organizational effects on the development of structures in the fetus, infant, or child. These effects persist regardless of subsequent hormone levels.

On the whole, evidence of organizational effects of hormones tends to support genetic explanations for sexual differentiation, since genetically determined levels of endogenously produced hormones are the principal source of hormone exposure. One type of organizational effect occurs during the prenatal testosterone surge. Other organizational effects occur later, especially in the first year of life and at puberty. However, hormone exposure can also result from environmental factors such as stress and, for a fetus, maternal hormone levels.

The organizational effects of hormones on cognitive performance can be studied directly by examining associations between prenatal hormone levels and cognitive performance. Studies on animals provide strong support for the hypothesis that hormones have organizational effect on cognitive behavior.[77] Research on humans is less conclusive, but points to at least a few links. Studies have clearly established that prenatal testosterone affects various non-cognitive sex-related behaviors such as play behavior and sexual orientation.[78] Evidence of a connection to cognitive behaviors is weaker. Relatively strong evidence suggests that prenatal testosterone exposure affects targeting ability.[79] Some research suggests that prenatal exposure to testosterone reduces interpersonal

---

[77] C. L. Williams & W. H. Meck, *The organizational effects of gonadal steroids on sexually dimorphic spatial ability*, 16 PSYCHONEUROENDOCRINOLOGY 155-176 (1991).

[78]See, e.g. M. Hines, *Prenatal testosterone and gender-related behaviour*, 155 EUROPEAN JOURNAL OF ENDOCRINOLOGY 115-121 (2006);

[79]M. Hines et al., *Spatial abilities following prenatal androgen abnormality: targeting and mental rotations performance in individuals with congenital adrenal hyperplasia*, 28 PSYCHONEUROENDOCRINOLOGY 1010-1026 (2003); C. M. Falter, M. Arroyo & G. J. Davis, *Testosterone: Activation or organization of spatial cognition?*, 73 BIOLOGICAL PSYCHOLOGY 132-140 (2006).

skills[80] and verbal ability.[81] Studies on mental rotation yield mixed results.[82]

The organizational effects of hormones on cognitive performance can also be studied indirectly, by examining the effects of prenatal hormones on dimorphic brain structures and then the association between those brain structures and performance. Again, animal studies provide compelling evidence of the organizational effects of hormones on brain structure,[83] while studies on humans provide weaker evidence.[84] The effect of at certain aspects of brain structure on cognitive ability is somewhat clearer.[85]

---

[80] S. Lutchmaya, S. Baron-Cohen & P. Raggatt, *Foetal testosterone and eye contact in 12-month-old human infants*, 25 INFANT BEHAVIOR AND DEVELOPMENT 327-335 (2002).

[81] S. Lutchmaya, S. Baron-Cohen & P. Raggatt, *Foetal testosterone and vocabulary size in 18-and 24-month-old infants*, 24 INFANT BEHAVIOR AND DEVELOPMENT 418-424 (2001).

[82] Studies finding no effect include J. Coolican & M. Peters, *Sexual dimorphism in the 2D/4D ratio and its relation to mental rotation performance*, 24 EVOLUTION AND HUMAN BEHAVIOR 179-183 (2003); E. J. Austin et al., *A preliminary investigation of the associations between personality, cognitive ability and digit ratio*, 33 PERSONALITY AND INDIVIDUAL DIFFERENCES 1115-1124 (2002). Studies finding an effect include Falter, Arroyo, and Davis, *supra* note___ (also finding effect on figure disembedding task); J. T. Manning & R. 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); G. M. Grimshaw, G. Sitarenios & J. A. K. Finegan, *Mental Rotation at 7 Years-Relations with Prenatal Testosterone Levels and Spatial Play Experiences*, 29 BRAIN AND COGNITION 85-100 (1995) .Some of studies suggest that high levels of prenatal testosterone raise spatial ability in females but lower it in males, consistent with the pattern that has arguably emerged in activation studies. *Id.* Hines and coauthors found no effect of prenatal testosterone on female spatial performance and concluded their study showed no effect of testosterone. However, it also found a negative correlation in males, consistent with the inverse U hypothesis. Hines et al., *supra* note___

[83] Bruce S. McEwen, *Sexual differentiation of the brain*, 291 NATURE 610 (1981); R. W. GOY & B. S. MCEWEN, SEXUAL DIFFERENTIATION OF THE BRAIN (1980).

[84]Sandra F. Witelson, *Neural sexual mosaicism: Sexual differentiation of the human temporo-parietal region for functional asymmetry*, 16 PSYCHONEUROENDOCRINOLOGY 131-153 (1991); C. C. C. Cohen-Bendahan et al., *Prenatal exposure to testosterone and functional cerebral lateralization: a study in same-sex and opposite-sex twin girls*, 29 PSYCHONEUROENDOCRINOLOGY 911-916 (2004).

[85]See, e.g., R. I. Siegel-Hinson & W. F. McKeever, *Hemispheric specialisation, spatial activity experience, and sex differences on tests of mental rotation ability*, 7 LATERALITY: ASYMMETRIES OF BODY, BRAIN, AND COGNITION 59-74 (2002); Geschwind and Galaburda, *supra* note___; Vogel, Bowers, and Vogel, *supra* note___.

(ii) *Activational effects*. 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,[86] but the results of studies in humans are inconclusive.[87] To the extent that studies find a relationship, they tend to suggest that the optimal level of testosterone is in the low-normal male range, so that increasing

---

[86] See, e.g., Bradley M. Cooke, Golnaz Tabibnia & S. Marc Breedlove, *A Brain Sexual Dimorphism Controlled by Adult Circulating Androgens*, 96 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 7538-7540 (1999)

[87] 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 R. Halari 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)(Finding no between-subject effect of hormone levels. Normal subjects, no intelligence controls, no direct measure of free T); L. 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)(medical administration of steroids for delayed puberty; between-subjects analysis using within-subject benchmark scores as controls; no effect); Falter, Arroyo, and Davis, *supra* note___)(no effect; between subjects, no control of general intelligence) with A. Aleman et al., *A single administration of testosterone improves visuospatial ability in young women*, 29 PSYCHONEUROENDOCRINOLOGY 612-617 (2004); A. Postma et al., *Sex differences and menstrual cycle effects in human spatial memory*, 24 PSYCHONEUROENDOCRINOLOGY 175-192 (1999)(exogenous testosterone correlated to normal female spatial performance);M. Hausmann et al., *Sex Hormones Affect Spatial Abilities During the Menstrual Cycle*, 114 BEHAVIORAL NEUROSCIENCE 1245-1250 (2000);. S. D. Moffat et al., *Longitudinal assessment of serum free testosterone concentration predicts memory performance and cognitive status in elderly men.*, 87 JOURNAL OF CLINICAL ENDOCRINOLOGY AND METABOLISM 5001-5007 (2002) (age related testosterone decline reduces spatial performance in males); I. Driscoll et al., *Virtual navigation in humans: the impact of age, sex, and hormones on place learning*, 47 HORMONES AND BEHAVIOR 326-335 (2005) (age related testosterone decline reduces spatial performance in males). One study that found a link also found that testosterone affected not mental rotation skill per se but an indirect contributor to rotational skill. C. K. Hooven et al., *The relationship of male testosterone to components of mental rotation*, 42 NEUROPSYCHOLOGIA 782-790 (2004).

testosterone in women should always increase spatial functioning, while in men it will have the reverse effect above a threshold level.[88]

If hormones turn out to have activational effects in humans, the implications for the nature-nurture debate will be complex. On the one hand, circulating hormone levels are determined primarily by each individual's genetic makeup. On the other hand, these levels can also be influenced by environmental factors.

h) *Sexual Orientation*

Men and women differ in the distribution of their cognitive ability; their hormone levels; and their neuroanatomy. Many studies have shown that homosexuals of both sexes often display some characteristics intermediate between their own sex and the opposite sex.[89] The environmental explanation for these observations would necessarily be based on the view that homosexuals are socialized to have characteristics of the opposite sex. Although this was once a common theory, many homosexuals have testified that, on the contrary, they were under great pressure to conform to the roles associated with their own sex. If this testimony is accurate, then their cognitive and anatomical characteristics

---

[88] S. D. Moffat & E. Hampson, *A curvilinear relationship between testosterone and spatial cognition in humans: Possible influence of hand preference*, 21 PSYCHONEUROENDOCRINOLOGY 323-337 (1996)(normal subjects); C. Gouchie & D. Kimura, *The relationship between testosterone levels and cognitive ability patterns*, 16 PSYCHONEUROENDOCRINOLOGY 323-334 (1991)(normal subjects);V. J. Shute et al., *The relationship between androgen levels and human spatial abilities.*, 21 BULLETIN OF THE PSYCHONOMIC SOCIETY 465-468 (1983).

