Chapter 3

Empirical Evidence for Cognitive Sex Differences

CONTENTS

INTELLIGENCE

> Are men smarter than women? The answer to the above burning question is: No, they are not. Data are now being laid on the table that show that, on average, men and women are equal in mental ability.
>
> —Dan Seligman (1998, p. 72)

The first question most people ask about sex-related cognitive differences is which is the smarter sex—males or females? Although this question has a long and acrimonious history, the question of who is the smarter sex has persisted for at least as long as modern measurements of intelligence have been possible and probably long before then. There are several ways to find answers for this question. One logical way is to obtain large random samples of women and men, give them a psychometrically sound intelligence test (one with good statistical properties), and compare the scores for women and men. The sex with the higher average score would be the smarter sex. Although this may seem like a logical, straightforward approach to answering the question of sex differences in intelligence, it won't work. Intelligence tests are carefully written so that there will be no average overall difference between men and women (Brody, 1992). During the construction of intelligence tests, any question that tends to be answered differently by males and females is either discarded or balanced with a question that favors the other sex. Even though intelligence tests are revised repeatedly to reflect changes in the population, all changes are carefully considered so that they do not benefit men or women as a group. Therefore, average scores on intelligence tests cannot provide an answer to the sex differences question because of the way the tests are constructed.

A second way to decide whether men or women are, on average, smarter might be to look at who performs the more intelligent jobs in society. Of course, one would have to decide which jobs require greater intelligence. Suppose that most people could agree in principle that jobs like government leader, architect, lawyer, physician, professor, mathematician, physicist, and engineer all require a high degree of intelligence. An examination of who performs these jobs would reveal that the overwhelming majority of these jobs are held by men. Does this mean that men are, in general, more intelligent? Looking at the types of jobs typically performed by women and men in society cannot provide an answer to the intelligence question because of differential sex roles for women and men. Many professions were formally or informally closed to women until recent years. Similarly, there are few male nurses, secretaries, and child care workers because of the constraints imposed by the male sex role. There are still considerable socially related differences between the sexes in background experiences, types of encouragement, amount and type of education, and social expectations for success. We cannot know if the differences in the numbers of men and women in the various job classifications are related to overall intelligence differences or to differential socialization practices or to some combination of the two. This issue is discussed in greater depth in the chapters on psychosocial hypotheses (Chapters 6 and 7).

A third way of answering the intelligence question is to look at school achievement. Which sex, on the average, gets better grades in school and is more likely to obtain advanced degrees? Numerous sources of data clearly show that women get better grades in school than men in every subject area, although the differences are not large. The U.S. National Center for Educational Statistics found that across elementary and secondary schools, females consistently receive higher grades than males in the classroom (Snyder, Dillow, & Hoffman, 2009). Other studies confirm this general conclusion. For example, one large-scale study of high school grades found that the average grade point average (GPA) was 2.83 for boys and 3.05 for girls (using a 4-point scale; U.S. Department of Education, 2004). As students move into college, women continue to get better grades, on average, than men: 61% of females and 49% of males have a college GPA higher than 3.0 (Clune, Nuñez, & Choy, 2001). Paradoxically, girls get better grades than boys even in "traditionally male" content areas, such as mathematics and physics, in which boys score higher on tests used for college and graduate school admissions such as the SATs and Graduate Record Exams (GREs; Coley, 2001). Figure 3.1 shows the combined high school GPA for girls and boys in the United States from 1990 to 2005. Note that although there is a consistent advantage for girls in grades, the difference is generally small, about 0.23 of a grade on a 4.0 GPA.

Although girls achieve higher grades on average in science and mathematics courses in high school, boys achieve higher scores on average on Advanced Placement (AP) examinations that are written to test knowledge of high school course materials. Many colleges will accept a score of 3 or higher as equivalent to college-level achievement in the subject that was tested. Sex differences in science and mathematics AP tests are shown in Figure 3.2.

The data favor the intelligence of females when looking at course grades, and they favor males when looking at most high stakes exams such as the Advanced Placement tests and SATs. As you will see as you read through this entire text, it is often true that whenever some data seem to provide a clear picture and understanding of cognitive sex differences, other data provide an equally compelling and a diametrically opposed view and understanding of the same phenomenon. Consider, as another example, the percentage of men and women who attain college degrees. As explained in Chapter 1, in the United States and most western countries, men are attending and graduating from college at higher rates than ever before (Mead, 2006). But the steady increase in men's increasing college-going rates since the 1970s to the present has been eclipsed by the much greater rise in women's college-going rates. Women's rates had a sharp rise in the 1970s and a small but steady increase in the 1980s

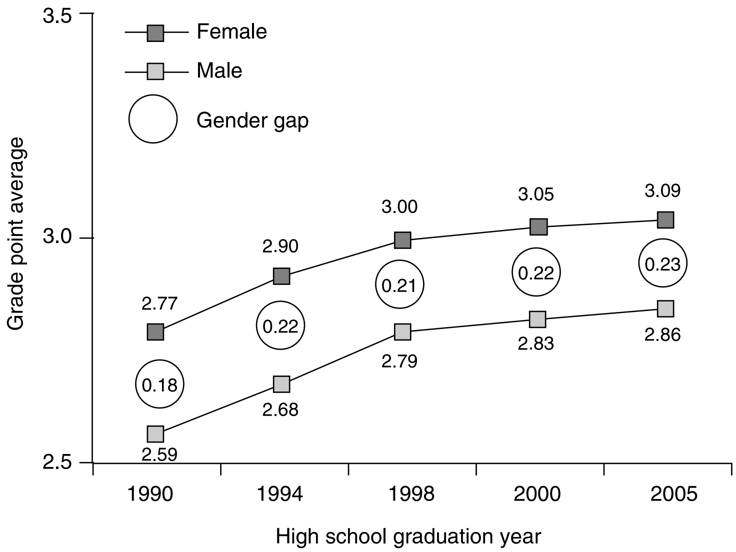


Figure 3.1 High school grade point averages for girls and boys in the United States from 1990 to 2005. From U.S. Department of Education, National Center for Education Statistics (2007b).

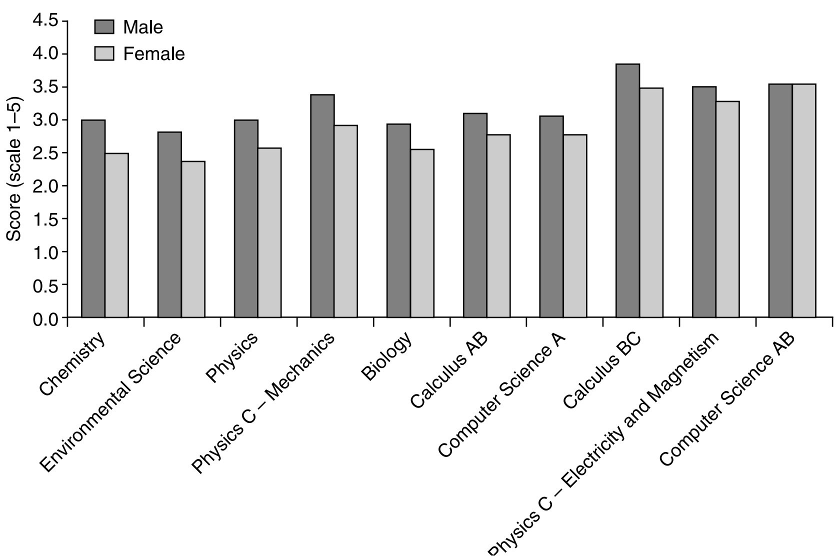


Figure 3.2 Average scores for males and females on Advanced Placement examinations in science and mathematics for 2009. From National Summary Report 2009. Copyright © 2010 The College Board. Reproduced with permission. www.collegeboard.com.

and beyond. Despite the higher rate of women attending and graduating from college, women are not selecting college majors that allow them to enter the higher-paying, more prestigious jobs at the same rate as men. There has been a large increase in the percentage of bachelor's degrees earned by women in biology (60%) and the agricultural sciences (60%), chemistry (52%), and mathematics (45%), with most degrees going to men in physics (79% male), engineering (80%), and computer science (80%; data from Hill, Corbett, & St. Rose, 2010, p. 9). There are large differences in the areas that men and women select as college and graduate school majors.

Consider graduate degrees in the sciences—there are many different kinds of sciences and the graduation rates of women vary widely among them. In the U.S. women are obtaining 50% of the MD degrees from medical schools, almost 78% of the DVMs from veterinary medical colleges, and 44% of dental degrees (Burns, 2010). The data for U.S. veterinary college enrollments are shown in Figure 3.3.

Clearly women are selecting careers and succeeding in some sciences. So, it is NOT a lack of cognitive ability for success in science that is responsible for their underrepresentation. Careers in the sciences and math can accommodate a wide range of abilities. Women are not achieving equally in all areas of science. In contrast, men make up a minority of enrollments in many of the majors that lead to careers in the helping professions such as clinical psychology, education, and nursing.

