5

The Idea of a Peculiarly Female Intelligence: A Brief History of Bias Masked as Science

Gerd Gigerenzer

Her philosophy is not to reason, but to sense.
Immanuel Kant (1764)

Her logical thought is slower, but her associations quicker than those of man, she is less troubled by inconsistencies, and has less patience with the analysis involved in science and invention.
G. Stanley Hall (1904)

Immanuel Kant's conviction that women's nature is sense rather than reason surprised few scholars during the Enlightenment. Learned ladies, Kant believed, were worse than useless, and the very thought of women intellectuals interested in Greek philosophy or the foundations of mechanics seemed almost comical in his eyes (Kant, 1764/2011). Kant stood in a long and tenacious tradition convinced that the mind of a woman differs from that of a man. It can be traced back to Aristotle's

G. Gigerenzer (✉)
Max Planck Institute for Human Development, Berlin, Germany
e-mail: gigerenzer@mpib-berlin.mpg.de

R. J. Sternberg, D. D. Preiss (eds.), *Intelligence in Context*,
https://doi.org/10.1007/978-3-030-92798-1_5

influential contention that "the female is softer in disposition, is more mischievous, less simple, more impulsive, and more attentive to the nurture of the young; the male, on the other hand, is more spirited, more savage, more simple and less cunning … She is, furthermore, more prone to despondency and less hopeful than the man, more void of shame, more false of speech, more deceptive, and of more retentive memory" (Aristotle, 350 BCE/1984, pp. 948–949). At the beginning of the twentieth century, modern psychology reiterated the idea that women are qualitatively different. G. Stanley Hall, founder and first president of the American Psychological Association, held that women "excel in mental reproduction rather than production" (1904/1976, p. 565) and are intuitive and emotional, slow in logical thought, and too impatient for analysis and science:

She works by intuition and feeling; fear, anger, pity, love, and most of the emotions have a wider range and greater intensity. If she abandons her natural naiveté and takes up the burden of guiding and accounting for her life by consciousness, she is likely to lose more than she gains, according to the old saw that she who deliberates is lost (p. 561).

Hall, then president of Clark University, consequently opted against coeducation. Like Clark, Harvard set up a female institution in the 1890s, Radcliffe, next to all-male Harvard College. But even there, women were not treated like men. Not until 1967 did Harvard's Lamont Library open its doors to female students (Masters, 1986), an opening vehemently opposed by the administration and the majority of male undergraduates, on grounds that females would distract male students and that there weren't even bathroom facilities for women. And it took another ten years before Harvard terminated its policy to admit only one female student for every four male students.

Hall expressed what psychologists at the time held to be a fact of nature, traces of which can be found in people's thinking today. When my colleagues and I asked representative samples of twenty-first-century Germans and Spaniards about gender differences, the result was surprising—or perhaps not. The vast majority of women and men, young and old, believed that women had better intuitions than men about matters of personal affairs, but not of science and finance (Gigerenzer et al., 2014). And the rejection of learned ladies persists: Most contemporary

American men in search of a partner on online dating sites find well-educated women with a master's degree or a PhD unattractive and prefer those with lower education (Bruch & Newman, 2018).

This chapter is a case study on how lack of theory about the nature of "intelligence" enabled cultural biases about women to be presented as science by major psychologists. A discipline that is unaware of the errors in its history is potentially hazardous: "Those who cannot remember the past are condemned to repeat it" (Santayana, 1905). I reconstruct the history of the idea of a peculiarly female intelligence in three overlapping views. In the first view, from Aristotle through to the mid-nineteenth century, the idea of intelligence as we encounter it today—as a general ability that is measurable and is largely independent of personality and moral character—did not exist. Instead, the difference between men and women was understood in terms of polarities that were a mixture of intellect, personality, and moral character, such as men's abstract versus women's concrete thought. The notion of these polarities wore away in the mid-nineteenth century and was supplanted by the concept of an inherited "natural ability" (soon to be named *intelligence*), mainly through the writings of the English polymath Francis Galton. As a consequence, in this second view, men and women differed no longer in quality but in quantity: On average, it was thought, women had inherited a smaller share of intelligence. The psychologist Louis Terman put an end to this view by eliminating particular test items from his Stanford-Binet test and balancing the rest so that girls and boys had the same mean IQ. The eliminated items landed in a personality scale called *masculinity-femininity*, which illustrates the arbitrariness of what counted as a measurement of intelligence rather than of personality. What remains debated to the present day is the third view, promoted by sexologist Havelock Ellis. It alleges that men's intelligence varies more than that of women, implying the existence of more male idiots and geniuses.

Whatever the hallmark of a peculiarly female intelligence has been—polarities, lower average, or lower variability—it has served the dubious purpose of justifying men's superior role in society. Similarly, whatever the supposed mental differences were, these became presented as part of the natural order, expressed in the female body and women's reproductive function (Daston, 1992).

Before Intelligence: Male–Female Polarities

Intelligence, as we know it from IQ tests, refers to a general ability that can be measured by a single number and is assumed to be largely independent of personality and moral character. IQ tests have been given to millions of children, recruits, and job applicants, and continue to influence access to education and jobs. The IQ has often been presented as a hard fact, and debates raged over how much of its variability is due to nature and nurture. These debates ignored the fact that intelligence as we know it was "invented" in the late nineteenth and early twentieth century.

Before that time, what we now call the intellect was considered neither a single general ability nor as largely unrelated to moral and personality traits. Rather, psychological theories conceived of the mind as a collection of faculties or talents. For instance, the key concept of *sensibility* in early eighteenth-century psychology encompassed both perceptual and emotional sensitivity as the precondition for empirical knowledge and the emotions of charity and compassion (Rifkin, 2002). Reason was even more closely identified with morality because the light of reason enabled one to recognize all forms of truth, including the distinction between good and evil. No single one of these faculties or a combination thereof corresponds to the contemporary concept of intelligence (Daston, 1992).

The prototypical male and female occupied opposite poles on the spectrum of these faculties. For instance, men were characterized by abstract thought, judgment, and genius, while women were considered to lack these and instead excel in concrete thought, imagination, and retentive memory. Male strength was opposed to female delicacy or bodily and mental weakness. This supposed weakness was in turn seen as evidence that nature intended women to confine themselves to the home and subordinate themselves to men. It was reasoned that because men's thought was abstract, they could comprehend truth, including moral truth, while women's concrete thinking prevented them from grasping abstract moral principles. Hence, women who lied or stole were considered incapable of understanding that their actions were evil. When Hall, in 1904, wrote that women were unfit for science and invention because they lacked patience, he was simply reiterating the timeworn conviction that women

did not have the necessary self-discipline and stamina to reason by following a lengthy chain of argument.

