I remember once being entreated not to read a certain newspaper lest it might change my opinion upon free-trade. "Lest I might be entrapped by its fallacies and misstatements," was the form of expression. "You are not," my friend said, "a special student of political economy. You might, therefore, easily be deceived by fallacious arguments upon the subject. You might, then, if you read this paper, be led to believe in protection. But you admit that free-trade is the true doctrine; and you do not wish to believe what is not true."

—Charles Sanders Peirce

"The Fixation of Belief," 1877
Contents

A Note on Presentation xi

Introduction 1

Part I
"Gender Is a Social Construct"

1. A Framework for Thinking About Sex Differences 19
2. Sex Differences in Personality 30
3. Sex Differences in Neurocognitive Functioning 44
4. Sex Differences in Educational and Vocational Choices 68
5. Sex Differences in the Brain 93

Part II
"Race Is a Social Construct"

6. A Framework for Thinking About Race Differences 133
7. Genetic Distinctiveness Among Ancestral Populations 141
8. Evolution Since Humans Left Africa 158
9. The Landscape of Ancestral Population Differences 182

Part III
"Class Is a Function of Privilege"

10. A Framework for Thinking About Heritability and Class 209
11. The Ubiquity of Heritability and the Small Role of the Shared Environment 218
A Note on Presentation

*Human Diversity* is grounded in highly technical literatures involving genetics, neuroscience, and statistics. It must satisfy two audiences with completely different priorities: my intended reader and the experts.

I’ve always thought of my intended reader as someone who enjoys reading the science section of the *New York Times*—curious about scientific matters, but someone who wants the gist of the science, not the minutiae. I need to keep the narrative moving. But I am conveying material that often has daunting technical complexities. Readers also need to be able to compare my claims with the details of the underlying evidence. I use my three favorite devices: Boxed text introduces related issues that are interesting but not essential. Appendixes provide full-scale discussions of important ancillary issues. Endnotes expand on points in the main text. But *Human Diversity* uses these devices, especially the endnotes, even more extensively than I have in the past. Some of the endnotes are full-scale essays, complete with tables. Brackets around a callout number for an endnote indicate that it contains at least a substantial paragraph of additional exposition.

For this complicated book, I have had to add a fourth device. In the past, I have usually been able to avoid technical jargon in the main text. *Human Diversity* doesn’t give me that option. Too much material cannot be discussed without using technical terms that will be new to many readers. I therefore insert periodic interludes in the text to explain them.

I have also tried to make the book more accessible by my treatment of charts and tables. Sometimes the information in a figure or table is complicated enough to warrant giving it a title and traditional formatting. But often a simple graph of a trendline or a few summary statistics don’t need the folderol. They can be integrated into the text so that you can absorb the simple point that’s being made and move on.
HUMAN DIVERSITY
Introduction

If you have picked up Human Diversity looking for bombshells, you'll be disappointed. I'm discussing some of the most incendiary topics in academia, but the subtext of the chapters to come is that everyone should calm down. The differences among human groups are interesting, not scary or earth-shaking. If that sounds boring, this isn't the book for you.

If, on the other hand, you have reached this page convinced that gender, race, and class are all social constructs and that any claims to the contrary are pseudoscience, you won't get past the first few pages before you can't stand it anymore. This book isn't for you either.

Now that we're alone, let me tell you what Human Diversity is about and why I wrote it.

The sciences form a hierarchy. "Physics rests on mathematics, chemistry on physics, biology on chemistry, and, in principle, the social sciences on biology," wrote evolutionary biologist Robert Trivers. If so, this century should be an exhilarating time to be a social scientist. Until now, we social scientists—for I am a member of that tribe—have been second-class citizens of the scientific world, limited to data and methods that cast doubt on our claim to be truly part of the scientific project. Now, new possibilities are opening up.

Biology is not going to put us out of business. The new knowledge that geneticists and neuroscientists are providing, conjoined with the kinds of analyses we do best, will enable us to take giant strides in understanding how societies, polities, and economies really function. We are like physicists at the outset of the nineteenth century, who were poised at a moment in history that would produce Ampères and Faradays.

We ought to be excited, but we aren't. Trivers again: "Yet discipline after
discipline—from economics to cultural anthropology—continues to resist growing connections to the underlying science of biology, with devastating effects."

Why the resistance? Because the social sciences have been in the grip of an orthodoxy that is scared stiff of biology.

**The Orthodoxy**

The core doctrine of the orthodoxy in the social sciences is a particular understanding of human equality. I don’t mean *equality* in the sense of America’s traditional ideal—all are equal in the eyes of God, have equal inherent dignity, and should be treated equally under the law—but equality in the sense of sameness. Call it the sameness premise: *In a properly run society, people of all human groupings will have similar life outcomes.* Individuals might have differences in abilities, the orthodoxy (usually) acknowledges, but groups do not have inborn differences in the distributions of those abilities, except for undeniable ones such as height, upper body strength, and skin color. Inside the cranium, all groups are the same.

The sameness premise theoretically applies to any method of grouping people, but three of them have dominated the discussion for a long time: gender, race, and socioeconomic class. Rephrased in terms of those groups, the sameness premise holds that whatever their gender, race, or the class they are born into, people in every group should become electrical engineers, nurture toddlers, win chess tournaments, and write sci-fi novels in roughly equal proportions. They should have similar distributions of family income, mental health, and life expectancy. Large group differences in these life outcomes are prima facie evidence of social, cultural, and governmental defects that can be corrected by appropriate public policy.