Figure 3.3 Data showing the change in veterinary school enrollment in the United States from 1970 to 2011. Figure from Burns (2010). Reprinted with permission from the Association of American Veterinary Medical Colleges.

There Are Data and There Are Interpretations of Data

> It has long been known that it is possible to create or eliminate differences in test scores by selecting different test items.
>
> —Beatriz C. Clewell and Patricia B. Campbell (2002, p. 264)

There are many types of information that researchers and the general public can use to decide whether women or men are the smarter sex. Even though many people consider it an unanswerable question that is best considered "a draw" given data that sometimes support the superiority of women and sometimes support the superiority of men, it is a question that just will not go away, probably because there are a handful of researchers who steadfastly maintain that women are less intelligent than men, and they seem to get coverage in the media with every pronouncement of male superiority. The superiority of males has been the consistent position of Rushton and his colleagues (who, by the way, also maintain that Blacks are less intelligent than Whites and Whites are less intelligent than Chinese; e.g., Rushton & Jensen, 2006). I hope that you are wondering what sort of data they use to support this contention.

As already explained, all of the major intelligence tests have been written so that there is no overall sex difference. The equality of intelligence test scores has been well known since 1942 when a leading intelligence test, the Stanford–Binet, was revised to "produce a scale which will yield comparable IQs for the sexes." After initial testing found women tended to score higher than men, the authors concluded that "intellect can be defined and measured in such a manner as to make either sex appear superior" (McNemar, 1942, p. 42). According to McNemar, "test developers sought to avoid using test items showing large sex differences in percentage passing" (p. 45) so they could produce a scale that would yield comparable IQs for males and females. In other words, the most important variable in determining whether females or males score higher on intelligence tests is the type of question that is asked.

Despite the fact that intelligence tests are written to show no overall sex differences, Jackson and Rushton (2006) claimed that males had a 4-point advantage on intelligence tests. They based this claim on calculations they made using the SATs, where in fact, males do achieve higher average scores than females. There are numerous reasons why SAT data cannot be used to decide that either sex is smarter. One problem with using SAT data to determine whether females or males are smarter is that many more females than males take the SATs; approximately 55% of SAT-takers are female. As explained in Chapter 2, when a greater percentage of one sex takes any assessment, especially a test that is used for college admissions, it is likely that more people from that group are "less-elite" than from the group with the smaller participation rate. We return to the SAT data in several places throughout this book because they play a central role in understanding how and when females differ in their cognitive abilities and because scores on the SATs are used, in part, to decide who gets admitted to competitive colleges and universities.

There are some researchers who report a small advantage for males on tests that were standardized to show no sex differences (Nyborg, 2005), but most do not (e.g., Colom, Garcia, Juan-Espinosa, & Abad, 2002; Deary, Thorpe, Wilson, Starr, & Whalley, 2003). In a recent review of the question of sex differences in intelligence, Dykiert, Gale, and Deary (2009) found that reported sex differences on intelligence tests can be explained by the use of samples that are not representative of females and males in general and, thus, reflect errors in the methods used to study this question. This conclusion was confirmed by Hunt and Madhyastha (2008) who provided a mathematical model of the subject selection problem that occurred in studies that report sex differences in intelligence. Hunt and Madhyastha concluded that there are numerous problems with studies that report sex differences in intelligence.

There are logical problems with any claim that one sex is smarter than the other. First, standardized measures of intelligence cannot be used to support this claim because one cannot use a test that was deliberately constructed and tested with a large standardization sample to ensure that there would be no overall sex differences to then support the conclusion that there are sex differences. Second, the conclusion that males are smarter than females ignores the many other sorts of mental measures on which females score higher than males, such as grades in school, writing tests, and many types of memory assessments. Any similar conclusion with females as the smarter sex faces the same logical problems because there are many areas that males excel in, as described later in this chapter.

In an interesting twist, Jensen (1998) joined the debate over sex differences in intelligence. Jensen is no stranger to heated controversies about intelligence. In a 1969 paper, he asserted that African Americans are, on average, less intelligent than European Americans, a position that he has maintained to the present. As a means of addressing the question of male–female differences in intelligence, Jensen analyzed tests that "load heavily on *g*" (*g* is the generally accepted term for general intelligence). In his analysis, Jensen used only tests that had not been deliberately written to eliminate sex differences, thus making it more likely that he would find evidence for sex differences in intelligence, if they existed. Jensen used five different test batteries for which he had large, representative samples that encompassed the full range of ability in the general population. Jensen concluded, "No evidence was found for sex differences in the mean level of *g* or in the variability of *g*. . . . Males, on average, excel on some factors; females on others" (pp. 531–532). A study of the intelligence of children in Belgium and the Netherlands confirmed Jensen's conclusion that there are no sex differences in overall intelligence, although sex differences are found on some subtests (van der Sluis, de Jong, & van der Leij, 2007).

Perhaps the most important lesson to be learned from this debate is that researchers, like the rest of us, maintain a particular world view that they use in interpreting research findings. This point became clear to me during a recent discussion of these issues that I had with a developmental psychologist. After I explained to him that females get higher grades in school and males get higher scores on (some) standardized tests of cognitive abilities, his face brightened. He filtered this information according to his own world view and exclaimed, “That proves that schools are biased against boys.” “Perhaps,” I responded. “But it could just as easily be used to ‘prove’ that the tests are biased against girls.” This is a good example of two contradictory explanations of the same findings—each of us making the leap from data to our interpretations of data via our privately held world views.

The problem with questions like “which is the smarter sex” is that they begin with the assumption that there ***is*** a “smarter sex.” The research reviewed in this book suggests several areas in which sex differences are consistently found and other cognitive areas where sex differences are not found, but in no way does this mean that one sex is the “winner” and the other the “loser,” or that one sex is smarter and the other is dumber. The more meaningful questions are when, where, and why are cognitive sex differences found. Modern society is complex and diverse. There is no single best set of intellectual abilities for all of society’s tasks. It is important that we come to think of differences apart from value judgments about who and what is better. If society consistently values the abilities that are more frequently associated with one sex, then the problem lies in the way differences are valued, not in the fact that they exist.

A more fruitful approach to the cognitive sex differences question is to examine specific abilities, especially in light of the fact that intelligence is not a unitary concept. It is theoretically more useful to think of multiple “intelligences” than to consider intelligence as a single homogeneous mental ability. The question then becomes, “What are the sex differences in cognitive abilities?” Although intelligence tests are constructed so that there will be no overall sex difference in intelligence, the tests do differ in the pattern of intellectual abilities for the two sexes. Surprisingly, in an area as controversial as this one, there is little disagreement about which of the cognitive abilities differs by sex. As you will see, the most heated debates revolve around whether the differences are large enough to be important and why these differences exist.

THE WHEN, WHERE, WHO, AND HOW OF DIFFERENCES

Although there are no overall differences in intelligence between males and females, sex-related cognitive differences are found consistently on tests of some cognitive abilities. Between-sex differences show an uneven pattern of results that often depend on the portion of the ability curve being sampled (e.g., gifted individuals, low ability, or the average-ability range), the age of the sample (infancy, preschool, middle childhood, adolescence, adulthood, old age), the response format (multiple choice, essay, diagram), and probably many other variables, including socioeconomic status and the gender equality and level of economic development of the society in which people live. What this means is that simple answers that apply to all females and all males are impossible. In addition, the size of the between-sex difference depends on other moderating factors, such as education, home environment, testing conditions, personal and societal beliefs, and many more.

Tails of Distributions

As introduced in Chapter 1, researchers and commentators in the area of cognitive sex differences can be thought of as “difference maximizers” or “difference minimizers” (Wai, Cacchio, Putallaz, & Makel, 2010). Researchers who stress findings that show that males and females tend to be highly similar (Hyde, 2005; Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Hyde & Linn, 2006) focus on findings that relate to the average differences between the sexes, which tend to be smaller than differences in high achieving or low achieving tails of distributions. On the other hand, difference maximizers focus on the tails of distributions, usually the high achieving tail, arguing that the highest achievers become our scientists, engineers, economists, physicians, and other professionals (e.g., Park, Lubinski, & Benbow, 2007; Wai, Lubinski, Benbow, & Steiger, 2010). Recall from Chapter 2 that an important distinction was made between average differences and differences among the highest achievers (high ability tail) and lowest achievers (low ability tail). Most people score near the middle of the distribution—the mean or average—but often researchers are concerned with people who have particularly high or low ability. The difference between the average scores and tails of a distribution is important, so readers who are not certain what a “tail” of a distribution is should reread the relative sections in Chapter 2, and carefully examine the various “bell-shaped” curves that are presented in that chapter.