Women's and men's virtues were also seen as diametrically opposed. For centuries (and in many contexts even today), chastity was considered the chief female virtue, and its violation a cardinal sin for women alone. Timidity, in contrast, was a cardinal sin for men, but easily excused in women (Daston, 1992). The view that women's intellect, character, and moral traits are intimately connected to their biology survived in various forms into early twentieth-century philosophy. Consider the controversial Austrian philosopher Otto Weininger, hailed by Freud and Wittgenstein as a great genius (Dury, 1984). In his book *Sex & Character* (1903, translated into English in 1906), Weininger drew on a wide range of philosophers and psychologists to assert that reasoning and feeling are equivalent in women, who as a consequence are prone to suggestibility, hypnosis, and hysteria, as documented by Freud. These alleged flaws correspond to Aristotle's view that women's memory is easier to imprint. From biologists Geddes and Thomson (1890), Weininger borrowed the conviction that each cell in a woman's body is sexually marked to make the female in every respect passive, submissive, and lacking in personality. Unlike man, he wrote, "woman is non-logical and non-moral" (p. 297). Faced with the fact that more men stand trial for crimes, he argued that behind every lawbreaker there is a woman who proposes the crime and profits from it. Weininger gained great popularity when he killed himself at the age of 23 at a spectacular site, the room in which Ludwig van Beethoven had died. This dramatic finale led to huge book sales and an enthusiastic reception by many contemporaries, including the Swedish playwright and novelist August Strindberg, who claimed that Weininger's book had finally solved "the problem of women" (Abrahamsen, 1946).

In sum, for millennia, a fairly consistent view reigned about women's intellect as differing fundamentally from that of men. My brief account does scant justice to the variations of this view among scholars and centuries. Yet the common denominator between them is that there was no concept of a general intelligence, which was instead defined by a number of diametrically opposed polarities attributed to the prototypical male and female, a combination of what were later separated into intelligence, personality, and moral traits.

The Invention of General Inherited Intelligence

The idea of mental faculties was slowly abandoned in the mid-nineteenth century for that of a single overarching intelligence. In contrast, the associated idea that this intelligence combines cognitive abilities, personality, and moral traits faded away only in the early twentieth century. The transition from multiple mental faculties to a single intelligence was driven not by data or experiment but by concerns outside the realm of science, chief among them Francis Galton's interpretation of evolutionary theory, his fascination with measurement, and his involvement with the fateful eugenics program.

Women Are Granted the Same Kind of Intelligence as Men, But Less of It

Galton, a cousin of Charles Darwin, promoted a strict distinction between nature and nurture, which had not been considered mutually exclusive before his time (Daston, 1992). That artificial distinction later led to a flood of psychological research trying to find an answer to the (wrong) question of what percentage of the variation in intelligence is due to nature and nurture (as opposed to asking how genes and environment interact, as in epigenetics). For Darwin's theory of evolution to work, it was clear that something must be passed on to the next generation and inherited by both boys and girls. In *Hereditary Genius* (1869/1979), Galton called this something *natural ability* (later known as *intelligence*). As he saw it, evolution implied that men and women must have the same kind of natural ability and also that this ability shows variability between individuals, given that variation is a driver of evolution. Men and women were assumed (no measurements or tests were involved) to exhibit the same bell-shaped ("normal") distribution of intelligence, an assumption Galton justified by analogy with height. Using the same analogy, he assumed the female distribution to have a lower average. Consequently, in *Hereditary Genius*, women feature solely as the mothers or wives of male geniuses.

Galton maintained the view that natural ability is a combination of intellect, personality, and moral traits, such as capacity, zeal, and the power to do laborious work. With respect to morals, he wrote that it is the nature of all of us to believe blindly in what we love, rather than in what we think most wise. "We are indignant when others pry into our idols, and criticize them with impunity, just as a savage flies to arms when a missionary picks his fetish to pieces. Women are far more strongly influenced by these feelings than men; they are blinder partisans and more servile followers of custom" (p. 196).

The invention of a single form of intelligence, or natural ability, allowed Galton and his followers to compare men and women on a single dimension, similar to how he compared humans of different racial categories and even animal species. For instance, he conjectured that the "negro race" differed from the Anglo-Saxon in their lower mean (p. 338), not in the nature of their intelligence, and that certain gifted dogs had superior intelligence to some human "idiots and imbeciles" (Galton 1869/1979, p. 36).

Today, the idea of single kind of intelligence is mostly related to Charles Spearman's (1904) "g" factor. In fact, Spearman was strongly influenced by Galton, and his main statistical tool was correlation, developed by Galton. Like Galton, he thought that high sensory discrimination and high intelligence are part of the same universal intellectual function. Unlike Galton, however, Spearman (1904) steered clear of prejudices about women or nonwhites being genetically inferior in their intelligence.

The Failure to Measure Intelligence

After Galton had invented the concept of general intelligence, he tried to measure it in his Anthropometric Laboratory in London, which opened in 1884. He started with the hypothesis that intelligence, being inherited, can be found in mind and body—in the entire nervous system. Therefore, greater sensory acuity would be the external sign of higher intelligence. Inspired by Galton, James McKeen Cattell established another anthropometric laboratory in Cambridge University, which also focused on sensory acuity. However, Clark Wissler (1901), a student of

Cattell's, could not find a clear relationship between sensory acuity and mental ability when looking at college freshmen's grades. Moreover, the various acuity measures did not appear to correlate with each other (see Blum, 1978; Sternberg, 1990). Rather than acknowledging this failure as an invalidation of his hereditary theory of intelligence, Galton assumed a need for better measures of innate ability. His search failed.

The key to measuring intelligence was found later in the work of Alfred Binet and Théodore Simon in France. In contrast to Galton and his followers, however, neither Binet nor Simon conceived of intelligence as fixed or inherited, and Simon protested against the misuse of their test in England and the US for measuring an allegedly inherited ability (Wolf, 1973).