The intellectual origins of the orthodoxy go back more than three centuries to the early days of the Enlightenment and the concept of humans as blank slates. The explicit rejection of a role for biology in the social sciences occurred from the end of the nineteenth through the beginning of the twentieth centuries, with the leading roles played by Émile Durkheim in sociology, Franz Boas in anthropology, and John Watson in psychology.³

The political expression of the orthodoxy had its origins in the mid-1960s with the legal triumphs of the civil rights movement and the rise of feminism. In the beginning, the orthodoxy consisted of specific allegations and solutions: Racism keeps black unemployment high. Sexism stunts women’s careers. Affirmative action and antidiscrimination laws are needed. But the orthodoxy soon began to incorporate an intellectual movement that gained momentum in the mid-1960s with the publication of *The Social Construction of Reality* by Peter L. Berger and Thomas Luckmann.

The authors were dealing with an ancient problem: Each of us thinks we know what reality is, but different people have different perceptions of it. "The sociologist is forced by the very logic of his discipline to ask, if nothing else, whether the difference between the two ‘realities’ may not be understood in relation to various differences between the two societies," wrote Berger and Luckmann.⁴ This beginning, written in plain English, perfectly sensible, morphed during the 1970s and 1980s into the orthodoxy position that just about everything is a social construct, often argued in postmodern prose that is incomprehensible to all but the elect.⁵ The sources of human inequalities are artificial, made up, a reflection of the particular reality that a dominant segment of society has decided is the one we must all live by.

As I write, three of the main tenets of the orthodoxy may be summarized as follows:

*Gender is a social construct.* Physiological sex differences associated with childbearing have been used to create artificial gender roles that are unjustified by inborn characteristics of personality, abilities, or social behavior.

*Race is a social construct.* The concept of race has arisen from cosmetic differences in appearance that are not accompanied by inborn differences in personality, abilities, or social behavior.

*Class is a function of privilege.* People have historically been sorted into classes by political, economic, and cultural institutions that privilege heterosexual white males and oppress everyone else, with genes and human nature playing a trivial role if any. People can be re-sorted in a socially just way by changing those institutions.

I have stated these tenets boldly. If you were to go onto a university campus and chat privately with faculty members whose research touches on issues of gender, race, or class, you would find that many of them, perhaps a majority, have a more nuanced view than this. They accept that biology plays a role. Why then don’t they mention the evidence for a biological role in their lectures? Their writings?

A common answer is that they fear that whatever they write will be
misinterpreted and misused. But it’s easy to write technical articles so that the mainstream media never notice them. The real threat is not that the public will misuse a scholar’s findings, but that certain fellow academicians will notice those findings and react harshly.

Therein lies the real barrier to incorporating biology into social science. It is possible to survive on a university campus without subscribing to the orthodoxy. But you have to be inconspicuous, because the simplistic version of the orthodoxy commands the campus’s high ground. It is dangerous for a college faculty member to say openly in articles, lectures, faculty meetings, or even in casual conversations that biology has a significant role in creating differences between men and women, among races, or among social classes. Doing so often carries a price. That price can be protests by students, denial of tenure-track employment for postdocs, denial of tenure for assistant professors, or reprimands from the university’s administrators.

The most common penalties are more subtle. University faculties are small communities, with all the familiar kinds of social stigma for misfits. To be openly critical of the orthodoxy guarantees that a vocal, influential element of your community is going to come after you, socially and professionally. It guarantees that many others will be reluctant to be identified with you. It guarantees that you will get a reputation that varies from being an eccentric at best to a terrible human being at worst. It’s easier to go along and get along.

The risks that face individual faculty members translate to much broader damage to academia. We have gone from a shared telos for the university, exemplified by Harvard’s motto, “Veritas,” to campuses where professors must be on guard against committing thought crimes, students clamor for protection against troubling ideas, codes limiting the free expression of ideas are routine, and ancient ideals of scholarly excellence and human virtue are derided and denounced.6 On an individual level, social scientists have valid rationales to avoid exploring the intersection of biology and society. Collectively, their decisions have produced a form of de facto and widespread intellectual corruption.

Archaeological Digs

The good news is that some scholars have been exploring the intersection of biology and society despite the risks—so many that the orthodoxy is in the process of being overthrown. The heavy lifting is being done not within the social sciences, but by biologists and, more specifically, by geneticists and neuroscientists. They have been accumulating data that will eventually pose the same problem for defenders of the sameness premise that Aristotelian physicists faced when Galileo dropped objects from heights. Everyone could see that they didn’t behave as Aristotle’s theory predicted. No one could offer a counterargument. When our understanding of the genome and the brain is sufficiently advanced—and it is approaching that point faster than most people realize—the orthodoxy will be in the same position. Continuing to defend the sameness premise will make them look silly. It is my belief that we are nearing inflection points and that the triumph of the revolution will happen quickly. The key battles are likely to be won within the 2020s. This book is a progress report.

In the course of writing Human Diversity, it became apparent to me that progress is at strikingly different points for gender, race, and class. The analogy of an archaeological dig of a buried city comes to mind.

The dig for gender is well along. Excavations have been extensive, the city’s layout has been identified, and thousands of artifacts have been found. There’s lots yet to be done, but the outlines of the city and its culture are coming into focus.