[89] M. Peters, J. T. Manning & S. Reimers, *The Effects of Sex, Sexual Orientation, and Digit Ratio (2D: 4D) on Mental Rotation Performance*, 36 ARCHIVES OF SEXUAL BEHAVIOR 251-260 (2007);J. A. Y. Hall & D. Kimura, *Sexual orientation and performance on sexually dimorphic motor tasks*, 24 ARCHIVES OF SEXUAL BEHAVIOR 395-407 (1995); Qazi Rahman & Glenn D. Wilson, *Born gay? The psychobiology of human sexual orientation*, 34 PERSONALITY AND INDIVIDUAL DIFFERENCES 1337-1382 (2003);Qazi Rahman, Glenn D. Wilson & S. Abrahams, *Biosocial factors, sexual orientation and neurocognitive functioning*, 29 PSYCHONEUROENDOCRINOLOGY 867-881 (2004); Qazi Rahman & Glenn D. Wilson, *Sexual orientation and the 2nd to 4th finger length ratio: evidence for organising effects of sex hormones or developmental instability?*, 28 PSYCHONEUROENDOCRINOLOGY 288-303 (2003); Qazi Rahman, Glenn D. Wilson & S. Abrams, *Sexual orientation related differences in spatial memory*, 9 JOURNAL OF THE INTERNATIONAL NEUROPSYCHOLOGICAL SOCIETY 376-383 (2003). A partial exception to this pattern is the minority of homosexual men who display hypermasculine characteristics. Homosexual men are thus more bimodal in their sex-linked traits than heterosexual men or women of any sexual orientation. Rahman and Wilson, *supra* note___.

developed in spite of, not because of, social pressures. A primarily genetic view of homosexuality is hard to reconcile with a purely environmental account of sex differences.

4. IMPLICATIONS

The scientific study of sex differences is still at a relatively early stage. The current evidence does provide a moderately strong case that the distribution of a few aptitudes, especially targeting ability, has a biological basis for sex differences. For many other observed differences, such as mathematics, the evidence suggests some biologically based difference, but is far from conclusive about the size of that difference. Some observers argue that that magnitude is relatively small and cannot account for the observed magnitudes of occupational segregation.[90]

This reasoning, however, is mistaken. Even small statistical differences might produce large degrees of occupational segregation. Suppose that there are two economically valuable skills, A and B. Suppose that skill at one task is highly correlated to skill at the other. In other words, general abilities --intelligence, agility, and so on-- are the main predictors of success at both. Suppose further that the distribution of general abilities shows enormous intrasex variation and no intersex differences.

However, suppose that there are also small task-specific aptitudes 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 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.

This scenario will tend to produce significant occupational segregation even though the statistical differences between the sexes are small. The exact degree of segregation will depend on factors such as the relative value of A and B. If the products of both activities are perfect substitutes for each other, a competitive market with full information would produce complete occupational segregation even though the variation within sexes is huge and the variation between sexes small.

---

[90] MELISSA HINES, BRAIN GENDER (2004).

III. DISCRIMINATION

A. THE PARADOX OF STATISTICAL DISCRIMINATION

1. THE BENCHMARK MODEL

Some observers believe that the existing scientific evidence suggests large statistical sex differences for some skills. From this they infer that discrimination plays little if any role in the occupational segregation of fields that make use of those skills.[91]

But just as small aptitude differences do not imply a small degree of occupational segregation, large differences do not imply the absence of discrimination. This inference might be reasonable if employers had near-perfect information on candidates and employees. In fact, no real-life employer has complete information on the qualifications of any given job candidate. 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 aptitude of a particular individual.

The basic mechanism of statistical discrimination was described by Edmund Phelps and subsequently expanded by Dennis Aigner and Glen Cain.[92] In this benchmark model, employers care about the quality of the workers that they hire, but quality varies among workers and employers cannot observe it directly. Instead, they observe various types of evidence. This can be any easily observed objective measure such as a test score, educational qualifications and so on. For brevity this evidence is usually referred to as a test score. The value of the test score predicts an applicant's quality with error.[93] The employer's best estimate of an applicant's aptitude is a weighted average of the mean quality for all

---

91 KINGSLEY R. BROWNE, BIOLOGY AT WORK: RETHINKING SEXUAL EQUALITY (2002).

92 The terminology in text is taken from Edmund S. Phelps, *The Statistical Theory of Racism and Sexism*, 62 AMERICAN ECONOMIC REVIEW 659-661 (1972) and Dennis J. Aigner & Glen G. Cain, *Statistical Theories of Discrimination in Labor Markets*, 30 INDUSTRIAL AND LABOR RELATIONS REVIEW 175-187 (1977). Both Phelps and Cain and Aigner also address the issues that arise when the variance of ability or the reliability of the test differs between the two groups.

93 Specifically

$$y = q + u$$

where q is the quality of an individual, y is the test score, and u is the error term.

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 quality. The more predictive the test, the more weight it receives relative to the group average.[94]

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 quality. The employer thus estimates the qualifications of each individual male and female by examining the weighted average of group mean and individual score.[95] Since the group average for women is lower than the group average for men, the employer will estimate a lower productivity for a female candidate with a given score than for a male candidate with the same score. The greater the difference, the more discrimination will occur. Evidence of sex differences does not prove that discrimination plays no role in occupational segregation. Rather, it almost inevitably implies that employers have a powerful incentive to engage in at least some degree of statistical discrimination.

2. THE DYNAMICS OF STATISTICAL DISCRIMINATION

A longstanding debate in the discrimination literature examines 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 are effectively maximizing profits.

The simple benchmark model of statistical discrimination does not evolve over time. The benchmark model assumed, implicitly or explicitly, a long set of assumptions, including fixed human capital; fixed employer

---

[94] The employer estimate of an individual applicant's quality is

$$= (1- \gamma)\ \alpha + \gamma y$$

where
α=mean quality for all applicants,
$\gamma= r^2$=squared coefficient of correlation between y and q

[95]

$$^{f}= (1- \gamma)\ \alpha^{f} + \gamma y$$
$$^{m}= (1- \gamma)\ \alpha^{m} + \gamma y$$

where
α^f =mean value of quality q for females
α^m = mean value of quality for males

information, and fixed labor supply.[96] 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, such as females, leads to a lower wage for all members of the group, regardless of their individual ability or test scores. The model has no dynamic tendency to move to a new equilibrium, but it is also perfectly efficient in its initial one.

The dynamic analysis of statistical discrimination drops the benchmark assumption that human capital levels are fixed, and instead assumes that workers chose their human capital investments based on their future job market prospects.[97] Efficiency in this setting generally requires that higher ability people acquire more human capital.

With acquired human capital, statistical discrimination can occur even if both groups have identical underlying ability distributions. Although underlying ability distributions are identical, the distribution of worker output will depend on human capital as well as ability, and may therefore differ between two groups. The eventual market equilibrium

---

[96] Aigner and Cain, *supra* note___; Phelps, *supra* note___. For an illuminating discussion of the role of these assumptions, see Joseph E. Stiglitz, *The Theory of "Screening," Education, and the Distribution of Income*, 65 THE AMERICAN ECONOMIC REVIEW 283-300 (1975). From the beginning, the literature has also analyzed the consequences of different variances in the signal quality between the two groups. Phelps, *supra* note___ Aigner and Cain, *supra* note___. Differing variances can produce statistical discrimination even in the absence of any difference in the underlying ability distributions. This phenomenon may be of great interest in understanding occupational segregation, but the data in this article does not address this possibility.