How Women's and Men's Average Intelligence Were "Made Equal"

Binet, a member—and, later, director—of the French Ministerial Commission of Abnormal Children, was concerned about the unreliable diagnoses of children with intellectual disabilities in France. One and the same child might be classified as "imbecile," "idiot," "feeble-minded," or "degenerate" in different certificates (Binet & Simon, 1916/1973). Binet set out to classify these children in an objective way with scientific precision. His goal was to place children with intellectual disabilities in special schools geared to improve their abilities, as in the German school system at the time, and also to ensure that children without any intellectual disabilities would not be placed in special classrooms solely because they were behaviorally challenging. But Binet had no coherent idea how to measure intelligence. Like Galton, he searched in vain for correlations with sensory acuity and tried almost everything else that seemed viable, including assessing intelligence on the basis of facial features (physiognomy), measurements of the head (cephalometry), and handwriting (graphology). For instance, he presented handwriting samples from convicted murderers mixed with those from normal citizens and asked expert graphologists for character assessments, only to find out that even the most

eminent experts arrived at disastrously false assessments (Wolf, 1973). The results were consistently disappointing. It remained a mystery what intelligence was, or how to measure it.

Eventually, however, Binet and Simon found an ingenious answer to the question of finding a test that correlated with teachers' assessments. They developed questions about subjects that mirrored what was taught at school, such as reasoning skills, knowledge, memory, and attention. Children's answers to these questions now correlated with their school grades as well as with teachers' evaluations. By 1905, Binet and Simon had their first test of intelligence for classifying intellectually challenged children into several levels of developmental delay; in 1908, the test was revised and called a test of the "development of intelligence among children." Note that the test was intended to sort children into categories, not to assign them a single number such as an IQ. It was also not intended to measure innate intelligence, but to replace teachers' and physicians' unreliable diagnoses of children with intellectual disabilities, as a "means of prophylaxis, a means of escaping conscious and unconscious error" (Binet & Simon, 1914, p. 10).

Binet and Simon's test questions still reflected the meaning of intelligence as a combination of intellect, character, and moral traits. For instance, the test included questions such as the following: "If you are late for school, what would you do?" and "Why should one judge a person by his acts rather than by his words?" Today, one might call this social intelligence, but Binet and Simon thought of social judgment as inseparable from intelligence. Now they had a test, but without a theory of intelligence, apart from a loose definition of intelligence as "judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting one's self to circumstances. To judge well, to comprehend well, to reason well, these are the essential activities of intelligence" (Binet & Simon, 1916, pp. 42–43). Before his death, Binet (1911) wrote: "Thus we return to our favorite theory: intelligence is marked by the best possible adaptation of the individual to his environment" and "to this we really do not want to add another thing" (p. 172). To which his biographer Theta H. Wolf added: "How strikingly inept is such a pronouncement if we think of the excellent 'adaptation' to their environment of mice and moose!" (1973, p. 210). Measuring without precisely knowing what one

is measuring has been, and still is, one of the striking features of research on intelligence. And this feature conveniently allowed researchers to adjust the facts about female intelligence.

Binet's Intelligence Test Crosses the Atlantic and Becomes Seen as a Test of Genetic Ability

After getting his PhD from G. Stanley Hall at Clark University, Lewis Terman joined the faculty at Stanford University and became known as *the* leading U.S. researcher on intelligence. Terman was more interested in gifted children than in intellectually challenged ones. In line with Hall and Galton, he firmly believed that intelligence was inherited. He translated Binet and Simon's test into English, added and deleted some questions, and published the product in 1916, which became known as the *Stanford-Binet Intelligence Scale*s.

Yet Terman had made important alterations that went largely unnoticed in the US but were to have damaging implications. He named the test an *IQ test* (the term was originally introduced by the German psychologist William Stern), where IQ was the ratio between mental age and chronological age. He believed that whatever the test measured was fixed and inherited, or at least predominantly so. Whereas Binet and Simon thought of the test as a means to send children with intellectual disabilities to special schools so that they could ideally be channeled back into normal classrooms, Terman instead advocated special institutions and sterilization of the "mentally retarded" (Minton, 1988, p. 149). Terman had a strongly biased vision of what would happen once his test was widely applied: "There will be discovered enormously significant racial differences in general intelligence, differences which cannot be wiped out by any scheme of mental culture" (Terman, 1916, p. 92).

Under the leadership of Robert Yerkes, president of the American Psychological Association and a member of the Eugenics Record Office's Committee on the Inheritance of Mental Traits, the *Army Alpha and Beta Tests*, based on Terman's IQ test, were applied to 1.75 million men in World War I (Carson, 2007). Yerkes and his staff were convinced that the test measured native intelligence, even though it included items such as

"The Overland car is made in Buffalo/Detroit/Flint/Toledo" (Minton, 1988, p. 70). They recommended immediately discharging about 8900 men with low test results from service, many of whom were foreign-born or illiterate. The army officers disagreed with the psychologists, pointing out that these men would become good soldiers after training (Minton, 1988, p. 73). Nevertheless, Yerkes hailed the test a great success, despite little evidence that it had made recruiting more efficient or had contributed to winning the war. On the contrary, the war helped to win publicity for mass testing—if only because the psychologists had shown that such testing could be accomplished. On that wobbly basis, IQ testing spread across the US.

Binet, who died in 1911, did not live to see what happened with the Binet-Simon test once it crossed the Atlantic, but Simon did. He objected to the term *IQ* because it suggested a fixed, inherited mental age. In interviews with Binet's biographer Theta Wolf, Simon even called the term and its genetic interpretation a betrayal ("trahison") of their test's original objective (Wolf, 1973, p. 203).

Men and Women Are Assigned the Same Mean Intelligence

Without much fanfare, Terman eradicated the idea that females have lower average intelligence. In his revised Stanford-Binet test, he deleted questions for which boys and girls had different success rates and balanced the rest so that, on average, girls ended up with the same IQ as boys. Terman was not particularly explicit about this correction, nor about its reasons. But his decision finally made women equal to men in terms of IQ, at least on average.

What was Terman's motivation? Terman and Merrill (1937) explained that they plotted the difficulties of each item against age groups "for the sexes separately as a basis for eliminating tests which were relatively less 'fair' to one sex than the other" (p. 22). Moreover, "a considerable number of those retained show statistically significant differences in the percentages of success for boys and girls, but as the scales are constructed these differences largely cancel out" (p. 34). The explanation of "fairness"

appears strange in face of Terman's intention to measure largely genetic differences in intelligence. And fair to whom? Were boys or girls originally better, and whose mean was upgraded? Terman and Merrill did not say.