The dig for race is in its early stages. Topological analysis has identified a promising site, initial clearing of the site has been completed, and the first probes have established that there’s something down there worth investigating. Scientists are just beginning excavation.

The dig for class had been largely completed by the end of the twentieth century, and scholars in this century had until recently been kept busy analyzing the artifacts. They are now returning to the site with newly developed tools.

Analogies aren’t precise, but this one explains the organization of the book. I begin with gender differences and devote five substantial chapters to them. A lot has been securely learned about gender differences. Race gets shorter chapters describing how the site was located, how it has been cleared, and the evidence that there’s something down there worth investigating. The chapters on class summarize findings that for the most part have been known for decades.

Why Me?

I am neither a geneticist nor a neuroscientist. What business do I have writing this book?
The answer is that specialists are seldom good at writing overviews of their specialties for a general audience because they know too much—the forest and trees problem. It's often easier for an outsider to communicate the specialists' main findings to other outsiders. There are personal reasons as well. I think I'm skilled at making the findings of technical literatures accessible to a broader audience; I enjoy doing it, and I have been a fascinated observer of developments in genetics and neuroscience for years. I'm also at a point in my career when I'm immune to many of the penalties that a younger scholar would risk.

That career includes the firestorm that followed the publication of The Bell Curve more than a quarter of a century ago, an experience that has been on my mind as I have written Human Diversity. How can I avoid a repeat? Perhaps it's impossible. The background level of animosity and paranoia in today's academia is much worse than it was in 1994. But here is the reality: We are in the midst of a uniquely exciting period of discoveries in genetics and neuroscience—that's good news, not bad. My first goal is to describe what is being learned as clearly as possible, without sensationalism. I hope you will finish the book understanding that there are no monsters in the closet, no dread doors that we must fear opening.

My second goal is to stick to the low-hanging fruit. Almost all of the findings I report are ones that have broad acceptance within their disciplines. When a finding is still tentative, I label it as such. I know this won't deter critics from saying it's all pseudoscience, but I hope the experts will be yawning with boredom because they know all this already. Having done my best to accomplish those two things, I will hope for the best.

**WHY THERE IS SO LITTLE ABOUT EVOLUTIONARY PSYCHOLOGY IN HUMAN DIVERSITY**

Hundreds of millions of years of evolution did more than shape human physiology. It shaped the human brain as well. A comparatively new discipline, evolutionary psychology, seeks to understand the links between evolutionary pressures and the way humans have turned out. Accordingly, evolutionary psychology is at the heart of explanations for the differences that distinguish men from women and human populations from each other. Ordinarily, it would be a central part of my narrative. But the orthodoxy has been depressingly successful in demonizing evolutionary psychology as just-so stories, I decided that incorporating its insights would make it too easy for critics to attack the explanation and ignore the empirical reality.

I discuss some evolutionary material in my accounts of the peopling of the Earth and the source of greater male variance. That's it, however, ignoring the rest of the fascinating story. The note gives you some sources for learning more. 

---

**The 10 Propositions**

The propositions that accompany most of the chapters are intended to exemplify low-hanging fruit. I take on an extremely broad range of topics, but with the limited purpose of clarifying a handful of bedrock issues.

I apologize for the wording of the 10 propositions—they are not as snappy as I would prefer—but there's a reason for their caution and caveats. On certain important points, the clamor of genuine scientific dispute has abated and we don't have to argue about them anymore. But to meet that claim requires me to state the propositions precisely. I am prepared to defend all of them as "things we don't have to argue about anymore"—but exactly as I worded them, not as others may paraphrase them.

Here they are:

1. Sex differences in personality are consistent worldwide and tend to widen in more gender-equalitarian cultures.
2. On average, females worldwide have advantages in verbal ability and social cognition while males have advantages in visuospatial abilities and the extremes of mathematical ability.
3. On average, women worldwide are more attracted to vocations centered on people and men to vocations centered on things.
4. Many sex differences in the brain are coordinate with sex differences in personality, abilities, and social behavior.
5. Human populations are genetically distinctive in ways that correspond to self-identified race and ethnicity.
6. Evolutionary selection pressure since humans left Africa has been extensive and mostly local.
7. Continental population differences in variants associated with personality, abilities, and social behavior are common.
8. The shared environment usually plays a minor role in explaining personality, abilities, and social behavior.

9. Class structure is importantly based on differences in abilities that have a substantial genetic component.

10. Outside interventions are inherently constrained in the effects they can have on personality, abilities, and social behavior.

On all 10, the empirical record is solid. The debate should move on to new findings in the many areas where great uncertainty remains. That doesn’t mean I expect the 10 propositions to be immutable. On the contrary, I have had to keep in mind that *Human Diversity* is appearing in the midst of a rushing stream, reporting on a rapidly changing state of knowledge. Aspects of it are sure to be out of date by the time the book appears. My goal is to have been so cautious in my wording of the propositions that any outdated aspects of them will have been elaborated or made more precise, not overturned.

*How the Phrase Cognitive Repertoires Is Used Throughout the Rest of the Book*

The 10 propositions repeatedly refer to “characteristics of personality, abilities, or social behavior.” As I will occasionally put it, I am talking about the ways in which human beings differ above the neck (a loose way of putting it, but serviceably accurate).