Consider the comments made by Lawrence Summers in 2005, when he was president of Harvard University: “There are three broad hypotheses about the sources of the very substantial disparities that this conference’s papers document and have been documented before with respect to the presence of women in high-end scientific professions.” The three hypotheses that he outlined were (a) many women do not want to work the 80-hour weeks that are standard in high level careers, a topic that is addressed later in this book in the chapters on psychosocial hypotheses; (b) differences in how men and women are socialized and possible discrimination during the hiring process, which is also discussed in the later chapters that discuss psychosocial hypotheses; and (c) there are relatively few women with high-level aptitude to become the leaders in science and math. Thus, Summers used the relatively low ratio of women to men in the high achieving tails of cognitive abilities as a main reason why women are underrepresented in "tenured positions in science and engineering at top universities and research institutions."

An example from an area outside of cognition may help in demonstrating how important all of these factors are in understanding sex differences. Virtually all social scientists believe that males, on the average, are more aggressive than females. What does a conclusion like this really mean? No one believes that the meekest male is more aggressive than the brashest female. Everyone realizes that there must be overlap between the female and male distributions for aggression. As you will see in this example, the kind of conclusion we draw depends upon the portion of the distribution we study. Let's consider the extremes of aggression, that is people who are exceptionally high and exceptionally low on aggression.

If we considered only the most aggressive individuals in society, we would have to conclude that there are huge sex differences with respect to aggression. The overwhelming majority of violent crimes (sadistic murders, rape, mutilation, serial killings, slasher crimes) are committed by males, and this is true in every society for which we have reliable data. For example, "between 82 and 94 per cent of all offenders in England and Wales found guilty of, or cautioned for, violence against the person, criminal damage, drug offences and robbery and burglary were male" (Office for National Statistics, British Crime Survey, 2007/2008, para. 3). Similarly, data from the United States show that "gender is the single best predictor of criminal behavior: men commit more crime and women commit less crimes" (American Law and Legal Information, 2010, para. 1). In the United States, 93.2% of prisoners are male (Harrison & Beck, 2003). Even though males commit most of the violent crimes in all societies, in fact, only a relatively small percentage of males are criminals. These data show how large sex differences in the extreme end of the distribution of violent behavior can have large effects on how we think about sex and crime.

Look carefully at Figure 3.4. It shows the male:female ratio at the very highest tail (right-hand tail) of the distributions for the SAT-M and the ACT-M (American College Testing Service Math Test) and ACT-S (American College Testing Service Science Test). These data come from 1.6 million seventh grade students who took the SATs and ACT as part of the screening process to identify academically precocious youth. As shown in this figure, the ratios were approximately 13 males for every female in the early 1980s (Benbow & Stanley, 1980; 1983). These ratios decreased to approximately 7 males for every female in the late 1980s and then leveled off at between 3:1 and 4:1 since the early 1990s (Wai, Cacchio et al., 2010).

Because data from the SATs and other standardized tests are often used to document sex differences in cognition, it is important to keep in mind the fact that people who take college entrance examinations are a more select group than a sample of high school graduates, so the data from college entrance exams are drawn from high achieving students. The tails of these distributions

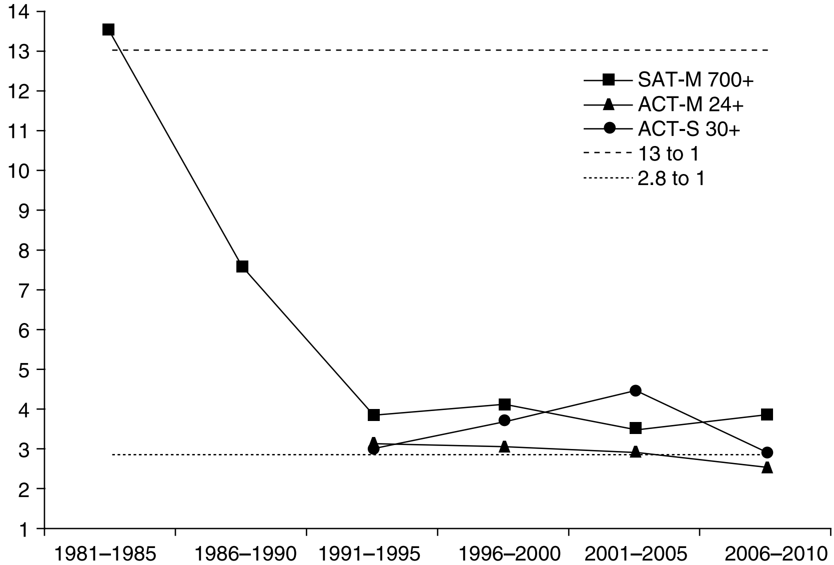


Figure 3.4 Ratio of males to females in right-hand (high scoring) tail of the SAT-M, ACT-M, and ACT-S from 1981 to 2010 for gifted youth. Notice that the proportion of males to females declined in the early 1990s and has stayed constant over the last two decades. Figure from Wai, Cacchio et al. (2010). Copyright © 2010, with permission from Elsevier.

are therefore data from high achieving students who are planning to attend college, and for these data, the sample is very high achievers who were identified as precocious youth. These data do not tell us why there are more males in the right-hand tail or whether the proportion of males to females will remain at between 3:1 and 4:1 or continue to decline as it did in the 1980s. They do, however, provide data about what is currently true about sex differences in the tails of these distributions for high achieving youth in the United States.

The high ability tails of standardized tests also show a clear advantage for girls on the SAT-Test of Written English (which was discontinued in 1994), the SAT-Writing sample (which was introduced in 2005), and the ACT-English test, which is a measure of verbal reasoning. The female advantage in writing is the most robust of these findings with the female to male ratio at 2.38:1 (in 2008; Wai, Cacchio et al., 2010).

Some researchers believe that data from standardized tests like the SATs are especially useful because they are based on very large numbers of test-takers and they can be used to show trends over time, such as the decline in the male to female ratios among the highest scorers from the 1980s to today. Others believe that data from college entrance exams should never be used because many more females than males take the SATs, which means that the sample of females is "less select" than the sample of males, a fact that would be expected to decrease the mean scores for females, but should not affect the number of females scoring at the top. (The reasoning behind these predictions may need some additional thought. Imagine a group of 100 males and 100 females who are approximately the same in some ability. If the top 50 from each group took a standardized exam, we would expect approximately the same average score from each group, with the same proportion of females and males at the top and bottom range of the scores. Now imagine that the top 60 of the women and the top 50 of the men take the same test. Now we would expect a lower mean score for the women because the additional 10 women were lower in ability than the top 50 women. By reaching "deeper" into the proportion of women, more women with less ability are taking the test and they are bringing down the mean score for women.)

Data from the American College Testing (ACT) program support the general results from the SATs in the direction of sex differences, but not in the size of the differences. The ACT is a national testing program, which is also used in decisions about college admissions. To get around the problem that many more females take the SATs than males do, we can examine the data from the state of Illinois which requires all high school seniors to take the ACT. Data from 2009 show that 52% of test-takers from Illinois were female, so even with the requirement that all students take this test, more females take the ACT (ACT, 2009). There are four separate scores, with females scoring higher on English ($d = 0.10$) and reading ($d = 0.02$) and males scoring higher on mathematics ($d = 0.17$) and science ($d = 0.13$). Although results are in the same general direction as found with the SATs, the effect sizes are small. The ACT also reports the percentage of students who are ready for college-level work. Using the data from Illinois, these are 64% for males and 68% for females in English, 44% for males and 37% for females in mathematics, 48% for both males and females in reading, and 30% for males and 19% for females in science.

Variability

When we turn our attention to cognitive abilities, researchers regularly (but not always) report that males are more variable than females. In other words, there are more males than females at the very high end and at the very low end in many tests of cognitive abilities and, correspondingly there are fewer males than females in the range of average abilities (Deary, Irwing, Der, & Bates, 2007; Hedges & Nowell, 1995; Hunt & Madhyastha, 2008; Willingham & Cole, 1997). Greater variability for males has been replicated with large samples in both the United States and United Kingdom (Lohman & Lakin, 2009). Thus, we find males to be overrepresented in both the top and bottom percentiles (e.g., more males in the top and bottom 10% on some tests) with smaller between-sex differences for those in the average range. It is important to understand that differences in average performance between males and females tend to be smaller than differences at the extreme ends (either very low or very high performance). So, using the same data, we can conclude 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.

Developmental Perspectives

Cognitive abilities change throughout the life span. We are born into the world with far fewer abilities than we have even a few years later, when we move from infancy into toddlerhood. We spend most of our lives as adults, a time when we use these abilities to earn a living, raise a family, and contribute to society. Of course, many people continue with these activities well into old age, but eventual cognitive slowing and decline is an inescapable fact of old age, even though there are some who can retain their cognitive abilities at high levels into old age. If we want to understand how males and females differ and are similar in their cognitive abilities, we will need to consider age as a critical variable because these abilities change over the life span.