Others proposed that Terman made the means equal to reckon with the fact that girls usually perform better in school, or in response to pressure generated by the increasing women's movement of the period (Blum, 1978). A third explanation is that Terman, working closely with a large number of women coworkers (according to his biographer, Henry Minton, 1988, sometimes too closely), was influenced by them to make the averages equal. Yet all three explanations assume that boys tested better than girls, and that item deletion served to upgrade the girls' average. Who really did perform better in the original set of test, girls or boys?

It took me a while to find an answer in Terman's writings. It appeared years later, in a different context, in the study on gifted children by Terman and Oden (1947), hidden in a side remark on another topic, the question of why there were more boys than girls in the group of gifted children. Terman and Oden discussed the possibility of a nomination bias (teachers nominate more boys than equally gifted girls), and also the possibility of "a real average superiority of boys in the intellectual function tested" (p. 13). They concluded that such a real average superiority is unlikely because for the 905 subjects on whom the 1916 Stanford-Binet was standardized, the mean IQ was slightly higher in girls. In other words, Terman appears to have found that girls had higher average scores in his intelligence test than boys, and then deleted items and balanced others to lower the mean of the girls to match the inferior mean of the boys!

One might ask what would have happened if girls had had the lower scores. Would Terman also have deleted items to make the averages equal? If not, the test might have been standardized such that females' average IQ was a few points lower than males'.

Terman's decision to make the average IQ of males and females equal put an end to the second idea of a peculiarly female intelligence. It also illustrates the deep problem of how to measure something in the absence of a theory, where there is wiggle room to make decisions about test items that produce the result one favors—for fairness or whatever other

reasons. In principle, Terman could have designed a test in which women are superior to men, or where certain cultures or races are superior to white Americans. The problem is this: One can measure whether women and men differ in a specific and clearly defined task, such as memory span. But if one has neither a clearly defined task nor a theory and instead selects dozens of test items and adds the points up to determine an IQ, there are many degrees of freedom that allow for tinkering with the test to fit its result with preconceived beliefs and biases.

This key problem of measuring IQ is not always acknowledged. Consider Hans-Jürgen Eysenck, who once was the most frequently cited living psychologist and one of the most controversial intelligence researchers. In his *Intelligence Controversy* with Leon Kamin (Eysenck & Kamin, 1981), he reified the equal averages, complaining that psychologists "are said to have selected items in such a way that equal scores are achieved regardless of whether there might or might not be genuine differences between the sexes. This accusation is false" (p. 40). He continued: "Given that unselected items give the sexes equal IQ scores, it was only reasonable for other test designers to avoid bias in favour of one or the other sex" (p. 41). However, there is no such thing as "unselected" items in the absence of a theory of what intelligence is and how it can be measured. Terman himself occasionally also reified the equality of mean IQ to support women's equality. In *Sex and Personality* (Terman & Miles, 1936), Terman and Catherine Cox Miles wrote: "Intelligence tests, for example, have demonstrated for all time the falsity of the once widespread prevalent belief that women as a class are appreciably or at all inferior to men in the major aspects of intellect" (p. 1). All in all, Terman's IQ test ended the view that females have lower average intelligence than males so that men and women were finally seen as equally intelligent—at the expense of favoring racial prejudice.

How Differences in Intelligence Became Differences in Personality

In the introduction to *Sex and Personality*, Terman and Miles (1936) noted that it appears impossible to explain sex differences in behavior

wholly in terms of biological factors and complained that the concepts of masculinity and femininity are even more vague than the nineteenth-century concepts of intelligence (pp. v–vi). As an example, they referred to the stereotype of the "occidental" woman whose moral life is shaped less by principles than by personal relationships, and whose everyday behavior is more determined by emotion, submissiveness, and inferior steadfastness of purpose.

Nevertheless, Terman and Miles did not present a theory that replaced the vagueness and stereotypes to which they objected. How then could they measure personality differences between men and women? Terman and Miles came up with an ingenious solution, which was initiated as subtly as Terman's strategy to discard test questions had been. It turns out that the discarded questions ended up in their "masculinity-femininity scale" (Terman & Miles, 1936). That action guaranteed differences between males and females on the new scale, which contained, among others, questions on interests such as movies and amusement, opinions such as "The unmarried mother deserves the scorn she gets" and "Blondes are less trustworthy than brunettes," and "information" such as "The most gold is produced in Alaska/NY/Tennessee/Texas." Once seen as items that measured inherited intelligence, these now served to measure personality and gender-specific knowledge. In the absence of a theory of intelligence that determines what questions are relevant, one-and-the-same item can be applied to measure sex differences in intelligence or in personality. In various forms, the masculinity-femininity scale is still in use and still presented as measuring sex differences in personality.

Larger Variability in IQ Justifies Male Superiority

In 2006, Harvard President Larry Summers resigned from his position in the wake of a no-confidence vote by his faculty. Among the reasons cited by the faculty was a remark he had made regarding women's intelligence and ability. On the question of women's aptitude for science, Summers said: "It does appear that on many, many different human

attributes—height, weight, propensity for criminality, overall IQ, mathematical ability, scientific ability—there is relatively clear evidence that whatever the difference in means—which can be debated—there is a difference in the standard deviation, and variability of a male and a female population" (2005). From that he drew the conclusion that the greater variability of males explains why top universities such as Harvard hired relatively few women as professors.

Summers's statement simply repeated a hypothesis discussed in psychological research for over a century: that the variability of women's physical and mental traits, including IQ, is smaller than that of men. This *variability hypothesis* both explains and justifies observations that there are more male geniuses than female ones and also explains why that there are more male idiots at the other end of the IQ distribution.

After Galton replaced the first version of intelligence—that men's and women's mental abilities were at opposite poles—with one common intelligence, and Terman in turn put an end to the subsequent idea of average differences, the only possible remaining difference on the bell curve was the variability, or standard deviation, in IQ. After all, a bell curve has only two parameters, mean and standard deviation. The variability hypothesis became the third and last bastion for the idea of a specifically female intelligence, contributing to Summers's fall. Its origins seem to be in an observation by Darwin in the second edition of *The Variation of Animals and Plants Under Domestication* (1875, p. 457) that male animals tend to be more variable than females, although Darwin himself devoted little attention to this issue. Instead, the claim of greater male variability was promoted by the English sexologist Havelock Ellis (1859–1939).