I use *personality* and *social behavior* in their ordinary meanings. *Abilities* is a catch-all term that includes not only intellectual abilities but interpersonal skills and the clusters of qualities that have been described as emotional intelligence and grit. A good way of thinking about the universe of abilities is through Howard Gardner’s famous theory of multiple intelligences.\(^{(8)}\)

From now on I will usually abbreviate *personality, abilities, and social behavior* to *cognitive repertoires*. *Cognitive* means that it happens in the cranium or is at least mediated there. *Repertoires* refers to different ways of doing things that need not be ordered from “bad” at one extreme to “good” at the other. Some of them can be so ordered, but few have bad-to-good extremes. If you’re an employer, where do you want a job applicant to be on the continuum from “extremely passive” to “extremely aggressive”? It depends on whether you’re recruiting Navy SEALs or care providers at nursing homes, and in neither case is the most extreme position the ideal one. The same is

true even of something generally considered to be an unalyzed good, such as high IQ. Google may be looking for the highest possible visuospatial skills among its applicants for programmers, but the qualities that often accompany stratospheric visuospatial skills would make many of them dreadful choices as SEALs or care providers.

For most of the human qualities we will be discussing, “bad” and “good” don’t capture human differences. How many kinds of lovable are there? How many kinds of funny? How many kinds of annoying? Using the word *repertoires* allows for these kinds of apples and oranges too. So take note: For the rest of the book, *cognitive repertoires = characteristics of personality, abilities, and social behavior*.

As we embark on this survey of scientiﬁc discoveries about human diversity, a personal statement is warranted. To say that groups of people differ genetically in ways that bear on cognitive repertoires (as this book does) guarantees accusations that I am misusing science in the service of bigotry and oppression. Let me therefore state explicitly that I reject claims that groups of people, be they sexes or races or classes, can be ranked from superior to inferior. I reject claims that differences among groups have any relevance to human worth or dignity. The chapters to come make that clear.
PART I

“GENDER IS A SOCIAL CONSTRUCT”

From earliest recorded human history, everywhere and in all eras, women have borne the children and have been the primary caregivers. Everywhere and in all eras, men have dominated the positions of political, economic, and cultural power. From those two universal characteristics have flowed a cascade of secondary and tertiary distinctions in the status of men and women, many of which have nothing to do with their actual capabilities. In today’s language, gender has indeed been partly a social construct. Many of those distinctions were ruthlessly enforced.

The legal constraints on women in the modern West through the eighteenth century were not much short of de facto slavery. Mary Astell, often regarded as the first feminist (though she had precursors), made the point in response to John Locke’s cramped endorsement of women’s equality in the Second Treatise. She italicized phrases borrowed from Locke’s philosophical case for freedom: “If all men are born free, how is it that all women are born slaves? As they must be if the being subjected to the unconstant, uncertain, unknown, arbitrary will of men, be the perfect condition of slavery… And why is slavery so much condemned and strove against in one case, and so highly applauded and held so necessary and so sacred in another?”

If Astell’s language seems extreme, consider: An English woman at the time Astell wrote and for more than a century thereafter rarely got any formal education and had no access to university education, was prohibited from entering the professions, and lost control of any property she owned...
when she married. She was obliged to take the “honor and obey” marriage vow literally, with harsh penalties for failing short and only the slightest legal protections if the husband took her punishment into his own hands. Men were legally prohibited from actually killing their wives, but just about anything less than that was likely to be overlooked. When the first wave of feminism in the United States got its start at the Seneca Falls Convention of 1848, women were rebelling not against mere inequality, but against near-total legal subservience to men.

Under those conditions, first-wave feminists were too busy to say much about questions of inborn differences between men and women. An exception was Kate Austin, who compared the plight of women to those of Chinese women with bound feet: “We know that at birth the feet of the little baby girl were straight and beautiful like her brothers, but a cruel and artificial custom restrained the growth. Likewise it is just as foolish to assert that woman is mentally inferior to man, when it is plain to be seen her brain in a majority of cases receives the same treatment accorded the feet of Chinese girls.”43 As Helena Swanwick put it, “There does not seem much that can be profitably said about [the alleged inferiority of women]… until the incubus of brute force is removed.”44 Men joined in some of the strongest early statements on nature versus nurture. John Stuart Mill coauthored “The Subjection of Women” with his feminist wife, Harriet Taylor.6 George Bernard Shaw wrote, “If we have come to think that the nursery and the kitchen are the natural sphere of a woman, we have done so exactly as English children come to think that a cage is the natural sphere of a parrot—because they have never seen one anywhere else.”[7]

After the great legal battles of first-wave feminism had been won during the first two decades of the twentieth century, a new generation of feminists began to devote more attention to questions of nature versus nurture. The result was second-wave feminism, usually dated to the publication of Simone de Beauvoir’s Le Deuxième Sexe, a massive two-volume work published in 1949. Its argument sprawled across philosophy, history, sociology, economics, and psychology. The founding statement of second-wave feminism opened the second volume: “On ne naît pas femme: on le devient.” One is not born, but rather becomes, a woman.[8]

It was an assertion that required an explanation of how and why the change from birth to adulthood takes place. The intuitive explanation of “how” is that little girls are taught to be women—what is known now as socialization theory. It refers to the ways that children are exposed to influences that shape their gender identities. The pressure can come from parental interactions in infancy and toddlerhood, as girl babies are dressed differently from boy babies and female toddlers are given dolls to play with while boys are given trucks. The pressure may take the form of encouragement by parents, teachers, or playmates to engage in sex-typed play and discouragement of behaviors that go against type, as in the case of tomboy girls and effeminate boys. Parents may teach different lessons about right behavior, emphasizing the importance of being helpful and cooperative to daughters and the importance of standing up for themselves and taking the initiative to sons. Children may be encouraged to model themselves on the parent of their own sex. In these and many other ways, sometimes subtle or unconscious, children are constantly getting signals that track with the stereotypes of males and females.