Measurement Variables

How can we best measure amorphous concepts like intelligence or cognitive ability? This is a difficult question to answer. The way one chooses to measure these psychological constructs will also affect results. For example, Bridgeman and Morgan (1996) found that females tend to score higher on written measures than on multiple choice questions, with the reverse for males. In a recent analysis of whether different types of question formats affected responses by males or females, Kelly and Dennick (2009) looked at responses on examinations in 359 medical school courses. They found that males did better, on average, when questions were answered with a true-false-abstain format, but females did better with short answer questions. (The advantage for females on short answer questions was small compared to the much larger advantage males had on true-false-abstain questions.) These authors concluded that all assessments should be evaluated for the possibility of bias caused by the format used to assess knowledge.

The SATs were revised in 2006 so that they now consist of three tests, a critical reading portion (SAT-CR), a mathematics portion (SAT-M), and a writing portion (SAT-W). Taken together, these three tests are called the SAT-I to differentiate them from subject area tests, which are called SAT-II. The addition of a writing sample was expected to boost the scores of women, who tend to perform better on tests that require a written response. And it did. Women outperform men on the writing portion of the SAT by 13 points (male = 486, female = 499; College Board, 2009). But males perform slightly better, on average, on the critical reading portion of the SAT, which requires a multiple choice response (male = 503, female = 498).

In 1996, the Department of Education arrived at a court-mediated agreement with the College Board concerning the Preliminary Scholastic Assessment Test (PSAT), a pre-college achievement test which is taken by more than 1.5 million students annually (College Board, 2009). The PSAT is used for determining semi-finalist status for the award of the prestigious National Merit Scholarships for college. In the past, a disproportionate number of these scholarships was awarded to males because they obtained higher average scores on the PSAT. As a result of the agreement, the PSAT now includes a writing component—one area in which girls usually excel. The addition of the writing component has resulted in a greater number of scholarship awards being won by females. The academic areas tested on the PSAT remain the same and, of course, the addition of a writing sample did not alter the average abilities of girls or boys; the only change is in the way their cognitive abilities were assessed. As expected, girls are now outscoring boys on the test of writing, a fact that balances the fact that boys outscore girls in the mathematics portion of these exams. As you can see from this example, measurement issues are critically important to how we understand sex differences in cognitive abilities and they also have important social consequences.

PERCEPTION

> Better that a girl has beauty than brains because boys see better than they think.
>
> —Author unknown (quoted in The Quote Garden, 2011)

All of our information about the world around us comes from our sensory systems. The cognitive or thinking process begins with the ability to sense changes in the environment and to make meaning out of the sensory stimuli constantly impinging on us. We no longer think that the infants' world is a bewildering array of random stimuli because it is now well known that newborns have innate capacities that they use to construct knowledge from sensory input. The first steps in the cognitive process are perception and attention. Sex differences in perception and attention are of particular interest because if there are sex differences at the earliest stages of information processing, this would provide a theoretical basis for positing sex-related differences in later stages. If we found early perceptual differences for boys and girls, these differences could create behavioral dispositions that vary as a function of sex. Differences in perception or attention would not mean that cognitive sex differences are inevitable or unalterable, only that they probably have an early physiological basis.

Audition, Olfaction, Vision, Taste, and Time Perception

Baker (1987) reviewed a variety of sex-related differences in perception and attention; and despite being over two decades old; it is still among the best overviews on this topic. She documented numerous sex differences in each of the sensory systems. The general conclusions have held up over intervening decades of research. In hearing, for example, females are better at detecting pure tones (tones of one frequency) during childhood and most of adulthood. Several studies have shown that adult females have more sensitive hearing for higher frequencies than males do and there have been several reports that school-aged girls have lower auditory thresholds than boys, which means that they can hear faint sounds better, but these differences are small and often are not statistically significant (reviewed in Al-Mana, Ceranic, Djahanbakhch, & Luxon, 2008). There are also sexually distinct patterns of hearing loss in middle age, with males beginning to lose the ability to detect high tones at about age 32 and females beginning to lose this ability at about age 37. A large-scale study of hearing in older adults in China found no differences in auditory function problems between ages 65 and 90 (Wang, Zheng, Kurosawa, & Inaba, 2009). A comparable large-scale study in the United States found greater hearing loss among men in their 70s and 80s than among women in this age range (Pratt, Kuller, Talbott, McHugh-Pemu, Buhari, & Xiaohui, 2009). Environmental causes for hearing loss in the sample of older adults in the United States can explain some, if not all, of the differences in hearing loss with aging—smoking, exposure to loud noises, and cardiovascular disease all contribute to hearing loss—and many of these causal factors are more prevalent among men than among women in the United States and most of the world.

In an extensive set of studies on sex differences in the auditory system, McFadden (1998; 2008) reported a wide array of sex differences in auditory perception ranging from binaural beats (a somewhat abstruse auditory phenomenon in which the brain produces a tone or "beat" when two tones of slightly different frequencies are presented separately to each ear), otoacoustic emissions (which are sounds generated by the auditory system that can be measured by sensitive microphones—females have stronger otoacoustic emissions than males). It is unlikely that differences of this sort can be attributed to sex-differentiated socialization practices. One reason for believing that sex differences in otoacoustic emissions are not caused by environmental experiences is that these differences are found in newborns (Berninger, 2007). Many of these sex-related differences in perceptual thresholds are detectable soon after birth, suggesting that they do not reflect learning, response biases, or postnatal environmental factors. For example, Reinisch and Sanders (1992) reported that newborn females are more sensitive to touch than newborn males. They found evidence of sex differences in the functional development of the central nervous system as early as 3 months of age. These are important findings in our quest to understand sex differences in cognition because these early perceptual differences could create different behavioral dispositions for boys and girls. Slight behavioral predispositions could then be exaggerated, reduced, eliminated, or ignored, depending on the way cultures respond to male–female differences, although I note here that the size of the differences is small and there is so much overlap in the perceptual data for girls and boys that it is unlikely that they have any meaningful effect on behavior.

One sensory system where females have a clear advantage is the ability to detect faint smells. The term for this ability in the research literature is "olfactory threshold sensitivity." Although women are more sensitive to some odorants and are better at identifying, discriminating (knowing that two very similar smells are indeed different), and remembering odors, the differences that are reported are usually quite small (Doty & Cameron, 2009). The female advantage in odor identification extends across the entire life span with females performing better at odor identification tasks from age 5 to 99 (Doty, Shaman, Applebaum, Giberson, Sikorsky, & Rosenberg, 1984). In a recent study of olfactory abilities across childhood, researchers found that the ability to detect odors and to identify them increased as children developed from ages 4 to 12, and girls performed better than boys (Monnery-Patris, Rouby, Nicklaus, & Issanchou, 2009). The researchers thought that one explanation for the superior ability of girls was that they had better verbal abilities than comparably aged boys and thus what appeared to be an advantage at an olfactory task really was a difference in verbal abilities—they were better able to label odors and remember the labels. This is exactly what they found. When they compared boys and girls who had the same verbal ability, the advantage that girls had shown in olfactory tasks disappeared. Thus, the difference was in the way females and males labeled different odors and their memory of the labels and not in the olfactory system. We cannot use this single study to conclude that males and females really have the same olfactory abilities because there is a large literature that found an advantage for girls, but this study does suggest that at least some of that advantage may be due to better verbal abilities among children aged 4 to 12.

Vision is a critically important sensory system for humans. In general, males under the age of 40 have better dynamic visual acuity (ability to detect small movements in the visual field) than females. Age-related loss of far vision occurs earlier for females (between ages 35 and 44) than for males (between ages 45 and 54). The question as to why there should be sex and age differences in dynamic visual acuity is more complex than it may seem at first. In a review paper on training in perceptual-cognitive skills, the authors discuss successful training programs that could, at least in part, explain these differences (Ward, Farrow, Harris, Williams, Eccles, & Ericsson, 2008). With deliberate and appropriate practice, people improve in these abilities. It may be that males, in general, who are more likely to play football, soccer, and other sports and video games where the quick detection of visual motion is key to winning, may develop these skills through sports and other experiences. Several studies report sex difference in color vision, beyond what would be expected by the fact that many more males are color blind than females (McIntyre, 2002). For example, recent research on color vision found that when a random sample of adults were asked to match colors, the men and women used different ratios of red and green in their color matching (Pardo, Perez, & Suero, 2007). These results suggest that women and men may “see” colors somewhat differently.

In a review article on sex and aging effects in taste perception, Mojet (2004) reported that many studies reported sex differences, with males making more errors than females in recognizing the basic tastes of sour, sweet, salty, and bitter. In general when differences in taste perception are found, females are more sensitive to different tastes, but there are many researchers who did not find any sex differences. In her own studies, Mojet found that women perceive the intensity of aspartame as stronger than men did, with the reverse effects for quinine, so simple conclusions about one sex always being more sensitive to tastes are wrong. Other evidence shows that older men have higher taste thresholds than older women (i.e., they are less able to taste tiny concentrations of chemicals on their tongue). There are many possible reasons for these results, including the likelihood that men, in general, smoke more than women do and smoking impairs the ability to taste foods.