The Variability Hypothesis

Ellis rebelled against the conspiracy of silence surrounding the sexes and decided to devote his life to their scientific study. For him, women and men were different but complementary—in contrast to Galton, who did not see much usefulness in women's lower average natural ability. In the first edition of *Man and Woman* (1894, p. 367), Ellis wrote: "From an

organic standpoint, therefore, men represent the more variable and the more progressive element, woman the more stable and conservative element, in evolution. It is a metaphorical as well as a literal truth that the center of gravity is lower in women and less easily disturbed." (In the fourth and fifth editions, Ellis left out the "progressive element," indicating second thoughts about the generalizability of biological variation, particularly to politics.) He wrote that women's smaller stature approximated that of humans' ancestors, and that women—as in witches and soothsayers—preserved ancient custom and methods of thought. Women had "an organic tendency to stability and conservatism, involving a diminished individualism and variability" (p. 369). As an example, he made the case that women had opposed the French Revolution, albeit also noting that the revolutionary movement of Christianity was to a considerable extent furthered by women (p. 370). He acknowledged that the facts are very complex and that that the claim of absolute inferiority for either sex is untenable, but nonetheless concluded: "It is undeniably true that the greater variational tendency of the male is a psychic as well as a physical fact" (p. 370).

Man and Woman received scant attention when it first appeared (Grosskurth, 1980, p. 170). Yet that changed when the statistician Karl Pearson (1897) vigorously attacked Ellis's variability hypothesis. Pearson was a committed socialist and promoted feminism and eugenics, both of which were considered progressive and revolutionary at the time. Pearson argued that the claim of greater male variability contradicts Darwin's theory of evolution by natural selection, which emphasizes variability as one of the driving forces of evolution but postulates that the more intense the struggle, the less is the variability. Therefore, he expected men, not women, to be less variable. Next, he criticized Ellis's inconclusive evidence, based almost entirely on pathological variation such as criminality and color blindness. And, finally, Pearson contended that measuring the variability of absolute variables such as the length of bones (as opposed to ratios such as cephalic index) by the standard deviation, as Ellis did, was an error. Instead, one needed to calculate the coefficient of variation, that is, the standard deviation divided by the mean. After all, women's bodies were smaller than men's and so, therefore, was the standard deviation of bodily measures. Pearson concluded from his own physical measures that

the coefficient of variation is slightly larger for women, not smaller, reflecting their "slightly less severe struggle for existence" (1897, p. 297).

In an Appendix in *Man and Woman*, Ellis rejected Pearson's "hostile" criticism at length, which Pearson did not deem worthy of a response. Pearson's sole reaction was a footnote in an article unrelated to variability, in which he noted that Ellis's response required no reply, as Ellis did not appear to understand that scientific evidence, not vague generalities, was what counted (Pearson & Lee, 1903, p. 372). Afterward, Pearson did not pursue the variability hypothesis any further.

Why did this bitter controversy over females' allegedly lower variability erupt? According to Ellis's biographer Phyllis Grosskurth (1979), one likely reason was personal resentment. Many women of the time found Ellis, who with his flowing beard resembled "a combination of archetypal Father and sensual Faun," irresistibly attractive (p. xvi). The South African writer Olive Schneider was one of the women upon whom Ellis had a strong influence, before she fell in love with Karl Pearson. Whatever its motivation, Pearson's critique of the variability hypothesis in fact contributed to making the hypothesis popular.

Ever since, psychologists, biologists, and statisticians have debated the variability hypothesis. Whereas Ellis and Pearson related it to both physical and mental traits, psychologists have focused largely on intelligence. McNemar and Terman (1936) reported greater variability in boys on the Stanford-Binet and other tests but, given the inconsistent evidence, were careful not to draw any general conclusions. In 1932, Scotland undertook the ambitious project of testing all 11-year-old Scottish children with the goal of discovering the amount of mental deficiency in the country (Scottish Council for Research in Education, 1933). Because suppliers demanded too much money for the nearly 100,000 commercial tests, the Council used the *Morey House Test* in place of the Stanford-Binet. The conclusion was that boys and girls did not differ in average IQ but that the standard deviation of boys was one IQ point larger than that of girls. In 1947, the same project was repeated with all 11-year-olds at that time, and again the standard deviation was one point larger for boys (Scottish Council for Research in Education, 1949). That appeared to support the hypothesis of both higher and lower male intelligence. Although this result was hailed as the most comprehensive demonstration of the greater

variability of mental ability among males (Deary et al., 2009, p. 185), the small difference in variability in the 1947 study was mainly due to an excess of males with very low scores, not to male genius (Deary et al., 2009, pp. 21, 184). The primary impetus of the 1947 study (and that of 1932) was not variability, but rather the concern that the nation's intelligence would decline because people with lower mental ability tended to have more children. Yet the children scored no worse than those studied 15 years earlier; in fact, their average IQ went up by about one point in boys and three points in girls.

Follow-ups of the Scottish children have shown similarly inconclusive results. In 1939, the Council found no significant difference in variability between boys and girls; in 1949, it reported slightly larger standard deviations in boys; and in 1958, it reported a greater proportion of females than males at the lower end of the IQ scale. Thus, one could find support for or against the variability hypothesis in intelligence, depending on the age group and study. More fundamentally, findings about variability—like mean differences—always depend on how the test items are selected and weighted. Just as Terman made the means between males and females equal, one can select items to make the variability equal.

Outspoken advocates have presented greater male variability as a biological fact, possibly due to sex linkage, speculating that intelligence might be located on the X chromosome. According to this line of reasoning, intelligence in males can express itself without interference of a second X chromosome, thereby causing greater variability in IQ (Johnson et al., 2009; Lehrke, 1978). This ignores the fact that the same hypothesis could be likewise used to predict that females have higher average intelligence than men, given their two X chromosomes, once again illustrating the utter arbitrariness of genetic explanations in the absence of a theory. Whereas the first two ideas about a peculiarly female intelligence had been conceived and debated virtually entirely by men, the variability hypothesis was challenged by an early generation of women scientists (Shields, 1982). Helen Bradford Thompson (1903) conducted her own studies and criticized Ellis's conclusions. Her critique of the variability hypothesis was widely read yet had no equivalent impact. In the most systematic critique of the variability hypothesis at the time, Leta Setter Hollingworth (1914) reported no evidence of it in her review of the

literature. Beth Wellman (1933) found in her review slight support for greater variability in boys, which, however, depended on the measure of variability used, the selection of children, and other details. The variability hypothesis remains a matter of discussion. In her review of the state of art in sex differences in cognitive abilities, Halpern (2012, p. 103) concludes "that females and males are very similar when we consider the average performance, and they are highly dissimilar when we consider performance at the high and low extremes."