[97]The original paper pointing out the role of human capital acquisition is Kenneth J. Arrow, *The Theory of Discrimination*, *in* DISCRIMINATION IN LABOR MARKETS 3–33 (Orley Ashenfelter & Albert Rees eds., 1973); See also George Akerlof, *The Economics of Caste and of the Rat Race and Other Woeful Tales*, 90 QUARTERLY JOURNAL OF ECONOMICS 599-617 (1976). This observation was incorporated into the Phelps-Aigner-Cain approach, Phelps, *supra* note___, Aigner and Cain, *supra* note___ by S. J. Lundberg & R. Startz, *Private Discrimination and Social Intervention in Competitive Labor Market*, 73 AMERICAN ECONOMIC REVIEW 340-347 (1983), which like Arrow, concluded that human capital effects tended to make statistical discrimination inefficient. 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) Andrea Moro & Peter Norman, *Affirmative action in a competitive economy*, 87 JOURNAL OF PUBLIC ECONOMICS 567-594 (2003); Andrea Moro & Peter Norman, *A general equilibrium model of statistical discrimination*, 114 JOURNAL OF ECONOMIC THEORY 1-30 (2004); Peter Norman, *Statistical Discrimination and Efficiency*, 70 REVIEW OF ECONOMIC STUDIES 615-627 (2003) although that paper itself addressed only affirmative action rather than prohibitions on statistical discrimination..

depends on the human capital distribution when the market starts. If members of both groups start with efficient levels of human capital, future members of both groups will continue acquire efficient levels of human capital, 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 lower human capital levels. Any initial deviation from efficient skill levels perpetuates itself: even employers without any animus correctly observe that members of the disadvantaged group have less skill and the employers therefore statistically discriminate against them. This in turn discourages members of the group from investing in human capital, perpetuating the true stereotype about skill levels and disguising the actual distribution of ability. This result may hold even when employers use sophisticated strategies for updating their assessment of ability distributions.[98] These equilibria will generally be stable, and show no tendency to converge to an efficient equilibrium.

Anti-discrimination policy can sometimes move the economy to the non-discriminatory equilibrium.[99] As these policies change incentives for human capital accumulation, differences diminish, and eventually the optimal stable equilibrium is reached and there is no further need for anti-discrimination rules.

The assumption of identical ability distributions is a helpful way of thinking about discrimination in areas such as race where current skill levels diverge between groups because of historical forces rather than because of underlying abilities. It is not, however, a complete way of thinking about sex. To the extent that differences are not changeable over time, the paradox has more far reaching implications. Some output differences will be attributable to differences in human capital, and anti-discrimination policy may reduce these differences by changing the incentives of both groups to accumulate human capital. Yet even if antidiscrimination policy moves the labor market towards an efficient

---

[98] A. Farmer & D. Terrell, *Discrimination, Bayesian Updating of Employer Beliefs and Human Capital Accumulation.*, 34 ECONOMIC INQUIRY 204-19 (1996); George J Mailath, Larry Samuelson & Avner Shaked, *Endogenous Inequality in Integrated Labor Markets with Two-Sided Search*, 90 AMERICAN ECONOMIC REVIEW 46-72 (2000); Lawrence E Blume, *Learning and Statistical Discrimination*, 95 AMERICAN ECONOMIC REVIEW 118-121 (2005)

[99] Lundberg and Startz, *supra* note___; S. J. Lundberg, *The Enforcement of Equal Opportunity Laws under Imperfect Information: Affirmative Action and Alternatives.*, 106 QUARTERLY JOURNAL OF BUSINESS AND ECONOMICS (1991);Coate and Loury, *supra* note___.

human capital equilibrium, that equilibrium never becomes self-sustaining. Since underlying differences never disappear, employers have a continued incentive to statistically discriminate. At best, policy can permanently eliminate certain historically contingent elements of discrimination, leaving an equilibrium with the statistical discrimination associated with the underlying difference in ability distributions. With underlying differences, the paradox is an immutable fact in the labor market. Perpetual difference implies perpetual discrimination.

B. 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 skills 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.

1. THE EXISTENCE OF DISCRIMINATION

a) *Studies of Labor Market Data*

Studies of labor market data by economists tend to confirm the general hypothesis of job discrimination. Economists traditionally measured discrimination by estimating wage regressions on observable proxies for productivity such as education. Residual discrimination is then measured in various ways, such as unexplained residual or the coefficients on dummy variables for race or sex.[100] Such studies tend to support the view that discrimination, though declining, continues to contribute to the wage gap.[101]

---

[100] This approach was first developed in R. 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).

[101] Stephen B. Jarrell & T. D. Stanley, *Declining Bias and Gender Wage Discrimination? A Meta-Regression Analysis*, 39 JOURNAL OF HUMAN RESOURCES 828 (2004).

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,[102] with the balance apparently attributable to discrimination.[103] However, if occupational segregation itself is in part a result of discrimination, these studies 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.[104] 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.[105]

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 fifteen years after graduation for graduates of the University of Michigan Law School.[106] 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

---

[102] Kimberly Bayard et al., *New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data*, 21 JOURNAL OF LABOR ECONOMICS 887-922 (2003); D. A. Macpherson & B. T. Hirsch, *Wages and Gender Composition: Why do Women's Jobs Pay Less?*, 13 JOURNAL OF LABOR ECONOMICS 426 (1995); J. Fields & E. N. Wolff, *Interindustry Wage Differentials and the Gender Wage Gap*, 49 INDUSTRIAL AND LABOR RELATIONS REVIEW 105-20 (1995).

[103] Bayard et al., *supra* note___.

[104] Claudia Goldin & Cecilia Rouse, *Orchestrating Impartiality: The Impact of Blind Auditions on the Sex Composition of Orchestras*, 90 AMERICAN ECONOMIC REVIEW 715–741 (2000).

[105]Judith K. Hellerstein, David Neumark & K. R. Troske, *Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations*, 17 JOURNAL OF LABOR ECONOMICS 409-446 (1999).

[106] Robert G. Wood, Mary E. Corcoran & Paul N. Courant, *Pay Differences among the Highly Paid: The Male-Female Earnings Gap in Lawyers' Salaries*, 11 JOURNAL OF LABOR ECONOMICS 417 (1993).

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.[107]

One 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 attempt to match less tangible qualities as well. The study found that females were less likely to be hired than comparable males.[108]

b) *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 "job candidates" are actors, and sometimes the subject receives only information about the candidate, such as a resume. Many different employment-related decisions have been

---

[107] Catherine J. Weinberger, *Race and Gender Wage Gaps in the Market for Recent College Graduates*, 37 JOURNAL OF ECONOMY & SOCIETY 67-84 (1998).

[108] David Neumark, Roy J. Bank & Kyle D. Van Nort, *Sex Discrimination in Restaurant Hiring: An Audit Study*, 111 QUARTERLY JOURNAL OF ECONOMICS 915 (1996). This appears to be the only audit study focusing on sex discrimination in hiring. Audit studies have been often used to study race discrimination, see, e.g. MICHAEL FIX & RAYMOND STRUYK, CLEAR AND CONVINCING EVIDENCE: MEASUREMENT OF DISCRIMINATION IN AMERICA (1993) although they sometimes incidentally provide evidence of sex discrimination. Marc Bendick, Charles Jackson & Victor Reinoso, *Measuring employment discrimination through controlled experiments*, 23 REVIEW OF BLACK POLITICAL ECONOMY 25 (1994). The audit method has been criticized on the grounds that (1) no two auditors can truly be identical; (2) the studies are not double blind and the testers may be influenced by a desire to please researchers. J. J. Heckman & P. Siegelman, *The Urban Institute Audit Studies: Their Methods and Findings,* CLEAR AND CONVINCING EVIDENCE: MEASUREMENT OF DISCRIMINATION IN AMERICA 187–258 (1993); J. J. Heckman, *Detecting Discrimination*, 12 JOURNAL OF ECONOMIC PERSPECTIVES 101 (1998). Neumark et al. addressed these concerns by minimizing the contact that applicants had with decisionmakers, who instead based decisions primarily on their resumes. Cf. M. Bertrand & S. Mullainathan, *Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination*, 94 THE AMERICAN ECONOMIC REVIEW 991-1013 (2004) (race study, resumes only, significant discrimination found).