There are also sizable sex differences in temporal cognition, that is our knowledge of and judgments about the passage of time. Hancock (2011) conducted an extensive review of the research relating to sex differences in time perception and made sense out of a massive and inconsistent research literature that is well over 100 years old. He concludes that “there are consistent temporal processing differences between the sexes” (p. i). Interestingly, the nature of the difference in temporal judgments depends on how it is measured. A recent study of adult men and women between the ages of 20 and 69 found that women consistently underestimated short intervals (up to 20 seconds) and men consistently overestimated the same time intervals while using the production method of time estimation (Hancock & Rausch, 2010). In the production method, participants are asked to create a tone, often by pressing a lever that lasts as long as the estimated time. So with the production method, if participants were asked to estimate 2 seconds, they would depress a lever that would emit a tone and release it when they estimated that 2 seconds had passed. Results from the study in which women and men were asked to estimate 1-, 3-, 7-, and 20-second intervals are shown in Figure 3.5.

It has been hypothesized that sex differences in time perception are caused by differences in body temperatures between females and males (Hancock, Vercruyssen, & Rodenburg, 1992). Females have higher resting body temperatures than males, which may affect the way each sex estimates time intervals. A recent study of sex differences in the ability to sense changes in temperature found at most a difference of 1 to 2 degrees (Fahrenheit) for children, so these

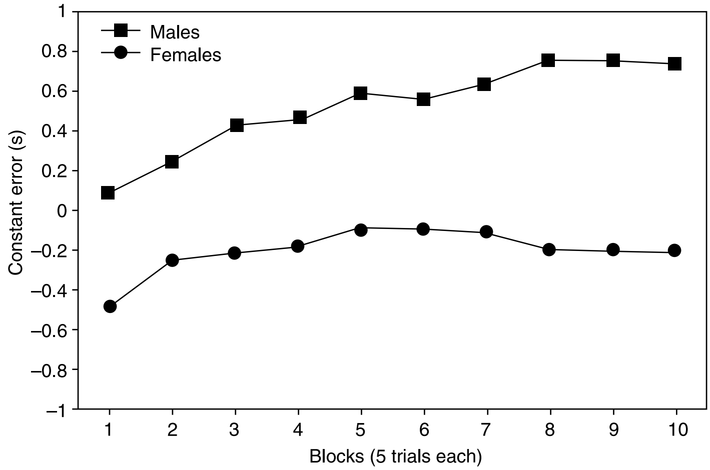


Figure 3.5 This figure shows the results of a time estimation task in which women and men were asked to press a button to estimate various time intervals (1, 3, 7 and 20 seconds). Over ten trials, men tended to overestimate the intervals with a positive level of constant error while women tended to underestimate the same intervals showing a negative level of constant error. Figure from Hancock and Rausch (2010). Copyright © 2010, with permission from Elsevier.

are very small differences (Blankenburg et al., 2010). In an extensive review of sex differences in time perception, the authors concluded that there are small, but reliable differences in the way women and men judge the passage of time (Block, Hancock, & Zakay, 2000), but these differences are moderated by many variables including age, the number of times (trials) each individual estimates the respective interval, and the way time judgments are assessed (Hancock, 2011).

Perceptual Motor Tasks

Sex differences favoring females are also reliably found on some speeded perceptual tests and some perceptual motor tasks. There are several tests that tap these abilities, each somewhat different in what it measures. Tests of perceptual speed and perceptual motor skills may require the rapid matching of stimuli, such as novel shapes, the “Finding A’s Task,” which requires scanning long columns of words and crossing those that contain the letter “A,” and copying simple forms from one line to another (Gouchie & Kimura, 1991). Another similar test of perceptual speed is digit-symbol coding. (It was a subset on the Wechsler Intelligence Tests, but was dropped from the Wechsler Tests in 2008 when the latest version of the WAIS—Wechsler Adult Intelligence Scale— was introduced.) As the name implies, the task for the participant is to copy a row of symbols where each symbol corresponds to a number (e.g., 2 = ˆ). In general, females copy more symbols correctly in a short period of time than males do (Burns & Nettlebeck, 2005; Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003.). One test of perceptual speed involved a series of three outlined faces, one of which is different from the other two (Jäger & Althoff, 1994). The participant's task is to identify the "odd" face, which differs in one detail, such as missing an eye. There are no sex differences in this test (Hausmann, Schoofs, Rosenthal, & Jordan, 2009), which shows, once again, that although tests may be categorized as being similar (in this example all measures of perceptual speed), there are differences among the tests which could explain why sex differences are found on some tests of perceptual speed, but not others. These tasks usually require rapid, fine motor movements such as quickly marking a symbol on a paper, another area in which females excel. Jensen (1998) found a very large female advantage on tests of perceptual speed, with the effect size as large as d = 0.86 among 12th grade students. This is a very large difference between the boys and girls on these tests. (Effect sizes were discussed in Chapter 2. Readers who are not sure what this term means should review the section on effect sizes so that they can understand what it means to say that an effect is large, medium, or small.)

Recall that general intelligence tests are written so that there are no overall differences in IQ scores for females or males, but there are differences in the subtests that make up the overall IQ. In a review of sex differences on the subscales of the WAIS-III (the newest version is the WAIS-IV, which was introduced in 2008), Longman, Saklofske, and Fung (2007) found a female advantage on the Processing Speed Index (d = 0.31), which is large enough to be meaningful, although like most of these indices, it is difficult to know what this difference would mean in everyday life.

Numerous studies have shown that females are usually superior at tasks that require fine motor manipulations. Kimura (1993) defined motor dexterity as "quick and effective use of the hands in the manipulation of small objects" (p. 1107). Nicholson and Kimura (1996) determined that women were faster than men when the task involved rapid repetitions of a sequence of movements. There is a large research literature on the topic of sex differences in motor tasks. Recent research shows that sex may be an important factor in motor performance and especially in motor learning. Most of the differences that are found in laboratory tasks are small, with little apparent application to everyday motor skills. These tests are sometimes labeled "clerical skills tests" and have been used to argue that females are naturally suited for clerical tasks like typing. I note here that fine motor skills are also needed in a variety of other professions such as brain surgery, dentistry, and the repair of small engines. One could just as easily use these experimental results to argue that females are naturally better suited for these other professions as well. Once again, I stress the distinction between research findings and the interpretation of these findings.

Findings of female superiority on fine motor tasks were questioned by Peters and Campagnaro (1996), who hypothesized that female superiority on fine motor tasks is an artifact of sex differences in finger size. To test this possibility, they had males and females perform a task that required subjects to rapidly move pegs in a peg board. They used both thick and thin pegs and had the subjects perform the task with and without tweezers as a way of controlling for finger size. They found that the female advantage on this task disappeared when the subjects had to use tweezers, a result that they interpreted as support for their hypothesis. As a follow-up to these studies, Rohr (2006) used a computer pointing task to investigate whether the distinction made by Peters (2005) would be supported with data from this task. Rohr found that although the women took longer than the men to make pointing movements, especially on more difficult tasks, they made fewer errors than the men did. Thus he concluded that "it is therefore important to consider gender-specific movement biases when interpreting performance differences and similarities between men and women" (p. 436). Thus, simple conclusions about which sex is better at fine movements will depend on whether speed or accuracy is assessed.

Numerous studies have found that women tend to be better at perceiving fine surface details by touch (e.g., Goldreich & Kanics, 2006). This finding holds up for blind as well as sighted participants, so it is not related to visual acuity. Women's smaller finger size is responsible for their ability to feel fine spatial structures because small fingers have a great density of Merkel cells, which are the type of cells that give rise to the perception of textures (Peters, Hackeman, & Goldreich, 2009). As the authors of this study explain, a man and a woman with fingers of equal size will experience the same tactile sensitivity.

On the other hand, motor tasks that involve throwing a projectile or otherwise aiming at a moving or stationary target show large advantages for males (Hall & Kimura, 1995; Watson & Kimura, 1991). This conclusion is based on studies that required throwing darts at a dart board or balls at a target and computer games that require subjects to "hit" a moving target on the screen, a task formally known as projectile interception. This is an important finding in the literature because of its implications for evolutionary theories. Recall from Chapter 1 that males were the hunters in hunter-gatherer societies and would need these skills in order to kill prey and human and animal enemies. Of course, starting at an early age, males in western societies have more practice throwing balls and other objects, so it is also possible, perhaps likely, that these differences in throwing accuracy result from differential life experiences and are not a legacy from our evolutionary past.