As with the question of whether males and females differ in their average IQ, the absence of a theoretical understanding of what a test actually measures opens the door to including or excluding items that make the mean and variance of IQ equal or different.

Lessons Learned

The idea of a peculiarly female intelligence emerged in three different and unrelated versions: male-female polarities, female lower mean intelligence, and female lower variability. The idea that men and women occupy opposite poles on a continuum, such as analytic versus intuitive, reigned for millennia. It began to fade away when Francis Galton invented intelligence (natural ability) as a single dimension, which later morphed into IQ or g (general intelligence), so that the mind of men and women now had the same quality, but with women having less of it. The idea that women have lower intelligence expired in the hands of Louis Terman, who eliminated test items so that both males and females had the same average IQ—otherwise, female means would in fact have been higher. The third idea was that while the means are the same, woman's variability is smaller, resulting in more male geniuses and idiots. This variability hypothesis is still debated today.

Despite the differences in these three ideas about a peculiarly female intelligence, their justifications are strikingly similar, and the supposed nature of women features prominently in all three. Woman's mind was said to be determined by her reproductive biology, her body, her genes, and her naturally ordained functions. The first president of the American Psychological Association, G. Stanley Hall, staunchly believed that the

female mind was created for nursing and motherhood, serving the production of men of genius and of daughters to bear future male geniuses (Diehl, 1986). Education, he felt, would damage women's reproductive organs, particularly coeducation in competition with men. Like many others at the time, Hall did not think of women as generally inferior but instead idealized them. In his view, women who entered men's world of education and business became innocent victims of man's evil nature, losing their purity and sainthood (Schofer, 1976).

The historian of psychology, Edwin Boring, famously said that intelligence is whatever the IQ tests measure. But that is precisely the problem. The idea of a peculiarly female intelligence is a striking case of measurement without understanding what one is measuring, paired with the hope that sophisticated correlation statistics and factor analyses could fill this theoretical void. From Galton to Binet to Terman, researchers variously believed that one could measure intelligence in terms of sensory acuity, head size, facial features, handwriting, memory capacity, or knowledge of facts, or by asking questions about proper social behavior.

This absence of theory left too many points of entry for biases and preset convictions, to the detriment of many. Galton's vision was to promote the eugenics program: to detect the less-well-endowed and prevent them from reproducing. Both Ellis and Pearson were early feminists but also proponents of eugenics, both of which were considered progressive movements at the time. Binet and Simon intended to give children with intellectual disabilities a second chance through special education. Yet when adapted "to American conditions and needs," as the editor's introduction to the 1916 edition of Terman's *The Measurement of Intelligence* put it, their test came to serve the various goals of eugenics, sterilization, racism, feminism, and, last but not least, a multibillion testing industry.

Why Is History Relevant?

Knowing one's history provides an opportunity to learn from errors and avoid repeating these. Differences between men and women, as well as their causes, have been an emotionally and politically charged topic for centuries. Firm convictions continue to be enforced in the guise of new

technology. For instance, Diane Halpern warns that modern neuroscience is being misused to justify sex role stereotypes in how men and women think, a program dubbed "neurosexism" (Halpern, 2012, p. xi). Basing conclusions about human thinking and behavior on the firing of neurons or changes in blood oxygen levels entails a long leap in logic. We have seen such leaps before, as with the argument that the smaller brain of females is responsible for woman's alleged intellectual inferiority. The stereotypes of the past also tenaciously survive in popular psychology bestsellers that present men and women as if they were alien species, as in *Men Are from Mars, Women Are from Venus* (Gray, 1992). In a throwback to the view of women being submissive by nature, such books imply that a wife's role is to hide her intelligence, to admire and appreciate her husband, and not to offer him advice unless he asks.

What is the current consensus about differences in cognitive abilities between men and women? According to Halpern (2012; Halpern & Wai, 2020), the list of differences is relatively small, and the similarities between the sexes are larger in number. Few of the differences that have been claimed over the years are stable across age, task, and culture. Among the few exceptions are that women have better memories than men (p. 119) and excel in reading and verbal abilities, while males excel in science and math (pp. 126–127). What causes these differences is far from being understood.

This history of the idea of a peculiarly female intelligence can teach us several general lessons. The first is to beware of research that evaluates the sexes in terms of polarities and, in general, uses polarities as a means to understand the human mind. Second, beware of composite index numbers, such as IQ. Unless there is a strong theory, test items can be selected to verify any existing bias "scientifically." And third, keep in mind that intelligence is about cognitive processes. Therefore, we would be well-advised to replace polarities and IQ numbers with the study of the actual processes underlying intelligent behavior, a scientific research agenda that would also leave little room for individual and cultural biases.

The Persistence of Polarity-Based Theorizing

In his paper "You can't play 20 questions with nature and win," Newell (1973) criticized that psychological explanations are often in the form of binary opposites, such as nature versus nurture, serial versus parallel processing, conscious versus unconscious, and intuitive versus analytic. Newell thought of these general dichotomies as the nadir of theorizing where, instead of achieving clarity, "matters simply become muddier and muddier as we go down through time" (pp. 288–289). Together with Herbert Simon, Newell instead set out to study the heuristic decision processes people use to solve problems and make intelligent decisions. Yet half a century later, theorizing in terms of polarities (as opposed to processes) remains popular in cognitive psychology. Here is a prominent case.