This brief characterization of socialization theory skips over a number of intense scholarly debates between learning theorists and cognitive theorists, but the debates differ about the mechanisms at work. All agree on the basic tenet that girls are taught from infancy to be girls and boys are taught from infancy to be boys.9

Is socialization theory true? It’s natural to think so, if only because almost everybody can think of something during their childhood that involved references to what girls are supposed to be and what boys are supposed to be. Those of us who have had children of both sexes know that our interactions with our daughters and our sons have been somewhat different even if we tried hard to be gender-neutral in encouraging their abilities and ambitions.

But it’s one thing to have such personal experiences and another to demonstrate empirically that these differences in treatment as children produce the sex differences in personality, abilities, and social behavior that we observe in adult women and men. Little boys and little girls are treated differently, but how differently? "Several theoretical models suggest mechanisms that are consistent with the differential treatment of boys and girls," wrote four Dutch scholars of childhood socialization. "However, to date there is no consensus in the literature about the extent to which parents do treat their sons and daughters differently, in which areas of parenting this mostly occurs, and whether fathers and mothers differ in the extent of gender differentiation."[10] [Emphasis in the original.]
children, little girls need to be socialized into the personality traits and skills that facilitate nurturance. If society has come to depend on men being providers and leaders, little boys need to be socialized into the personality traits that facilitate acquiring resources and status.

Social role theory includes a role for biology. “Men and women selectively recruit hormones and other neurochemical processes for appropriate roles, in the context of their gender identities and others’ expectations for role performance,” Eagly and Wood write. “Testosterone is especially relevant when, due to personal identities and social expectancies, people experience social interactions as dominance contests. Oxytocin is relevant when, due to personal identities and social expectancies, people define social interactions as involving bonding and affiliation with close others.”

Biology interacts with psychology in two ways. Men and women alike psychologically internalize their gender roles as “self standards” for regulating their own behavior. They also regulate their behavior according to the expectations that others in the community have of them. “Biology thus works with psychology to facilitate role performance.”

The interdisciplinary sweep of social role theory means that it calls upon a wide variety of empirical observations about social roles across history and across cultures, evidence from psychology about internalization of norms, social psychological experiments, the nature of sex differences in personality, demographic trends, and economics, among many others. There is no equivalent to the meta-analyses of socialization studies that permits a short characterization of the state of knowledge about the validity of social role theory. But social role theory does what socialization theory does not: It provides a comprehensive explanation of why sex is a social construct.

But is sex exclusively a social construct? That the woman in a heterosexual couple does more housework than the man even when both have full-time jobs is at least largely a gender difference—the product of culture. It may have biological roots (perhaps men have evolved to be more tolerant of a messy living space than women are). But the issue is whether differential effort in doing the housework is sustained today by culture or genes. Think of it this way: How many women who can afford to hire someone to clean the house do so? A lot.

But simple quickly becomes complicated. Is the difference between the time men and women spend tending to young children artificially created by
culture or driven by inborn male-female differences? How about the attraction of girl toddlers to dolls and boy toddlers to trucks? Male-female differences in college majors? Male-female differences in attraction to casual sex? Are they sex differences or gender differences?

The sensible answer would seem to be "probably some of both," with arguments about how much of each. At one level, that's actually how the academic debate is conducted. The following chapters have hundreds of references to highly technical articles, adhering to normal standards of scientific rigor, published in refereed journals, arguing questions of nature and nurture, with male and female scholars making contributions on all sides on all topics. The tone is usually civil, and the conclusions are usually nuanced and caveated.

But the women and men who are engaged in this endeavor are a rarefied group of neuroscientists and quantitative social scientists. Few of them seek publicity (many do their work as unobtrusively as possible), and they do not set the mood on college campuses. Since American second-wave feminism took off in the 1960s, the most visible feminist academics have rejected the possibility that there are any significant sex differences from the neck up. In my terminology, they have denied that men and women have any inborn differences in cognitive repertoires. A person's gender "is an arbitrary, ever-changing socially constructed set of attributes that are culture-specific and culturally generated, beginning with the appearance of the external genitals at birth," in the words of one of the most widely read feminist scientists in women's studies courses, Ruth Bleier. It's not a position with a lot of nuance. Gender is a social construct. End of story.

The most famous illustration of what happens to those who question the orthodoxy is what befell economist Larry Summers. On January 14, 2005, Summers, then president of Harvard University, spoke to a conference on diversifying the science and engineering workforce. In his informal remarks, responding to the sponsors' encouragement to speculate, he offered reasons for thinking that innate differences in men and women might account for some of the underrepresentation of women in science and engineering. He spoke undogmatically and collegially, talking about possibilities, phrasing his speculations moderately. And all hell broke loose.