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.[109] The most recent meta-analysis by Davison and Burke found a much larger effect.[110] This larger effect 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,[111] and is consistent with observation that in real labor markets, men in female-dominated professions earn less than women in those professions.[112] Field studies likewise find that each sex receives better performance ratings in

[109] J. D. Olian, D. P. Schwab & Y. Haberfeld, *The impact of applicant gender compared to qualifications on hiring recommendations: A meta-analysis of experimental studies*, 41 ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES 80-1 (1988); H. L. Tosi & S. W. Einbender, *The effects of the type and amount of information in sex discrimination research: A meta-analysis*, 28 ACADEMY OF MANAGEMENT JOURNAL 712-723 (1985). See also V. Nieva & B. Gutek, *Sex differences on evaluation*, 5 ACAD. MANAGE. REV 267-76 (1980) (narrative review finding most but not all studies find some sex discrimination).

[110] Heather K. Davison & Michael J. Burke, *Sex Discrimination in Simulated Employment Contexts: A Meta-analytic Investigation*, 56 JOURNAL OF VOCATIONAL BEHAVIOR 225-248 (2000)

[111]The first meta-analyses did not address the sex role congruence issue, Tosi and Einbender, *supra* note___, while the second was unable to reach a conclusion because of the sample size. Olian, Schwab, and Haberfeld, *supra* note___. Tosi and Einbender 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. A meta-analysis of sex bias across contexts, including employment, found no overall discrimination but discrimination towards individuals acting in a non-sex-role congruent field.

[112] Macpherson and Hirsch, *supra* note___; Patricia A. Roos, *Sex Stratification in the Workplace: Male-Female Difference in Economic Returns to Occupation*, 10 SOCIAL SCIENCE RESEARCH 195 (1981); G. Johnson & G. Solon, *Estimates of the Direct Effects of Comparable Worth Policy*, 76 AMERICAN ECONOMIC REVIEW 1117-25 (1986); MARK KILLINGSWORTH, THE ECONOMICS OF COMPARABLE WORTH. (1990). For evidence that sex stereotypes track actual labor market occupational segregation, see M. A. Cejka & A. H. Eagly, *Gender-Stereotypic Images of Occupations Correspond to the Sex Segregation of Employment*, 25 PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN 413 (1999).

performance measures that are sex-role congruent, but these studies cannot fully control for possible differences in actual performance.[113]

2. Type of Discrimination

Sex discrimination may take several forms, all of which will tend to produce occupational segregation. Discrimination is statistical only if based on a rational estimate from all available information on the relevant aptitudes of both groups. Stereotype-based discrimination occurs when employers have stereotypes about the relative productivity of men and women that are not based on available evidence. Taste-based discrimination may result if employers, co-workers, or customers have a discriminatory preference for one sex or the other.

The strongest reason to believe that discrimination is statistical would come from evidence that employers will update their decisions based on new information. Stereotypes are, almost by definition, resistant to updating: if stereotype 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 available information and should incorporate new information. Those inferences may be false, as when an employer draws inferences about the relative competence of the sexes from the current labor market, which may reflect the effects of past discrimination. Nonetheless, if the inference is based on the best cost-effective information available, it is a rational statistical judgment. A rational employer should be willing to override a generalization given sufficiently detailed information about an individual candidate. As a corollary, 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.

The 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. The most complete meta-analysis of these studies supports the view that a subject's reliance on information about the

[113] Chieh-Chen Bowen, Janet K. Swim & Rick R. Jacobs, *Evaluating Gender Biases on Actual Job Performance of Real People: A Meta-Analysis*, 30 Journal of Applied Social Psychology 2194-2215 (2000).

applicant's sex decreases when the subject is given more information about applicant productivity,[114] suggesting that discrimination is at least in part rational and updatable.

Other studies, in contrast, point to the role of at least some non-rational forces producing discriminatory outcomes. A rational agent may use statistical generalizations in addition to individualized data about an applicant but should not allow those generalizations to color the process of gathering individualized information. In fact, many studies find that task specific evaluations of individuals are often influenced by sex based generalizations. For example, two otherwise identical resumes with male and female names attached will be evaluated differently.[115]

IV. THE LAW AND POLICY OF DISCRIMINATION

Statistical discrimination is not self-evidently the type of wrong that Title VII was intended to address. 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 similar to statistical discrimination.[116] Moreover, statistical discrimination can sometimes be economically efficient, and some might argue that the efficiency losses of prohibiting it must be weighed against gains in equity.

These considerations might cause some to reconsider the Supreme Court's conclusion that Title VII forbids employers to use even true

---

[114] Davison and Burke, *supra* note___.

[115] C. C. Bauer & B. B. Baltes, *Reducing the Effects of Gender Stereotypes on Performance Evaluations*, 47 SEX ROLES 465-476 (2002). This observation is an example of a more general phenomenon in which an observer's correct information about one dimension of performance or of prior performance inappropriately influences evaluation of other dimensions or current performance. See, e.g. B. B. Baltes & C. P. Parker, *Understanding and Removing the Effects of Performance Cues on Behavioral Ratings*, 15 JOURNAL OF BUSINESS AND PSYCHOLOGY 229-246 (2000); R. F. Martell & C. E. Willis, *Effects of Observers' Performance Expectations on Behavior Ratings of Work Groups: Memory or Response Bias?*, 56 ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES 91-91 (1993).

[116] Many observers have a reflexive tendency to assume that any distributive effect of the use of group characteristics is inherently invidious. For a thoughtful critique of this, see C. S. Diver & J. M. Cohen, *Genophobia: What Is Wrong with Genetic Discrimination*, 149 U. PA. L. REV. 1439 (2000)

statistical generalizations about sex as the basis for decisions.[117] For present purposes, though, this holding will be taken as given. Policy considerations about the precise effects of statistical discrimination do remain relevant to two important questions: how should the law achieve a minimal level of enforcement of the prohibition on statistical discrimination, and how far past the minimum level should this policy be pushed?

A. STATISTICS IN DISCRIMINATION LAW AND POLICY

Discrimination is sometimes inferred when the representation of a group in a given job is not proportional to the composition of the population comparisons as a whole. Some critics argue that such inferences are central to 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. [118]

Under current law, comparisons to population figures can seldom be the basis for the imposition of Title VII liability. Both the formulation of affirmative action plans and their legal status often depend on population comparisons. These comparisons are problematic if ability distributions in fact vary.

1. STATISTICAL EVIDENCE IN THE PLAINTIFF'S CASE

Statistical evidence can be introduced in two types of Title VII claim. 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.[119] To establish discrimination, class plaintiffs introduce evidence of discrimination against individual class members and provide

[117] City of Los Angeles Department of Water and Power v. Manhart , 435 U.S. 702 (1978). See also *Code of Federal Regulations*, Section *1604.2.*

[118] Kingsley R. Browne, *Statistical Proof of Discrimination: Beyond "Damned Lies"*, 68 WASHINGTON LAW REVIEW 477 (1993); RICHARD A. EPSTEIN, FORBIDDEN GROUNDS: THE CASE AGAINST EMPLOYMENT DISCRIMINATION LAWS (1992).