Attention

The ability to attend to stimuli and to switch attention is both a precursor to and a consequence of the thinking process. There are huge sex discrepancies in attention disorders. Sex ratios for attention deficit disorder (a psychiatric diagnostic category that often includes hyperactivity) range from 3:1 to 10:1 with the larger value corresponding to the proportion of males diagnosed (Biederman et al., 2002). One area of heated controversy concerns the question of whether infant boys and girls prefer to attend to (look at) different sorts of things, and the related question of whether there are sex differences early in life in the way infants understand the world. Spelke (2005) made a strong case against the idea that males have an intrinsic aptitude for mathematics and science. One of her main points was that males and females do not differ in what they attend to early in life. She cited numerous studies that support this contention. However, like almost every claim we will review pertaining to cognitive sex differences, there is also good evidence for the opposite conclusion. One reason why differences among female and male infants are so important is that they are unlikely to have been created by differential life experiences. In general, we tend to think of sex differences in infants as more likely caused by biological variables, but infant girls and boys are treated differently from birth, so we cannot know the influence of environmental variables even at very young ages.

Two related types of studies have emerged as critical in the dispute about differences early in life in female and male preferences. The first concerns whether infant girls and boys prefer different types of toys. Research has shown that sex differences in play start to emerge during the second year of life and are well established by age 3 (Ruble, Martin, & Berenbaum, 2006). Since children's sex-typed toy preferences (e.g., trucks for boys and dolls for girls) tend to resemble the activities of adult men and women, it seems logical that these toy preferences are caused by sex role socialization practices. To determine if sex-typed toy preferences occur early in infancy, Alexander, Wilcox, and Woods (2009) presented a doll and a truck to 5-month old girl and boy infants and recorded the number of times each infant looked at each toy. Looking time is a common measure of interest for infants. They found that boys looked at the truck and doll about equally; whereas, the girls looked at the doll much longer than the truck. These data are shown in Figure 3.6. The authors concluded that their results show different patterns of attention to toys because girls and boys are attracted to different visual characteristics of the objects. They speculate that these innate preferences are likely to be enhanced by the continuous process of gender socialization. But, note an important point—the only difference was that girls were less likely to look at trucks than boys were. In theory at least, the baby dolls resemble humans and both the infant girls and boys looked equally long at the dolls (i.e., the difference was not statistically significant). The idea that these differences in toy preferences are biologically determined is discussed in more detail in Chapter 4, where toy preferences for girls who were exposed to high levels of prenatal androgens are presented along with (possible) data showing sex-typed toy preferences in some nonhuman mammals.

The other type of finding regarding infant perception and cognition that is hotly disputed involves sex differences in an infant's ability to understand

Figure 3.6 Girl infants spent more time looking at dolls than at toy trucks; boy infants looked equally long at both dolls and trucks. Figure from Alexander, Wilcox, and Woods (2009). Reproduced with kind permission from Springer Science + Business Media.

complex events. There has been an explosion of research on infant cognition, and we have come to understand that infants have a much better understanding of the physical world than researchers ever imagined. One way infant cognition is studied is with "occlusion events." As its name implies, an occlusion event occurs when something happens behind a screen, so it is occluded from view. In a standard set-up, an infant would be held on a parent's lap with a small stage immediately in front of them. A ball would roll into view and pass behind a screen and then roll out a second or so later depending on how fast it was moving. Infants show that they expect the ball to roll out the other side of the screen by shifting their gaze to the other side of the screen. In a variation on this paradigm, infants watched while either a box or a ball was moved behind the screen. The screen was then removed and a ball was revealed. The ball would have been expected on those trials when a ball was moved behind the screen, but it would not be expected when a box was moved behind a screen. Infants showed their surprise at seeing a ball when they expected a box by looking longer at the surprising event. The researcher (Wilcox, 2007) found that male and female infants at 9.5 months did not look longer at what should have been the surprising event and therefore did not understand that if a box is moved behind a screen it should be there when the screen is removed; at 10.5 months, only the boy infants looked longer at the surprising event, and by 11.5 months both girl and boy infants registered their surprise with longer looking times. Of course, if this were the only study with this finding, it would need extensions and replications before it could be considered as good evidence for sex differences in early event processing. Other studies using other sorts of events with infants have also reported that infant boys found some impossible results surprising at an earlier age than girls (e.g., Schweinle & Wilcox, 2004). But, in interpreting these results, keep in mind the possibility that early differences in how boys are played with and the toys they are given could be causing these sex differences.

Thus, sex differences are found in both perception and attention—the earliest stages of information processing—and in some fine and gross motor skills, although again I urge caution in interpreting these results. Even though there is considerable evidence for some sex differences in perception and attention, it is difficult to translate findings like differential touch sensitivity and hearing thresholds into predictions about cognitive performance. A conservative conclusion is that while there seem to be perceptual and attentional differences between females and males, we can only speculate about their influence on cognitive abilities, especially for males and females in the middle range of intellectual ability—the portion of the abilities distributions where most people (by definition) belong and where differences on most tasks are the smallest.

What Sex Differences in Perception Mean and How They Have Been Distorted

> Put all of this information together, and it's clear that at birth, boys and girls do not differ dramatically in their perceptual abilities.
>
> —Lise Eliot (2009, p. 61)

Reading through a long list of ways that males and females differ in perception can create a strong belief that women and men are living in separate worlds where sounds have different qualities, colors have different hues, and food has different tastes. But, in reality, all of the sex differences in perception are quite small, even for the most reliable ones such as differences in the ability to smell faint scents, remember odors, or estimate time intervals. Even more importantly, these differences have no effects for the vast majority of people in their day-to-day lives. Consider the finding that, on average, females can detect softer sounds than males can. First, it is critical that we remember that all of these differences are average differences, with wide individual differences. But even more importantly, these data do not support the idea that males and females have different perceptual worlds. Second, it is easy to lose sight of the many ways in which males and females do NOT differ in their perceptual abilities. Consider, for example, a study of sensitivity to and preference for the taste of calcium (Leshem, Katz-Levin, & Schulkin, 2003). Researchers believed that women would be sensitive to the taste of calcium and would prefer it relative to males because it is essential for growth and reproduction. These beliefs did not hold up empirically because there were no sex differences in either the preference for or the ability to detect low concentrations of calcium. This example with calcium is intended to show that there are many more ways in which males and females are similar in perception than there are ways that they differ.

Unfortunately, the data on sex differences in perception have been misused to make statements like girls and boys need different lighting in their classrooms or that we need to talk to girls in a softer voice than we should use with boys or that girls need the temperatures in their classrooms 6 degrees warmer than boys do! These are the claims of some popular writers who advocate sex-segregated education (e.g., Sax, 2005). These statements show a misunderstanding of the perception literature. Perception thresholds are measures of the smallest amount of stimulation (e.g., intensity of a sound) that is needed for someone to just be able to perceive it. Girls are often able to perceive stimuli (a light or a sound) at lower levels of stimulation than boys. Differences in absolute thresholds (minimum amount of stimulation needed for detection) do not mean that boys and girls live in perceptually different worlds. If a constant level of sound is used in normal conversations both sexes will recognize that it is a normal talking voice, and it will NOT sound like shouting to girls or whispering to boys (which is what some advocates for single-sex education claim—see the National Association for Single Sex Public Education website for a misunderstanding of these data, http://www.singlesexschools.org/research-learning.htm). In addition, almost all studies of differences in perception thresholds have been conducted with adults and there is no reason to believe that they would be found in children because there are many developmental differences in perception between adults and children. What is even more important is that there is so much overlap in perceptual thresholds between males and females that any attempt to sort people on the basis of their perceptual thresholds would result in large percentages of both females and males in any grouping. Liberman (2006) calls claims that girls and boys need different sorts of educational experiences based on supposed differences in perception "neuroscience in the service of sexual stereotypes." The claims that boys and girls need different types of education based on sex differences in perception are based on faulty reasoning and, in some cases, fiction. Bottom line: There are some differences in absolute thresholds and other perceptual measures for females and males (e.g., males are more likely to be color-blind; Rodriguez-Carmona, Sharpe, Harlow, & Barbur, 2008), but these differences do not mean that they see, hear, or process information differently or that girls and boys need different learning environments.

A COGNITIVE ABILITIES APPROACH

As stated earlier, there are no practical differences in the scores obtained by males and females on intelligence tests. Sex differences are, however, found in the subscores on intelligence tests. Intelligence tests are comprised of several subscores, each presumably reflecting a separate cognitive component that also requires some general intelligence. One of the most widely used intelligence tests was devised by David Wechsler. The adult version is known as the Wechsler Adult Intelligence Scale (WAIS), and the children's version is the Wechsler Intelligence Scale for Children (WISC). The newest version of the WAIS is designated as "WAIS-IV." It yields four subscores of intelligence and an overall IQ score which does not show sex differences. The four subscores, which are usually called scales, include (a) a Verbal Comprehension Score comprised of scores on verbal subtests (e.g., similarities, vocabulary, information and comprehension; (b) a Working Memory Score (remembering digits, arithmetic, and letter–number sequences); (c) a Perceptual Reasoning Score (making block designs, matrix reasoning, visual puzzles, picture completion, figure weights); and (d) Processing Speed Score (symbols search, coding, cancellation) (Pearson Assessment, 2008).