An MIT biologist, Nancy Hopkins, told reporters that she "felt I was going to be sick," that "my heart was pounding and my breath was shallow," and that she had to leave the room because otherwise "I would've either blacked out or thrown up." Within a few days, Summers had been excoriated by the chairperson of Harvard's sociology department, Mary C. Waters, and received a harshly critical letter from Harvard's committee on faculty recruiting. One hundred and twenty Harvard professors endorsed the letter. Some alumnies announced that they would suspend donations. Summers retracted his remarks, with, in journalist Stuart Taylor Jr.'s words, "groveling, Soviet-show-trial-style apologies." As to validate that image, Lizabeth Cohen, a Harvard history professor, told reporters after attending the Summers self-criticism session that "[h]e regrets what he said, and I hope that he will prove that by taking constructive steps. We're going to be in intense discussions with him over the next week."

Since 2005, expanding knowledge about male-female differences has substantiated Summers's speculations. The next five chapters review that evidence. The basics have been available to interested lay readers for years. And yet elite gender studies departments still refuse to acknowledge the biological side of gender differences. The degree to which the standard social science disciplines have also ignored this literature is an intellectual scandal. Evolutionary biologist Robert Trivers, whom you met in the introduction, has not held back:

Once you remove biology from human social life, what do you have? Words. Not even language, which of course is deeply biological, but words alone that then wield magical powers, capable of biasing your every thought, science itself reduced to one of many arbitrary systems of thought.

And what has been the upshot of this? Thirty-five wasted years and counting. Years wasted in not synthesizing social and physical anthropology. Strong people welcome new ideas and make them their own. Weak people run from new ideas, or so it seems, and then are driven into bizarre mind states, such as believing that words have the power to dominate reality, that social constructs such as gender are much stronger than the 300 million years of genetic evolution that went into producing the two sexes—whose facts in any case they remain resolutely ignorant of.
A Framework for Thinking About Sex Differences

A few decades from now, I expect we will have a widely accepted comprehensive theory of sex differences that is grounded in neuroscience, genetics, and evolutionary psychology. Progress has already been made in that regard, but it is still at the frontiers of scholarship and bears no resemblance to low-hanging fruit. In any case, my purposes don’t require that level of sophistication. A simple framework for thinking about phenotypic sex differences is supported by a growing number of scholars. This framework also links up with recent findings about sex differences in the brain.

A WORD ABOUT USAGE

From now on I will usually refer to "sex differences" instead of "gender differences." "Gender" was popularized in the 1960s to designate socially constructed differences. But it turns out that there is no clear division between biological and socially constructed differences and no point in trying to pretend otherwise—which is what the widespread use of "gender" amounts to. In the technical literature, many scholars who write on these topics have resumed the use of "sex" to apply to all kinds of differences between males and females. So do I.

The People-Things Dimension

More than a century ago, Edward Thorndike, one of the founders of educational psychology, asserted that the greatest cognitive difference between men and women is "in the relative strength of the interest in things and their mechanisms (stronger in men) and the interest in persons and their
effective therapists for the psychologically troubled—but, using the same
eurocognitive tools, they can also make effective arbitrators of disputes,
terrogators of criminal suspects, managers of people, or election-winning
politicians.

Other scholars of sex differences have been finding differences in ac-
demic interests, careers, and life choices that break along the lines of sys-
temizing and empathizing but that also lend themselves to the broader and
simpler difference that Thorndike identified—in choice after choice, men
are attracted to options that have more to do with things while women are
attracted to options that have more to do with people. That’s the simple
theory of the case I bring to the chapters on sex differences: Women and men
divide along the People-Things dimension.

Let there be any misunderstanding: I am talking about statistical ten-
dencies, not binary divisions. Many men and women possess trait profiles more
typical of the other sex. But these tendencies are strong enough to create
distinctively different distributions on important traits of personality, abili-
ties, and social behavior.

First Interlude: Interpreting How Big a Sex Difference Is

I warned you in “A Note on Presentation” that I would occasionally be inter-
rupting my narrative to explain technical terms. This is the first such inter-
lude. Some of you are already familiar with the term I will be explaining,
ext size, but I urge you to continue reading nonetheless. The interpretation of
effect sizes plays a significant role in how one interprets the evidence.

In the following chapters, I compare men and women on dozens of traits.
They are based on many kinds of measures—answers to questionnaire
items, scores on tests, and ratings of observed behavior, to name just a few.
Researchers need a common metric for expressing the differences that these
comparisons reveal.

To see what this metric must do, think in terms of a simple measure like
height. In one sense, an inch gives a common metric for measuring height.
You can express the height of anything with it. In another sense, it doesn’t
tell us much. For example, how big is a difference of six inches in height? In
absolute terms, it’s always the same. But how big is a six-inch difference if we
are talking about the height of elephants? The height of cats? The answer
depends on the average height of the things you are measuring and how much height varies among the things you are measuring. You need a way to express height in a way that means the same thing for elephants relative to other elephants and cats relative to other cats.

We need the same kind of metric to talk about sex differences across cognitive repertoires. That metric is based on a statistic called the standard deviation, described in detail in Appendix 1. In many cases, including the ones we will be dealing with, the standard deviation applies to a normal distribution, also known as a bell curve. To get from bell curves to effect sizes, let’s stick with the example of height.