[119] International Brotherhood of Teamsters v. United States, 431 U.S. 324 (1977).

statistical evidence of defendant's past treatment of the protected group.[120] Statistics alone suffice only if the discrepancies are extreme.[121]

In a disparate impact case, the plaintiff does not need to show that the employer discriminated intentionally. Instead, 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.[122] If the plaintiff succeeds, the defendant may show that that practice is job-related and consistent with business necessity.[123]

Under either disparate treatment or disparate impact theory, the appropriate statistical comparison differs slightly in hiring and promotion cases. In hiring cases, the Supreme Court did authorize the use of the comparisons between the employer's workforce and the general population or the entire workforce in the 1977 case *Brotherhood of Teamsters v. United States*.[124] 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."[125] Initially, some lower courts did not consistently apply the *Hazelwood* rule, and permitted

---

[120]Statistical evidence is not an absolute requirement in systemic disparate treatment cases, but virtually all plaintiffs attempt to provide such evidence .Pitre v. Western Elec. Co., 843 F.2d 1262 (10th Cir. 1988) (holding statistical evidence not an absolute requirement)

[121]Hazelwood School District v. United States, 433 U.S. 299 (1977). The Teamster's Court indicated that individual testimony brings "the cold numbers convincingly to life," Teamsters, *supra* note___.

[122] The plaintiff also has the burden of identifying the challenged practice with particularity and of showing a causal nexus between the practice and the disparate outcome. The elements of a disparate impact case are now governed by statute. 42 USC § 2000e-2(k).

[123] 42 USC § 2000e-2(k)

[124]

> absent explanation, it is ordinarily to be expected that nondiscriminatory hiring practices will in time result in a work force more or less representative of the racial and ethnic composition of the population in the community from which employees are hired. Evidence of long-lasting and gross disparity between the composition of a work force and that of the general population thus may be significant even though § 703(j) makes clear that Title VII imposes no requirement that a work force mirror the general population.

Id

[125]Hazelwood School District v. United States, *supra* note___.

Formatted: footnote quote*

general workforce comparisons for skilled jobs.[126] In 1989, in *Ward's Cove Packing Co. v. Atonio,* the Court ruled unambiguously that statistical comparisons must be limited to the " 'qualified …population in the relevant labor market.' "[127] Although *Ward's Cove* was a disparate impact case, the Court relied on *Hazelwood* and made clear that the same basic principles applied in both disparate impact and disparate treatment cases.[128]

Since *Ward's Cove*, courts have adhered to the *Hazelwood* rule that general population comparisons can be used only in relatively low skill jobs. Under the *Hazelwood* rule, population comparisons have been disallowed for numerous positions.[129] In hiring cases, the *Hazelwood* rule is typically implemented in one of two ways. Courts may 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 including the EEOC.[130] Often, however, even relevant labor market data is regarded as unfair on the grounds that it may not take into consideration considerations unique to employer.[131] In such cases, courts may look at actual applicant data,

---

[126] Laborde v. Regents of the Univ. of California, 495 F. Supp. 1067(C.D. Cal. 1980), aff'd, 686 F.2d 715 (9th Cir. 1982) (university faculty); United States v. City of Miami, 614 F.2d 1322, 1330(5th Cir. 1980), reh'g on other grounds, 664 F.2d 435, 27 FEP 913 (5th Cir. 1981)(all City employees); EEOC v. Radiator Specialty Co., 610 F.2d 178 (4th Cir. 1979)(clerical, sales, professional, and managerial job classification); Jurgens v. Thomas, 29 FEP 1561 (N.D. Tex. 1982)(EEOC personnel); Quigley v. Braniff Airways, Inc., 85 F.R.D. 74, 23 FEP 1811 (N.D. Tex. 1979)(flight attendants). To be sure, many courts did follow Hazelwood and reject general population comparisons. Khambata v. National Ass'n of Counties, 32 FEP 353 (D.D.C. 1982) (program manager of a grant program for county governments) Ramon v. Texas Dep't of Human Res., 31 FEP 1041 (S.D. Tex. 1983)(administrative positions); O'Neal v. Riceland Foods, 684 F.2d 577 (8th Cir. 1982) (clerical positions).

[127] Ward's Cove Packing v. Atonio, 490 U.S. 642, quoting Hazelwood School District v. United States, *supra* note___.

[128]The Civil Rights Act of 1991, PL 102-166, modified some aspects of *Ward's Cove*, but left the rule on workforce comparisons intact. Sections 2 and 3 of the Act, as well as the interpretive memorandum, make clear that Congress objected only to the Court's rulings on business necessity and job relatedness, and no Section of the Act addresses the issue of the relevant workforce to be used in comparisons.

[129]Carter v. Ball, 33 F.3d 450(4th Cir. 1994)(naval intellgince) Cross v. City of Ontario, 67 FEP 725 (C.D. Cal. 1995)(law enforcement); Alexander v. Fulton County, 207 F.3d 1303, (11th Cir. 2000)(law enforcement)..

[130] Census 2000 Special EEO File, http://www.eeoc.gov/stats/census/index.html.

[131] EEOC v. Turtle Creek Mansion Corp., 70 FEP 899 (N.D. Tex. 1995), aff'd, 82 F.3d 414 (5th Cir. 1996); see generally Ramona L. Paetzold and Steven L. Willborn, The

reduced to include only those applicants deemed qualified. For some time now, courts seem to have a preference for applicant flow data.

Where employees challenge the employer's promotions and other policies towards current employees, general labor market comparisons are not relevant, but courts have found other ways of implementing the *Hazelwood* principles. In early cases, plaintiffs would attempt to make comparisons with the entire class of employees theoretically eligible for a particular promotion. Courts today, however, restrict the comparison class to the subset of eligible employees who are meet relevant qualification standards. In addition, of course, a showing of statistical disparity is almost never enough to compel judgment for plaintiff. In a disparate treatment case, the plaintiff must also produce evidence of intent, and in a disparate impact case the defendant has the opportunity to demonstrate business necessity.

Perhaps the only situation in which general population statistics are used for skilled jobs occurs in cases where the employer recruits and trains unskilled applicants, a fact which may explain why firefighting and law enforcement are one of the few job categories in which class action challenges are still common. In such cases, criticism of the use of statistical evidence has some merit, and these doctrines deserve re-examination. The perception that statistical comparisons are based on general labor market data may be based on these cases, which receive a great deal of publicity. Such cases seem to have done great damage to the image of class action litigation, and may contribute to its current disfavor among the public and courts.

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 relevant workforce comparisons will tend to perpetuate the effects of prior discrimination, whether statistical or otherwise, since that discrimination will be reflected in the workforce numbers. The composition of the applicant pool will also reflect current discrimination, since the members of the disadvantaged sex will be discouraged from applying.

2. VOLUNTARY AFFIRMATIVE ACTION

The imposition of Title VII liability does not embody any presumption that all groups be proportionately represented in all positions.

Statistics of Discrimination: Using Statistical Evidence in Discrimination Cases 4.3 (2007).

Such a presumption, may, however, may be implicit in some policy analysis and voluntary affirmative action schemes.[132]

The Supreme Court last considered affirmative action under Title VII in *Johnson v. Santa Clara County*, decided in 1987.[133] The Transportation Agency of Santa Clara County had voluntarily promulgated an affirmative action plan for hiring and promotion. The Agency stated that its long-term goal was "a work force whose composition reflected the proportion of minorities and women in the area labor force." The Supreme Court approved the use of a goal based on total workforce comparisons subject to several caveats, of which two are important here. First, the actual implementation of the plan required not immediate parity but incremental improvements over the number of women in the qualified workforce. Second, "the Agency's Plan was intended to attain a balanced work force, not to maintain one."