Recall the opposition between intuition and analysis, as in Immanuel Kant's and Stanley Hall's view of female and male cognition. By the twenty-first century, its association with gender was mostly dropped in psychology, albeit continuing in parts of the general public (Gigerenzer et al., 2014). The polarity itself, however, has survived in psychological theorizing and is now used to characterize two allegedly opposite poles of a continuum of thinking, despite a meta-analysis of 75 studies that showed that measures of intuition and analysis are *not* negatively correlated (as opposites should be) but instead independent (Wang et al., 2017). Theorizing in terms of polarities also has survived in dual-process models (e.g., Evans & Stanovich, 2013; Kahneman, 2011). These consist of poles such as intuitive versus analytic, unconscious versus conscious, fast versus slow, and automatic versus deliberate, not unlike those in the first version of the idea of a peculiarly female intelligence. The poles are said to be aligned and form two systems of cognition, System 1 and System 2, despite the absence of evidence for such an alignment (Kruglanski & Gigerenzer, 2011; Melnikoff & Bargh, 2018). The intuitive, impulsive, and impatient System 1 has been linked to women's thinking. Consider the Cognitive Reflection Test (CRT), a short test comprising three numeracy questions. Women score on average lower than men, which is attributed to their supposedly higher reliance on the intuitive System 1 (Frederick, 2005, p. 37), a reinstatement of the old

stereotype about women. Yet leaping to the conclusion that lower numeracy results from higher intuition or impatience is neither necessary nor supported by the evidence (Bago & De Neys, 2019; Easton, 2018).

As this case illustrates, intuition and analytic thinking are still perceived by some psychologists as contraries, with intuition as the inferior pole that requires (male) analytic thinking to prevent it from error. This devaluation of intuition ignores the empirical evidence that experts need to rely on intuition to achieve better performance (Gigerenzer, 2007; Klein, 2017). As Albert Einstein famously said, "The intuitive mind is a sacred gift and the rational mind is a faithful servant. We have created a society that honors the servant and has forgotten the gift" (Calaprice, 2011). History is destiny. Despite Newell's warning, theorizing in terms of polarities persists and is still able to trump empirical evidence.

Beyond Polarities and IQ: Intelligent Decision Processes

The history of the idea of a peculiarly female intelligence shows, in my view, that the field of sex differences in intelligence, and of intelligence in general, could benefit from a fresh start. Herbert Simon's and Alan Newell's work on heuristics and artificial intelligence, which has inspired my own research on heuristic decision-making (Gigerenzer & Gaissmaier, 2011), can provide such a new framework. Heuristics are strategies that help to make decisions and solve problems in an intelligent and efficient way. After all, what we call intelligence manifests itself in the quality of the decisions we make. In the context of this chapter, I can only sketch out the research agenda, which centers on two questions: (i) What is the repertoire of intelligent strategies (such as heuristics) at a person's disposal for making decisions, and (ii) what is a person's ability to choose a proper strategy for the situation at hand (Gigerenzer, 2020; Gigerenzer et al., 2011)? In this framework, intelligence has a very concrete meaning that connects cognitive abilities with behavioral strategies, namely the adaptive toolbox of strategies available and the ability to choose a strategy wisely to achieve a goal.

While the study of intelligent heuristics is well established, it has had a blind spot for sex differences in how males and females search for information, when they stop searching, and how they make or delay decisions. One exception is the work of Meyers-Levy and Loken (2015), who reported that females search more extensively for information than males, while males are more selective in search and rely on faster stopping rules. Moreover, they concluded that females are more sensitive to environmental cues, whereas men more often ignore these and rely on the same heuristics across contexts, indicating less ability in adaptive choice. As for social heuristics, they found that women are more likely to base decisions on trust, are more likely to be trusted, and have higher ability in reading nonverbal cues and making inferences about the mental states of others. Note that these are preliminary findings, but they indicate a different kind of question to pursue: Abandon studying polarities and differences in IQ test outcomes and instead ask whether there are differences in the way males and females search for information and make decisions.

Conclusion

Does women's intelligence differ from men's? I believe it became clear over the course of the years that this question is ill-posed because the very idea of what intelligence is has shifted several times, and the various answers have been polluted by preconceived beliefs and biases in the absence of a theory of the nature of intelligence. Moreover, understanding potential sex-based differences in intelligence appears not to have been the primary course that history took, nor was it always the goal of measurement. Rather, measurement served to fortify preconceptions and biases. In my opinion, progress in the field requires going beyond polarities and IQ and analyzing the very strategies (heuristics) that males and females use to make intelligent decisions. Such a research program would also eliminate loopholes through which persisting strong beliefs about the nature of men and women can distort science. That said, some of the arbitrary decisions in the study of male and female intelligence have nevertheless contributed to an erosion of the millennia-old idea that nature has assigned women a subordinate social position. The idea that men and

women's intelligence is polarized has largely been eradicated, as has the idea that women are on average less intelligent than men. The supposedly greater male variability remains the last bastion of those who cling to the idea of male supremacy.

References

Abrahamsen, D. (1946). *The mind and death of a genius*. Columbia University Press.

Aristotle. (1984). History of animals. In *The complete works of Aristotle* (Vol. 1, revised Oxford transl. & J. Barnes, Ed.). Princeton University Press. (Original work published ca. 350 B.C.E.)

Bago, B., & De Neys, W. (2019). The smart System 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning, 25*, 257–299.

Binet, A. (1911). Nouvelles recherches sur la mesure du niveau intellectuel chez les enfants d'école [New studies on the measurement of schoolchildren's intellectual levels]. *L'Année Psychologique, 17,* 145–201. [Cited and translated in Wolf, 1973, pp. 209–210].

Binet, A., & Simon, T. (1914). *Mentally defective children* (W. B. Drummond, transl.). Edward Arnold.

Binet, A., & Simon, T. (1973). *The development of intelligence in children.* (E. S. Kite, transl.). Arno Press. (Original work published 1905; English translation first published 1916).

Blum, J. (1978). *Pseudoscience and mental ability*. Monthly Review Press.

Bruch, E. E., & Newman, M. E. J. (2018). Aspirational pursuit of mates in online dating markets. *Science Advances, 4*(8), eaap9815.

Calaprice, A. (2011). *The ultimate quotable Einstein*. Princeton University Press.

Carson, J. (2007). *The measure of merit: Talents, intelligence, and inequality in the French and American republics*. Princeton University Press.

Darwin, C. (1875). *The Variation of Animals and Plants Under Domestication* (Vol. 1, 2nd ed.). London: John Murray.

Daston, L. (1992). The naturalized female intellect. *Science in Context, 5*, 209–235.

Deary, I. J., Whalley, L. J., & Starr, J. M. (2009). A lifetime of intelligence. *Follow-up Studies of the Scottish Mental Surveys of 1932 and 1947*. American Psychological Association.