The contemporary mean height of American women ages 20 or older is 63.6 inches. The comparable mean for men is 69.0 inches. Most people are clustered within a few inches of those means, but successively smaller numbers of people are three, four, five, and six inches from the mean. A tiny proportion of people are a foot or more from the mean. The nationally representative database of people that produced those numbers had these distributions:

![Diagram of bell curves for male and female distributions](image)

The dotted vertical lines show the means for women and men. The gray horizontal bar shows the difference between the two, which I call the “raw effect size.” Dividing it by the pooled standard deviations of the two groups gives us a way to express magnitude that can be compared across different traits.

An effect size is denoted as $d$. To calculate $d$ for height, I subtracted the male mean from the female mean, producing a difference of $-5.4$ inches. The pooled standard deviation is 2.9 inches, so $d$ equals $-5.4 / 2.9$, which works out to an effect size of $-1.86$. This is an extremely large effect size. Most sex differences are much smaller and the distributions have much more overlap.

Note that the sign of $d$ (negative or positive) is arbitrary. If I had subtracted the female mean from the male mean, the effect size wouldn’t have changed, but the sign would have been positive. Just so you know, in this book my default will be to subtract the male mean from the female mean in calculating sex differences. Therefore negative $d$ values will always indicate that males are higher than females on the trait in question, whether “higher” means something good, bad, or neutral.

Two questions are crucial to assessing the importance of sex differences: When is an effect size big enough to be interesting? Should individual effect sizes be treated individually or aggregated?

When Is an Effect Size Big Enough to Be Interesting?

Jacob Cohen, who originated Cohen’s $d$, inadvertently set the standard for interpreting effect sizes (he had a different purpose in mind). His list was subsequently expanded by Shlomo Sawilowsky. Under these guidelines, a $d$ value of 0.01 = very small, 0.20 = small, 0.50 = medium, 0.80 = large, 1.20 = very large, and 2.00 = huge.\[60\]

The guidelines were well-intended but have often proved to be pernicious in practice. As Cohen himself took pains to point out, the importance of a given value of Cohen’s $d$ depends on the specific topic you are examining. In 2019, psychologists David Funder and Daniel Ozer took on what they called the “nonsensical” standard set by Cohen, arguing that the interpretation of effect sizes should be guided by their consequences. In the case of a drug for curing a deadly disease that has a relatively small success rate, the effect of a success is a saved life—a consequence that can be important even if the effect size is small. In the case of a small effect size that has many repetitions, it’s the cumulative effect that’s important. For example, a study that tracked two million financial transactions found that the correlation between a person’s score on a measure of extraversion and the amount spent on holiday shopping is just +.09. “Multiply the effect identified with this correlation by the number of
heavily conditioned by the costs of making Type I and Type II errors for the particular question under consideration.”

Type I error refers to a false positive finding—in this case, wrongly concluding that a sex difference has been found. Type II error refers to a false negative finding—mistakenly concluding that no difference exists. Hyde was worried about the consequences of making a Type I error. She went on to give examples of the ways that inflating sex differences have real-world costs. For example, the idea that women are more nurturing than men backfires when it comes to the workplace: “Women who violate the stereotype of being nurturant and nice can be penalized in hiring and evaluations,” Hyde wrote, citing evidence to that effect.

On these issues, everyone who writes about sex differences should put their personal perspectives on the table. Regarding the use of Cohen’s guidelines, I think Hyde’s reliance on them to defend the gender similarities hypothesis is misplaced. There are too many ways in which effect sizes defined as “small” by Cohen’s guidelines can have important aggregate effects when thinking about sex differences. I appeal to the arguments made by the scholars I have cited, including Cohen himself, in defense of my position.

I also disagree with Hyde’s position that Type I errors should still be more feared than Type II errors. If we were back in 1960, I would agree with her—many people assumed that men and women were separated by large differences, and research that falsely reinforced that assumption could perpetuate harmful stereotypes, just as Hyde argues. But I’m writing at the end of the second decade of the twenty-first century when so many things, from high school athletic programs to the military’s composition of combat units, are guided by the assumption that there are no relevant sex differences. My guess is that the situation in 1960 has been reversed: More harms are now inflicted by incorrectly ignoring sex differences than by incorrectly exaggerating them. At the least, it can be said that there’s no clear case that Type I error is still more harmful than Type II error. This is an argument that does not lend itself to data-driven resolution. Differences in perspective are embedded in the literature on sex differences. It is well to be transparent about them.

Should Individual Effect Sizes Be Treated Individually or Aggregated?

My more important difference with Hyde involves her insistence on treating sex differences as independent bits and pieces rather than as profiles. When
are traits of personality, ability, and social behavior rightly treated independently? When should they be added up? These questions come up all the time in the social and behavioral sciences, and there are no cookbook recipes to go by.

To illustrate, let’s say we’re investigating personality differences and discover that people in Group A (the group could be based on any kind of common membership, not just sex) are somewhat more outgoing on average than people in Group B, with “somewhat” meaning that \( d = +0.35 \).

We get to know these groups better and determine that Group A is also somewhat warmer on average than Group B, with \( d = +0.35 \). Should we represent the two groups as separated by a mean personality difference of +0.35? Add the two effect sizes and say they are separated by a difference of +0.70? Or something in between?

I say that the answer is something close to +0.35. **Outgoing** and **warm** are nearly synonymous. The additional information hasn’t given us reason to think that the two groups of people are much more different than we already knew.