Johnson's use of affirmative action to achieve parity may make be useful in ending statistically-based race discrimination. With no underlying ability differences, long-run population targets are appropriate. Economic models suggest that the labor market can be moved to the efficient equilibrium either by affirmative action policies or by monitoring the employment process to ensure that employers are not using statistical proxies in employment decision. As Steve Coate and Glenn Loury have pointed out, affirmative action may be more effective than costly monitoring of employers.[134] A well-designed policy aims for proportionate representation, but proceeds incrementally based on current skill levels to avoid creating disincentives for the disadvantaged to acquire skills.[135] As these policies change incentives for human capital accumulation, differences diminish, the optimal stable equilibrium and antidiscrimination policies become superfluous. Thus, a plan like the one in *Johnson*, that uses long term population goals based on short term current labor market measures may be able to attain a balanced workforce yet not be required to maintain it.

This framework for analyzing affirmative action, however, becomes highly problematic when applied to sex. Short run incremental goals based on current qualified work force may serve the important goal

---

[132] FEDERAL GLASS CEILING COMMISSION, BETTE WOODY & CAROL WEISS, BARRIERS TO WORK PLACE ADVANCEMENT: THE EXPERIENCE OF THE WHITE FEMALE WORK FORCE (1993).

[133] Johnson v. Transportation Agency, Santa Clara County, 480 U.S. 616 (1987).

[134] Coate and Loury, *supra* note___.

[135] *Id.*

of reducing statistical discrimination. But a long run target based on a comparison with general population figures does, as critics complain, assume the absence of underlying ability distributions. Yet calculating a nondiscriminatory equilibrium is a daunting task that requires understanding of the exact nature of the abilities required for a job and also the true distribution of those abilities in a nondiscriminatory world.

Even if appropriate targets could be formulated, a plan that used them could not meet *Johnson*'s requirement that the plan be a means to attaining rather than maintaining a balanced workforce. With no differences in underlying ability distribution, affirmative action plans eventually become unnecessary. Once the nondiscriminatory equilibrium is reached, it maintains itself. But with permanent difference, affirmative action must be permanent or else the labor market will revert to a statistically discriminatory point.

Numerically-based theories of liability are thus too weak to deal with sex based statistical discrimination, and affirmative action policies potentially overreach.

B. SCRUTINY OF EMPLOYER DECISIONS

The vast majority of discrimination claims involve neither statistical theories of liability nor affirmative action. Instead, they examine individual decisions made by employers. Current doctrine may not be fully adequate to the task of combating statistical discrimination, but understanding how statistical discrimination works suggests ways to improve the tools available to courts.

1. UNCONSCIOUS STATISTICAL BIAS

If employers today use sex-based statistical generalizations in employment decisions, they are unlikely to do so by openly issuing guidelines dictating the use of these generalizations. Instead, statistical generalizations are most likely to be translated into action through subjective decisions made by individual supervisors.

To the extent that an employer's supervisory personnel consciously albeit informally make use of statistical generalizations, plaintiffs may be able to provide conventional evidence of intent in either a class or individual disparate treatment case. For example, a plaintiff might introduce statements of managers, made in the context of an individual employment decision, that women lack the skills necessary to perform the particular job in question.

Such evidence, however, is increasingly rare. In an important article, Linda Hamilton Krieger documented evidence of biased decision-making even by individuals who had a sincere desire to make unbiased choices.[136] Subsequent psychological research discussed by Krieger and others has provided additional evidence that much biased decision-making is unconscious.[137]

The problem of unconscious bias is compounded by a phenomenon described by Susan Sturm. Sturm has noted that current discrimination does not fit the model of what she calls first generation discrimination, which often took the form of easily identified decisions in which discriminatory intent played an overt role. So-called second generation employment discrimination results from the gradual accretion of smaller discriminatory choices made by a variety of actors. These day-to-day occurrences cannot be readily documented.[138]

For the most part, the literature on unconscious bias has focused on irrational bias. Though there is evidence that such bias continues to exist, statistical discrimination is in many ways a more serious policy problem, since market pressures not only do not operate against it, but reinforce it.

The policy tools used to address rational and irrational bias may sometimes vary. To use the taxonomy of Christine Jolls and Cass Sunstein,[139] debiasing mechanisms work only against irrational stereotypes, since rational stereotypes are, of course rational. However, strategies to insulate decisions against the effects of bias can be useful regardless of the rationality or irrationality of the biased belief.

---

[136]Linda Hamilton Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 STANFORD LAW REVIEW 1161-1248 (1995).

[137] The literature is too large to list in its entirety. Some representative articles include Ann C. McGinley, *Viva La Evolucion: Recognizing Unconscious Motive in Title VII*, 9 CORNELL J.LAW & PUB. POL'Y. 415 (1999); Tristin K. Green, *Discrimination in Workplace Dynamics: Toward a Structural Account of Disparate Treatment Theory*, 38 HARV. C.R.-C.L. L. REV. 91 (2007); Melissa Hart, *Subjective Decisionmaking And Unconscious Discrimination*, 56 ALABAMA LAW REVIEW 741 (2005); Christine Jolls & Cass R Sunstein, *The Law of Implicit Bias* , 94 CALIFORNIA LAW REVIEW 969 (2006); Linda Hamilton Krieger & Antony G. Greenwald, *Implicit Bias: Scientific Foundations*, 94 CALIFORNIA LAW REVIEW 945 (2006).

[138] Susan Sturm, *Second Generation Employment Discrimination: A Structural Approach*, 101 COLUMBIA LAW REVIEW 458-568 (2001).

[139] Jolls and Sunstein, *supra* note___.

One potentially valuable insulating strategy is the regulation of subjective decisions. In a disparate treatment suit, the practice of giving unconstrained subjective discretion to supervisors may be considered evidence of intent. In disparate impact litigation, the employer delegation of subjective decision-making power to supervisors may be challenged as a "practice" and the employer can be held liable if the discretion is exercised in a discriminatory manner.[140] 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.

Courts have in principle shown sympathy to the principle that excessively subjective decision making opens the door to discrimination.[141] They have been less willing in practice to base liability on the use of such practices,[142] in part because of an understandable concern about whether courts are competent to evaluate employment practices.[143] Any effort to increase monitoring of employer practices must be based on a solid foundation of empirical work supporting the used of the practices that the law encourages.

2. STRUCTURED VERSUS UNSTRUCTURED ASSESSMENT

Psychologists have generated a voluminous literature on the accuracy of various types of assessment procedures in a variety of contexts, including medicine, law enforcement and business. Perhaps the largest and most rigorous has emerged from the distinction between algorithmic and clinical judgments.[144] Algorithmic judgments, sometimes also called statistical, mechanical, or objective judgments, are based on a clearly specified decision procedure that contains no discretionary elements. Clinical judgments, roughly equivalent to holistic or subjective judgments, are made by a human observer, often an expert, without the guidance of an explicit decision procedure.

An important line of research has concluded that algorithmic decision-making is almost always superior to clinical judgment. More specifically, mechanical methods are almost invariably superior for

---

[140] Watson v. Fort Worth Bank and Trust, 487 U.S. 977 (1988).

[141] Hart, *supra* note___.

[142] *Id.*

[143] Watson v. Fort Worth Bank and Trust, *supra* note___.

[144] The distinction was first proposed and analyzed in P. E. MEEHL, CLINICAL VERSUS STATISTICAL PREDICTION: A THEORETICAL ANALYSIS AND A REVIEW OF THE EVIDENCE (1954)

combining data, although both clinical and mechanical methods are useful for collecting data,[145] and clinical judgment may be the best starting point for constructing mechanical evaluation algorithms.[146] Hybrid methods that use algorithmic combination derived from clinical methods and along with clinical data are sometimes called structured decision processes.

A large body of research now supports the view that structured methods are almost always as good as, and usually superior to, unstructured ones.[147] The algorithmic-clinical literature cuts across subject matter, but its findings are generally confirmed by studies that focus on personnel decision making. Academic research has reached a strong consensus that structured interviews are more accurate than unstructured ones in predicting future performance.[148]

The strength of these results has been somewhat obscured by the existence of an apparently contradictory literature on so-called subjective versus objective evaluation of job performance.[149] Selection methods can

---

[145] J. Sawyer, *Measurement and prediction, clinical and statistical.*, 66 PSYCHOL BULL 178-200 (1966).