Diehl, L. A. (1986). The paradox of G. Stanley Hall: Foe of coeducation and educator of women. *American Psychologist, 41*(8), 868–878.

Dury, M. O. C. (1984). *Some notes on conversations with Wittgenstein. Recollections of Wittgenstein*. Oxford University Press.

Easton, C. (2018). Women and 'the philosophical personality': Evaluating whether gender differences in the Cognitive Reflection Test have significance for explaining the gender gap in philosophy. *Synthese, 198*, 1–29.

Ellis, H. (1984). *Man and woman*. Scott.

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223–241.

Eysenck, H.-J., & Kamin, L. (1981). *The intelligence controversy*. Wiley.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*, 25–42.

Galton, F. (1979). *Hereditary genius*. Julian Friedman Publishers. (Original work published 1869).

Geddes, P., & Thomson, J. A. (1890). *The evolution of sex*. Scribner & Welford.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. Penguin.

Gigerenzer, G. (2020). Intelligence and decision making. In R. J. Sternberg (Ed.), *Cambridge Handbook of Intelligence* (Vol. I, pp. 580–601). Cambridge University Press.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision-making. *Annual Review of Psychology, 62*, 451–482.

Gigerenzer, G., Galesic, M., & Garcia-Retamero, R. (2014). Stereotypes about men's and women's intuitions: A study of two nations. *Journal of Cross-Cultural Psychology, 45*, 62–81.

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press.

Gray, J. (1992). *Men are from Mars, women are from Venus*. HarperCollins.

Grosskurth, P. (1980). *Havelock Ellis: A biography*. Knopf.

Hall, G. S. (1976). Biological and anthropological differences between the sexes. In P. C. Lee & R. S. Stewart (Eds.), *Sex differences* (pp. 371–379). Urizen Books. (Original work published 1904).

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). Erlbaum.

Halpern, D. F., & Wai, J. (2020). Sex differences in intelligence. In R. J. Sternberg (Ed.), *The Cambridge handbook of intelligence* (2nd ed.). Cambridge University Press.

Hollingworth, L. S. (1914). Variability as related to sex differences in achievement. *American Journal of Sociology, 19*, 510–530.

Johnson, W., Carothers, A., & Deary, I. J. (2009). A role for the X chromosome in sex differences in variability in general intelligence? *Perspectives on Psychological Science, 4*, 589–611.

Kahneman, D. (2011). *Thinking fast and slow*. Allen Lane.

Kant, I. (2011). *Observations on the feelings of the beautiful and sublime* (P. Frierson & P. Guyer, transl.). Cambridge University Press. (Original work published 1764).

Klein, G. (2017). *Sources of power* (20th Anniversary ed.). MIT Press.

Kruglanski, A., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review, 118*, 97–109.

Lehrke, R. G. (1978). Sex linkage: A biological basis for greater male variability in intelligence. In R. T. Osborne, C. E. Noble, & N. Weyl (Eds.), *Human variation. The biopsychology of age, race, and sex* (pp. 171–198). Academic Press.

Masters, B. A. (1986, April 18). *When the Cliffies finally conquered Lamont*. The Harvard Crimson, Retrieved from https://www.thecrimson.com/article/1986/4/18/when-the-cliffies-finally-conquered-lamont/

McNemar, Q., & Terman, L. M. (1936). Sex differences in variational tendency. *Genetic Psychology Monographs, 18*, 1–66.

Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Sciences, 22*, 280–293.

Meyers-Levy, J., & Loken, B. (2015). Revisiting gender differences: What we know and what lies ahead. *Journal of Consumer Science, 25*, 129–149.

Minton, H. L. (1988). *Lewis M. Terman: Pioneer in psychological testing*. New York University Press.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing: Proceedings of the eighth annual Carnegie symposium on cognition* (pp. 283–308). Academic Press.

Pearson, K. (1897). *The chances of death and other studies in evolution* (Vol. I). Arnold.

Pearson, K., & Lee, A. (1903). On the laws of inheritance in man. I. Inheritance of physical characters. *Biometrica, 2*, 357–462.

Rifkin, J. (2002). *Science in the age of sensibility: The sentimental empiricists of the French Enlightenment*. University of Chicago Press.

Santayana, G. (1905). *The life of reason*. Scribners.

Schofer, G. (1976). G. Stanley Hall: Male chauvinist educator. *The Journal of Educational Thought, 10*, 194–200.
Scottish Council for Research in Education. (1933). *The intelligence of Scottish children: A national survey of an age group.* (Publications of the Scottish Council for Research in Education V). University of London Press.
Scottish Council for Research in Education. (1939). *The intelligence of a representative group of Scottish children* (Vol. VX). University of London Press.
Scottish Council for Research in Education. (1949). *The trend of Scottish intelligence* (Vol. XXX). University of London Press.
Scottish Council for Research in Education. (1958). *Eleven-year-olds grow up* (Vol. XLII). University of London Press.
Shields, S. A. (1982). The variability hypothesis: The history of a biological model of sex differences in intelligence. *Signs: Journal of Women in Culture and Society, 7*, 769–797.
Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence*. Cambridge University Press.
Summers, L. H. (2005, January 14). *Remarks at NBER Conference on diversifying the science & engineering workforce*. Retrieved from http://web.archive.org/web/20080130023006/http:/www.president.harvard.edu/speeches/2005/nber.html.
Terman, L. M. (1916). *The measurement of intelligence*. Houghton Mifflin.
Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence*. Houghton Mifflin.
Terman, L. M., & Miles, C. C. (1936). *Sex and personality*. McGraw-Hill.
Terman, L. M., & Oden, M. H. (1947). *The gifted child grows up: Vol. IV. Genetic Studies of Genius*. Stanford University Press.
Thompson, H. B. (1903). *The mental traits of sex*. University of Chicago Press.
Wang, Y., Highhouse, S., Lake, C. J., Petersen, N. L., & Rada, T. B. (2017). Meta-analytic investigations of the relation between intuition and analysis. *Journal of Behavioral Decision Making, 30*, 15–25.
Weininger, O. (1906). *Sex & character.* William Heinemann. (Original work published 1903).
Wellman, B. (1933). Sex differences. In C. Murchison (Ed.), *Handbook of child psychology* (2nd rev. ed., pp. 626–649) Clark University Press.
Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review, Monograph Supplement, 3*(6).
Wolf, T. H. (1973). *Alfred Binet*. University of Chicago Press.