Sample items from three of the subtests that are new to the latest edition of the WAIS are shown in Figure 3.7.

The older versions of the WAIS showed sex differences favoring females on a verbal subscale and sex differences favoring males on a performance subscale. The newest version now has four subscales and at the time of writing this version is so new that there are not many published studies comparing women and men on the new subscales. One study found small advantages favoring men at ages 16 to 64 on Verbal Comprehension, Perceptual Reasoning, and Working Memory and a small advantage (although it was larger than the others) favoring women for Processing Speed (Salthouse & Saklofske, 2010). The between-sex differences are so small that they are not likely to be meaningful.

MEMORY

There is a long history of research showing a female superiority on many measures of memory. For example, a review article published in 1927 reported that "women are superior in all forms of memory" (Allen, 1927, p. 297). There are many varieties of memory, which means that cognitive psychologists do not think of memory as a unitary construct. Because there are many different types of memory, no single test can correspond to memory in general, and any conclusions about sex differences in memory will have to be modified to make it more specific to the task used to assess memory and what we believe to be true about the underlying cognitive processes. In an examination of sex differences in memory, Stumpf and Jackson (1994) analyzed a battery of tests that each assesses different aspects of memory. Their subjects were medical school applicants in Germany over a 9-year period. They found that women were substantially better on a battery of tests of memory than men (taken together, effect size, d = 0.56, over a half of a standard deviation). The authors of this study explained that memory is usually not studied in the context of sex

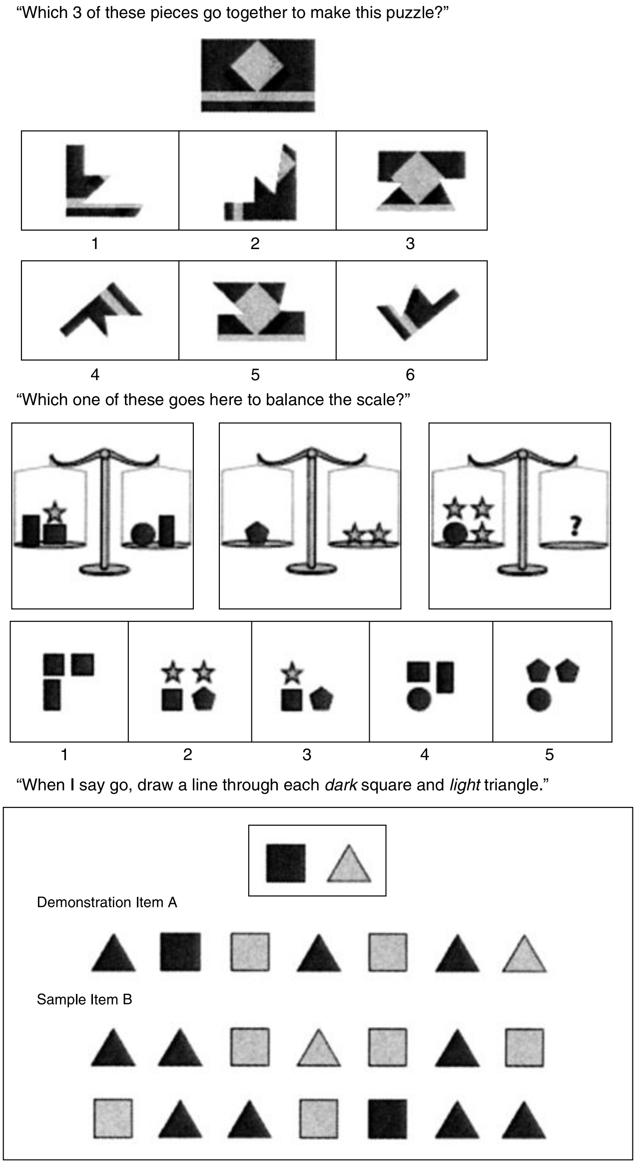


Figure 3.7 Sample items from three subscales on the WAIS-IV. Figure from http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=015-8980-808&Mode=summary Wechler Adult Intelligence Scale, Fourth Edition (WAIS-IV).

differences because it is not a single concept (in the jargon of cognitive psychology, it is not a pure factor), a fact that makes it difficult to obtain consistent findings among studies. They believe that the size of the female advantage on memory tasks has been underestimated because the tasks that researchers use are unreliable and memory is a multidimensional construct. Stumpf and Jackson's use of a test battery corrects for some of these problems. In a later study, Stumpf and Eliot (1995) examined academically talented students in middle school and high school in the United States. In this study, they also found an advantage for females, this time on tests of visual memory. Recall from an earlier discussion in this chapter that females also have better memory than males for odors.

Episodic memory refers to memories for which someone can remember where and when they were when they learned the information that is being recalled—memories that include time and place of the information remembered. For example, if you can recall that women have better memory for odors than men do and that you just read about this, then this is an example of an episodic memory. But memory is one part of a complex cognitive system, which means that the likelihood that you remembered that women tend to have better memory for odors depends on what you were doing when you read that section. If you thought it was an odd or interesting fact, or if you related this finding to your personal experiences, it is more likely that you would remember this fact than if you are merely trying to finish reading this chapter as quickly as possible. Motives, prior knowledge, and interest are among the many variables that affect what is remembered. If you cannot recall where or when you learned something, then this sort of memory is called semantic memory. Semantic memory refers to our general knowledge of concepts, which is usually recalled without knowing when or where the information was learned. Across a wide variety of memory tasks, women have been found to have better episodic memories than men. Women are better at recognizing faces they recently viewed and names they recently heard or read (Herlitz & Kabir, 2006; Larsson, Lövdén, & Nilsson, 2003; Rehman & Herlitz, 2006.) The female superiority for remembering faces can be explained, at least in part, by the finding that females across all adult ages are better at recognizing facial emotions (Sasson, Pinkham, Richard, Hughett, Gur, & Gur, 2010). A large-scale study that was conducted on an internet news site, with over 7,000 participants, showed that the generally better ability to recognize emotions as expressed in faces is a likely explanation for women's better memory for faces. It seems that the faces may be more distinct when the viewer can infer the correct emotion and processing facial emotions makes the faces easier to recall.

There are a variety of different measures that show that female college students have better memory for speech that they have heard than male college students do (Ely & Ryan, 2008). Cognitive psychologists use the term "autobiographical memory" for the kind of memory that pertains to memory for events in one's own life. When college students were asked to recall speech from their past—words that were spoken to them—women reported that they recalled many more instances of speech in their own personal lives (e.g., the words spoken when a parent told them the parents were divorcing). Women also reported younger ages for their earliest memories for feelings and emotions. In general, women recall more memories from childhood, and they report a younger age for their first memory than men do, with 7.78 years of age for women's earliest memory and 8.66 years of age for men's earliest memory (Davis, 1999). Of course, reports about one's memory are not the same as actual tests of memory, and it is possible that even though the women reported more early memories than the men this is a difference in how people report what they remember and not in memory per se.

As part of a national study in the United States with midlife adults, researchers found that women had better short-term memory than men (Pearman, 2009). Short-term memory was defined in this study as memory for events that are (approximately) up to 30 seconds old. The researchers measured short-term memory using a standard procedure that is part of the Wechsler Intelligence Scale that requires the participant to repeat lists of digits, first in the same order in which they are spoken (e.g., 5 9 4 3 0 7 7) and then to repeat them in the reverse order in which they were spoken (e.g., a participant would say 7 7 0 3 4 9 5 if she heard these digits in order starting with 5). In general, women recall more digits than men do using this research paradigm (a paradigm is a research procedure).

Jensen's (1998) extensive review of multiple tests showed that females scored higher on tests of short-term memory, with an effect size, *d*, between 0.20 and 0.30, depending on the nature of the test. These results have also been found with a sample of Chinese high school students where the girls had larger word spans (short-term memory for words), $d = 0.54$, and larger working memories, $d = 0.35$, than the boys (Huang, 1993). (Working memory refers to the processes used when both remembering something and processing information—such as remembering a list of numbers while also answering questions about an unrelated topic.)

Females also have better memories for spatial locations. This is the conclusion from studies by Eals and Silverman (1994), who believe these data reflect their evolutionary origins from hunter-gatherer societies in which females needed good memory for the location of plants in their role as the gatherers. This study was updated recently using participants from 40 different countries who logged into a BBC (British Broadcasting Corporation) website that collected data on sex-related spatial competencies (Silverman, Choi, & Peters, 2007). In 35 of the 40 countries, women scored higher than men on a test of their ability to remember where an object was located. In a meta-analysis, which is a statistical review of many different studies, Voyer, Postma, Brake, and Imperato-McGinley (2007) analyzed the results from 36 different studies on memory for objects (remembering, for example, if participants saw a shoe or a house) and memory for object location (where an object was shown, for example, in the upper left corner of a photo). Women showed better memory for objects and object location (with few exceptions, such as memory for uncommon objects) from age 13 through adulthood. Thus, the female advantage in memory is found across a variety of task and ages.