[146] H. J. Einhorn, *Expert measurement and mechanical combination*, 7 ORGANIZATIONAL BEHAVIOR AND HUMAN PERFORMANCE 86-106 (1972).

[147] William M. Grove et al., *Clinical versus mechanical prediction: A meta-analysis.*, 12 PSYCHOLOGICAL ASSESSMENT 19-30 (2000); Stefania Aegisdottir et al., *The Meta-Analysis of Clinical Judgment Project: Fifty-Six Years of Accumulated Research on Clinical Versus Statistical Prediction*, 34 THE COUNSELING PSYCHOLOGIST 341-382 (2006).

[148]Major narrative reviews include M. A. Campion, D. K. Palmer & J. E. Campion, *A Review of Structure in the Selection Interview.*, 50 PERSONNEL PSYCHOLOGY (1997); R. D. Arvey & J. E. Campion, *The Employment Interview: A Summary and Review of Recent Research.*, 35 PERSONNEL PSYCHOLOGY 281-322 (1982); N. Schmitt, *Social and situational determinants of interview decisions: Implications for the employment interview.*, 29 PERSONNEL PSYCHOLOGY 79-101 (1976); M. M. Harris, *Reconsidering the employment interview: A review of recent literature and suggestions for future research.*, 42 PERSONNEL PSYCHOLOGY 691-726 (1989); E. C. Mayfield, *The selection interview: A re-evaluation of published research.*, 17 PERSONNEL PSYCHOLOGY 239-260 (1964). Meta-analytic reviews include A. I. Huffcutt & WJR Arthur, W.J.R., *Hunter and Hunter(1984) revisited: interview validity for entry-level jobs*, 79 JOURNAL OF APPLIED PSYCHOLOGY 184-190 (1994); F. L. Schmidt & J. E. Hunter, *The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings.*, 124 PSYCHOLOGICAL BULLETIN 262-274 (1998).

[149] R. L. Heneman, *The Relationship Between Supervisory Ratings and Results-Oriented Measures of Performance: A Meta-Analysis.*, 39 PERSONNEL PSYCHOLOGY 811-26 (1986) (distinguishing between structured and unstructured; structured correlate better with objective than unstructured); W. H. Bommer et al., *On the interchangeability of objective and subjective measures of employee performance: A meta-analysis.*, 48 PERSONNEL PSYCHOLOGY 587-605 (1995) (distinguishing between structured and unstructured but apparently coding some structured as unstructured; subjective and

be validated using later performance reviews, but there is no measure of performance to validate the performance reviews themselves. Thus, studies of performance evaluation are limited to comparisons between different performance measures without determining the accuracy of each method. To complicate matters, these studies classify all supervisory ratings as subjective, since they contain some subjective component. No distinction is usually made between structured and unstructured supervisory evaluation procedures. Even when the distinction is made, these studies at best compare structured and unstructured ratings to objective performance measures like sales, a correlation that tells us little about the relative accuracy of the various methods.

A number of laboratory experiments confirm the potential of structured decision-making to reduce discrimination. In one series of experiments, subjects were asked first to rank individual candidates for a traditionally male job. They were then asked to state which qualifications were most relevant to the job. Their after-the-fact ranking of job qualifications varied so as to justify a preference for a male candidate. After evaluating a male candidate with more education and less job experience than a female alternative, subjects stated that education was more important than job experience. When the male candidate had more job experience and less education, subjects afterwards stated that job experience was more important than education. Sex bias in candidate rankings greatly diminished in a second study when subjects were first asked to rank job qualifications and then individual candidates.[150]

---

objective measures different, but no difference between structured and unstructured); B. R. Nathan & R. A. Alexander, *A Comparison of Criteria for Test Validation: A Meta-Analytic Investigation.*, 41 PERSONNEL PSYCHOLOGY 517-35 (1988)(clerical workers; all supervisory ratings, both structured and unstructured combined, *Id.* at 521**; not different from objective measures); D.S. Ones, C. Viswesvaran & F. L. Schmidt, *Comprehensive meta-analysis of integrity test validities findings and implications for personnel selection and theories of job performance*, 78 JOURNAL OF APPLIED PSYCHOLOGY 679-703 (1993)(apparently combining structured and unstructured supervisor ratings; finding no difference between subjective and objective measures); W. F. CASCIO & H. AGUINIS, APPLIED PSYCHOLOGY IN HUMAN RESOURCE MANAGEMENT (1998)(same).

[150] Michael I. Norton, Joseph A. Vandello & John M. Darley, *Casuistry and Social Category Bias*, 87 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 817-831 (2004). See also Eric Luis Uhlmann & Geoffrey L. Cohen, *Constructed Criteria: Redefining Merit to Justify Discrimination*, 16 PSYCHOLOGICAL SCIENCE 474-480 (2005); G. Hodson, J. F. Dovidio & S. L. Gaertner, *Processes in Racial Discrimination: Differential Weighting of Conflicting Information*, 28 PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN 460 (2002); C. K. Hsee, *Elastic Justification: How Unjustifiable*

Interviews and supervisory evaluations are important parts of most real world personnel systems. For the most part these procedures are far less structured than research would recommend.[151] The legal system can encourage use of more structure by considering it favorably in evaluating employer liability.

3. *DIRECT COSTS*

Before the legal system adopts rules that encouraging more structured decision-making, the benefits of such rules must be weighed against potential costs. One obvious set of costs are the direct costs incurred by employers on implementing more structured personnel systems. Perhaps firms choose not to use structured practices because those practices are too expensive.

Several considerations suggest that the burden of direct employer costs should not be a basis for concern. A number of studies suggest that greater structure in hiring and job performance assessment is associated with greater profitability.[152] In surveys of hiring practices, the reason most often given for not using structured interviews was not cost but rather the belief that adding structure would not improve the quality of information received.[153] Typically, this view was expressed by managers with little or

---

Factors Influence Judgments, 66 ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES 122-129 (1996).

[151] Virtually all American firms use interviews, but only about 35% use fixed interview questions, much less more sophisticated forms of structure. A. M. Ryan et al., *An International Look at Selection Practices: Nation and Culture as Explanations for Variability in Practice.*, 52 PERSONNEL PSYCHOLOGY 359-362, 378 (1999); David E. Terpstra & Elizabeth J. Rozell, *The Relationship of Staffing Practices to Organizational Level Measures of Performance*, 46 PERSONNEL PSYCHOLOGY 27-48 (1993); R. L. Dipboye, *Structured selection interviews: Why do they work? Why are they underutilized*, 13 INTERNATIONAL HANDBOOK OF SELECTION AND ASSESSMENT 455-473 (1997). Various studies have found that between 75% and 90% of all firms use some type of formal appraisal system. K. R. MURPHY & J. N. CLEVELAND, UNDERSTANDING PERFORMANCE APPRAISAL 5 (1995). Because appraisal systems come in so many varieties it is hard to quantify the extent to which they are adequately structured but a large informal consensus exists among HR academics that the systems used in practice could stand considerable improvement.

[152]See, e.g. Terpstra and Rozell, *supra* note___; Mark A. Huselid, *The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance*, 38 THE ACADEMY OF MANAGEMENT JOURNAL 635-672 (1995).

[153] D. E. Terpstra & E. J. Rozell, *Why Some Potentially Effective Staffing Practices are Seldom Used*, 26 PUBLIC PERSONNEL MANAGEMENT 483-495 (1997) (random sample of 1000 U.S, firms employing over 200 workers; 9% cited costs as reason for not using structured interviews); F. Lievens & A. De Paepe, *An empirical investigation of*

training in personnel matters. Those with more training and experienced generally supported the use of structure both in theory and in their own practice. In another survey, managers said that the single factor that would most improve their own hiring processes was a better understanding of what makes an interview effective.[154]

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 behaviors that constitute good, average and poor performance.[155]

But formal job analysis is not the critical component of structure. The general literature on algorithmic and clinical judgment acknowledges that algorithms derived from clinical judgment maybe quite effective. The crucial factor is the ex ante determination of the structure of the evaluation that prevents the evaluator from introducing undesired considerations into the final decision.