Suppose instead that we determine that Group A is also more emotionally stable than Group B, with \( d = +0.35 \). Should we continue to represent the two groups as separated by an average of +0.35? An aggregate of +0.70? Or something in between?

This time, I argue that the answer has to be closer to +0.70. We’re comparing people who are both warmer and more emotionally stable with people who are more aloof and easily upset. The personalities of the two groups are (on average) definitely more different than we knew before.

We continue to learn more about the two groups. We learn that one group is more prudent, the other more happy-go-lucky; one group is more practical, the other more imaginative; and so on. In some cases, the additional traits on which the groups differ are so closely related that the new knowledge adds only a small amount to the difference; in other cases, the new information adds a lot to the degree of their difference. But whether increments are small or large, my view is that individual differences that are conceptually related should routinely be aggregated.

Psychologist Marco Del Giudice, a leading advocate for aggregating sex differences in personality, uses an analogy with the distance between towns. If I tell you that one town is 35 miles west and 35 miles north of another town and ask you the Euclidean distance between the two, it wouldn’t occur to you to take the average of the two and announce that the towns were 35 miles apart. Similarly, it wouldn’t occur to you to add the two and say that the towns are 70 miles apart. You realize that we’re talking about a right triangle and that the hypotenuse is the distance between the two towns. You remember the Pythagorean theorem and know that the distance is therefore the square root of \( 35^2 + 35^2 \), which works out to about 49.5 miles. If I were to tell you that the altitude of the two towns differed by 4,000 feet, you would have to recalculate, taking the third dimension of height into account.

I like the analogy in part because the correct answer is so intuitively satisfying: We neither treat the three measures of distance separately nor simply combine the raw measures. Some method of aggregation that falls between averaging and simple addition seems right.

If you still want to average traits or treat them separately, my argument does not compel you to change your mind. I’ve made it through analogy and an appeal to intuition. But you should come to grips with how radical your solution is. If two indicators are involved, averaging cuts the simple sum of the two effect sizes by half. With three indicators, it cuts the simple sum by two-thirds. Suppose 10 indicators are involved. Averaging the results gives you an estimate of the sex difference that is just one-tenth of the estimate you would get by adding up the effect sizes. Doesn’t that seem like too much of a discount? This is a nontechnical way of saying that cognitive repertoires commonly involve multidimensional constructs, and the measure of male-female differences must be multidimensional as well.\(^{17}\)

In the same way that it is possible to compute the geographical distance separating two towns given two measures of their distance on the cardinal points of the compass, it is possible to compute distance in multidimensional space. The most widely used statistic for expressing multivariate distance is called Mahalanobis \( D \), named after the Indian statistician, Prasanta Mahalanobis, who developed it. The algorithm for calculating \( D \) does what I have argued intuition tells us it should, taking correlations into account. Suppose that variables have correlations near zero. \( D \) converges on the Euclidean distance. The higher the correlation between variables, the less \( D \) is augmented by including them. When a new variable is a linear combination of variables already in the equation, \( D \) is not augmented at all.\(^{18}\) The note also gives you
references disputing his position (one of them by Hyde) and Del Giudice's response to them.

In assessing the various arguments for and against, three points need to be kept in mind. First, Mahalanobis $D$ or any other method of aggregation must be used cautiously. In all complex statistical analyses, the validity of the results depends on interpreting the statistic with its limitations in mind.

But that leads to my second point: When I talk about indicators of sex differences being "conceptually related," I am not appealing to esoteric social science abstractions. To go back to my example, traits like warmth and emotional stability are characteristics with which we're all familiar from everyday life. We can effortlessly think of them as continua from coldly aloof to gushingly friendly; from rock-solid calm to emotionally volcanic. We've had experience with people who have different combinations of the two traits. In the same way, given normal standards of technical care in the application of multidimensional measures of distance and a clear narrative description of the logic for combining traits, aggregated measures of multidimensional distance can enhance our understanding of sex differences.

My third point is that in the real world it is taken for granted that small differences add up. Imagine a tennis match. You know that both players are professionals, but that's all you know. You have to bet on one of them. You learn that one player is 10 percent taller than the other. That doesn't give you much to go on; all you need is fractionally better than 1:1 odds to bet on the other guy. But suppose you then learn that the taller player also has 10 percent greater wingspan, 10 percent greater strength, 10 percent more endurance, 10 percent faster foot speed, 10 percent faster serve speed, 10 percent higher percentage of first serves, 10 percent faster reaction time, and 10 percent more emotional control. Now what kind of odds do you require to bet on the other guy?

I should add that my position makes virtually no practical difference to the discussions in the next four chapters. Almost all of the effect sizes I report are plain vanilla Cohen's $d$. I have given so much space to this topic because I think that treating effect sizes individually or averaging them has underestimated male-female differences. If you are unpersuaded, I will rest my case with the example of sex differences in the human face. Adult female and male faces are distinguished by dozens of tiny morphological differences. But they add up. Consider the following two faces:

Describing precisely why those two faces are so obviously a female's on the left and a male's on the right is daunting. The individual differences are almost imperceptible. But one thing is sure: To average out all those tiny individual differences and conclude that "male and female faces are virtually indistinguishable" would be ridiculous. The estimate of overall sex difference in faces must be expressed as some sort of aggregation. I submit that the same holds true for all sex differences comprised of functionally distinctive but conceptually related traits.