4. INDIRECT EFFICIENCY COSTS

Statistical discrimination clearly places a burden on individuals who cannot engage in the activity at which they are most productive. On the other hand, statistical discrimination can sometimes be economically efficient because it provides employers with information about worker productivity. The benchmark analysis made a long list of assumptions that eliminated any efficiency effects of statistical discrimination: a single labor market; a single output; a single dimension of ability; independence of each worker's marginal product from that of coworkers; full employment; fixed human capital; fixed employer information, and fixed labor supply.[156] With more realistic assumptions, the decision to permit or

---

interviewer-related factors that discourage the use of high structure interviews, (76 large private Belgian firms; no question about cost; ease of preparation significant factor, but less than sum of others including desire for discretion and personal contact and personal characteristics of personnel director, such as training)

154 M. D. Nowicki & J. G. Rosse, *Managers'Views of How to Hire: Building Bridges Between Science and Practice*, 17 JOURNAL OF BUSINESS AND PSYCHOLOGY 157-170 (2002).

155 Michael A. Campion, David K. Palmer & James E. Campion, *Structuring Employment Interviews to Improve Reliability, Validity, and Users' Reactions.*, 7 CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE 77-82 (1998); Campion, Palmer, and Campion, *supra* note___.

156 Aigner and Cain, *supra* note___; Phelps, *supra* note___. For an illuminating discussion of the role of these assumptions, see Joseph E. Stiglitz, *The Theory of*

prohibit statistical discrimination could have efficiency effects. Perhaps in evaluating the vigor with which to combat it, such efficiency losses must be weighed against gains in equity. [157]

a) *Forces Making Statistical Discrimination Efficient*

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 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.[158]

The assumption that each worker's output is independent of the quality of other workers may be unrealistic. Lower ability workers may slow down the higher ability workers and production may be higher if workers are segregated by ability level.[159] Statistical discrimination increases this beneficial segregation.

In real labor markets, ability is not a single homogenous quality. Instead, varying abilities are useful in different job types, 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.[160] Employers in a competitive economy will use all observable information to match employees with jobs. For these purposes, there is no meaningful distinction between observables that are causally connected

"Screening," Education, and the Distribution of Income, 65 THE AMERICAN ECONOMIC REVIEW 283-300 (1975). From the beginning, the literature has also analyzed the consequences of different variances in the signal quality between the two groups. Phelps, *supra* note___ Aigner and Cain, *supra* note___. Differing variances can produce statistical discrimination even in the absence of any difference in the underlying ability distributions. This phenomenon may be of great interest in understanding occupational segregation, but the data in this article does not address this possibility.

157 S. W. POLACHEK & W. S. SIEBERT, THE ECONOMICS OF EARNINGS (1993). The origin for this conclusion appears to be a justly influential paper by George Stigler. G. J. Stigler, *Information in the Labor Market*, 70 THE JOURNAL OF POLITICAL ECONOMY 94 (1962).

158 Stiglitz, *supra* note___.

159 *Id.*

160 More formally, suppose that individuals vary in their ability *A* so that their output is higher when matched with an appropriate job *S*. Workers with a high value of A perform best in high S jobs and less well in low S jobs. Workers with low A perform best in low S jobs. Thus, worker output decreases as A-S, the gap between A and S, increases. A is determined by both observable factors Y and unobservable factors Z, and employers naturally only observe Y.

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 lower output, creating inefficiency.[161]

b) *Forces that may make Statistical Discrimination Inefficient*

The argument that statistical discrimination has no efficiency effects depends on the assumption that the economy consists of single sector whose identical employers all have the same imperfect evidence of ability. In an important 1986 paper, Stewart Schwab considered the consequences of relaxing this assumption.[162] As always, workers differ in their ability. However, the economy now consists of two sectors. In one sector, the "standardized" market, employers have imperfect information about potential employees. Workers may alternatively choose to be self-employed, including employment in home production, in which case they have better information about their own aptitudes than employers in the standardized market. These assumptions imply that market forces will misallocate high ability workers to self-employment. Schwab then considers the implications of statistical discrimination between 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.

The benchmark model assumed that 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 signal costs be lower for high

---

[161] M. Rothschild & J. E. Stiglitz, *A Model of Employment Outcomes Illustrating the Effect of the Structure of Information on the Level and Distribution of Income,"*, 10 ECONOMICS LETTERS 231–236 (1982); R. Haagsma, *Statistical Discrimination in the Labour Market: An Efficient Vice?*, 146 DE ECONOMIST 321-345 (1998).

[162] Stewart Schwab, *Is Statistical Discrimination Efficient?*, 76 AMERICAN ECONOMIC REVIEW 228-234 (1986); For an extension of Schwab's analysis, see R. Haagsma, *Is Statistical Discrimination Socially Efficient?*, 5 INFORMATION ECONOMICS AND POLICY 31-50 (1993).

ability workers than low ability workers. The canonical case of a signal is educational level; other signals include dress and deportment. The example of education can be somewhat misleading, since education (we hope) also increases marginal product as well as signaling ability. The signaling component of education is only the component that does *not* increase marginal product.

Signaling will redistribute wealth to high ability workers. Unless there are efficiency effects from identifying ability, signaling will be inefficient, since the signal is cost that is not offset by any allocative benefit. Market with signaling tends to be characterized by multiple equilibria, many of which will be inefficient, yet stable: The market shows no tendency to converge to an efficient equilibrium. If information about ability has offsetting allocative effects, such as increased labor supply by the more able, signaling will be efficient only if it occurs at the minimum level needed to achieve those effects. Higher levels of the signal will often lead to stable equilibria, however, and this excess signaling will be inefficient.[163]

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. If an employer invests resources in information about employee quality, other firms will have an incentive to poach that firm's employees. In the interests of employee freedom, the law offers little compensation to the original employer. Firms collectively then will tend to underinvest in information, relying on employee generated signals such as education and any available proxies for employee quality or job match. With imperfect information, a ban on statistical discrimination may improve welfare by inducing firms to invest more in screening employees.[164]

---

[163] M. Spence, *Job Market Signaling*, 87 QUARTERLY JOURNAL OF ECONOMICS 355-374 (1973); Stiglitz, *supra* note___; Haagsma, *supra* note___; Michael Spence, *Signaling in Retrospect and the Informational Structure of Markets*, 92 THE AMERICAN ECONOMIC REVIEW 434-459 (2002). See also G. Akerlof, *The Market for Lemons: Qualitative Uncertainty and the Market Mechanism*, 84 QUARTERLY JOURNAL OF ECONOMICS 488-500 (1970).

[164] A useful sketch of one variant has been made by Joseph Altonji and Rebecca Blank, who consider the problems raised by the vicious cycle effect of statistical discrimination on human capital investment. Depending on the level of costs necessary to obtain better information, prohibitions on statistical discrimination might increase employer investment, thus inducing disfavored workers to invest in human capital and break the cycle. Note that this assumes that firms are able to reap some private benefits from information investment, just not the full benefit. This is a plausible assumption, since the investing firms has the benefit of the full results of its research while other firms

c) *Implications*

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 gain or loss from allowing statistical discrimination but no useful generalizations seem possible now.

V. CONCLUSION

Those concerned about the status of women are understandably uncomfortable with any claims about differences in aptitude 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 aptitudes are indeed differently distributed, on average, between men and women. These differences have sometimes 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 this understanding.

---

have simply the signal value of the decision to hire. J. G. Altonji & R. Blank, *Race and Gender in the Labor Market*, 3 HANDBOOK OF LABOR ECONOMICS 3143–3259, 3190-91 (1999). See also Glen G. Cain, *The Economic Analysis of Labor Market Discrimination: A Survey*, *in* HANDBOOK OF LABOR ECONOMICS 693-785, 727 (Orley Ashenfelter & Richard Layard eds., 1986).