VERBAL ABILITIES

> Women appeared to perform relatively well with a format that requires written responses.
>
> —Bruce Bridgeman and Gordon McHale (1996, p. 16)

Evidence from a variety of sources supports the finding that, on the average, females have better verbal abilities than males, but the advantage is likely to be small and depends on the type of verbal ability that is measured. Like the other cognitive abilities, “verbal abilities” is not a unitary concept. The term applies to all components of language usage: word fluency, which is the ability to generate words (both in isolation and in a meaningful context), grammar, spelling, reading, writing, verbal analogies, vocabulary, and oral comprehension. There is also strong neurological evidence that separate brain subsystems are involved in generating language, comprehending language, using grammatical rules, and in producing and decoding speech sounds (Gazzaniga, Ivry, & Mangun, 1998). The size and reliability of the sex difference depend on which of these aspects of language usage is being assessed. Consider the various verbal questions that are shown in Figure 3.8. As you can see, they tap related but somewhat different abilities. Much of the confusion in the literature comes from the failure to distinguish among language tasks, some of which show no sex differences while others show large sex differences. When sex differences in verbal abilities are found, they virtually always show better performance by females.

According to a study released in 2010, there is “good news for girls and bad news for boys . . . overall male students in every state where data are available lag behind females in reading” (Robelen, 2010). Rivers and Barnett (2010) remind us that we should not “read too much into boys’ verbal scores” because a closer look at the data shows a pattern of results that provides a context for understanding these and similar data on differences between boys and girls. When the data are analyzed separately by race and social class, the data show a different picture. White and Asian boys in suburban schools are not behind in reading, and they do not drop out of school at high rates. Black and Hispanic boys, especially those in urban schools, do more poorly, but so do Black and Hispanic girls in poor-performing schools. Among White boys, those in rural areas and those in poverty also are behind in reading. Thus, what appears to be a simple sex difference is really more complex. In general, boys’ reading achievement has been improving, but at the same time, girls are improving at

TESTS OF VERBAL ABILITIES

1. Name as many words as you can that start with the letter "k."

2. Select the word that is most nearly the same in meaning:

   VIVACIOUS
   A) HONEST
   B) MEDIOCRE
   C) LIVELY
   D) BRAT

3. IGLOO: INDIAN :: TEPEE:

   A) ICE B) CANVAS C) ESKIMO D) HOME

4. Answer the questions based on the information provided in this passage.

   The literature with regard to sex differences in verbal abilities has been mixed with some researchers reporting large differences and others reporting no statistically significant differences. It seems that the controversy can be resolved by looking at the types of verbal tasks in which differences are found and determining how they differ from tasks in which differences are not found. It may be that tasks like solving verbal analogies are more similar to mathematical problem solving than to some of the other verbal tasks.

   a) What is the "controversy" that is referred to in the second sentence?

   b) Why does the author suggest that verbal analogies are similar to mathematical problems?

5. Which is correct?

   a) Give the money to Bob and I.

   b) Give the money to Bob and me.

6. Recall the words in a word list (e.g., book, sign, worry, sleep, justice, railroad, money, diamond, child, hospital, movie, lamp) or recall the objects in a room you were in recently.

7. Write a story about growing up as an immigrant in a foreign country.

Figure 3.8 Tests of verbal ability. Each of these tests may be tapping a different type of verbal ability.

an even faster rate. Although there is much to be concerned about in these data, the simple idea that boys are failing at reading or any other verbal skill is plain wrong. This conclusion is supported by a meta-analysis conducted by Hyde (2005) in which she concluded that the female superiority in verbal abilities is so slight that it is meaningless.

What about sex differences at the highest ability end of the distribution for verbal skills? The answer to this question was presented earlier in this chapter where sex differences in tails of distributions were discussed. Wai, Cacchio et al. (2010) answered this question using data from seventh graders who took college entrance tests as part of the screening process for a program for academically precocious youth. They examined data from the SAT-W, which is a writing test that was introduced in 2005, the SAT-V and SAT-Test of Standard Written English. For most years, the ratio of males to females was approximately equal among the highest scorers on the SAT-V. Wai et al. concluded that the female advantage in writing is the most robust, peaking at 2.38 females to every male among test-takers who scored above 700.

Scores of females and males on the ACT-English test are shown in Figure 3.9. On average, females outperform males on this college admissions test.

There are numerous indicators of sex differences in verbal abilities when we consider the low end of the verbal abilities distribution. Simply stated, "the overwhelming majority of children in special education today are boys" (Meyerhoff, 2008, para. 2). Boys are classified as learning disabled at approximately twice the rate of girls and as emotionally disturbed at 4 times the rate of girls, two factors that are probably related (Henning-Stout & Close-Conoley, 1992).

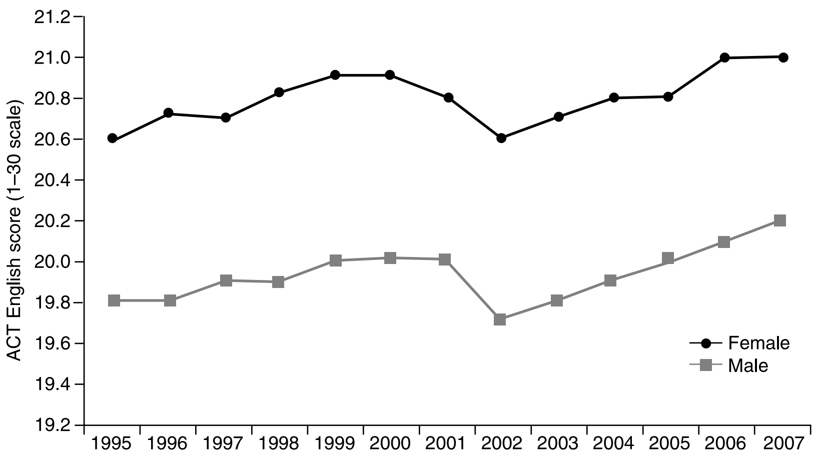


Figure 3.9 Mean scores for males and females on the ACT-English test, which is commonly used for college admissions decisions. Unpublished data provided by the ACT Statistical Research Department. Reprinted with permission of ACT, Inc.

Stuttering, a disability in the production of fluent speech, is overwhelmingly a male problem. Estimates about the percentage of children and adults who stutter vary, probably because of differences in who is sampled and how stuttering is defined. For example, one researcher reported that stuttering affects more than 15% of children between 4 and 6 years of age, but this drops to between 1% and 2% for adults (Gordon, 2002). According to the National Institute on Deafness and Other Communication Disorders (NIDCD, 2010), approximately 5% of all children stutter at some time in their childhood. Many sources on stuttering agree that there are 3 to 4 times more male stutterers than female stutterers (NIDCD, 2010). Other experts have estimated sex ratios in stuttering to be as high as 10:1 (Woodruff-Starkweather & Givens-Ackerman, 1997). Thus, the exact ratio of males to females depends on how stuttering is defined, but all measures clearly show that stuttering is overwhelmingly a male disorder.

There are also sex differences in other measures of speech disability. For example, men are more likely to exhibit aphasia (impairment in producing oral speech) when they have a stroke than women are (Di Carlo, Lamassa, Baldereschi, Pracucci, Basile, & Inzitari, 2003). A stroke occurs when a blood vessel ruptures and results in a loss of oxygen to the brain. In general, women experience their first stroke at an older age than men do, so any findings about sex differences in strokes is also confounded by age. Although earlier research showed that women were better able to regain language following strokes and brain surgery (Witelson, 1976), this finding has not held up in more recent research (Cloutman, Newhart, Davis, Heidler-Gary, & Hills, 2009), so any conclusion about the recovery of speech following a stroke will have to await more research.

Boys are much more likely to have reading disabilities than are girls (Rutter et al., 2004). Dyslexia, a severe reading disability found in individuals whose other cognitive abilities are within normal ranges, is also predominantly a male problem, with most samples showing at least twice the incidence of reading disorders for boys as compared to girls. These conclusions are from research conducted at four different sites in New Zealand and the United Kingdom. Although approximately 2% of the school population is dyslexic, mild dyslexia is 5 times more likely to occur in males than in females, and severe dyslexia is 10 times more likely to appear in males than in females (Sutaria, 1985). The conclusion that boys are, in general, less skilled in reading is supported by international data from the Programme for International Student Assessment (PISA; Organisation for Economic Co-operation and Development, 2009, p. 79). In a cross-national study of 30 countries, 3.7% of females and 0.8% of males are in the category designated as “top performers in reading,” so girls are both more likely to be among the best in reading and less likely to have a reading disability.

The female advantage in verbal abilities can be seen in creative writing. Kaufman, Niu, Sexton, and Cole (2010) found that women’s poems were judged as more creative than men’s poems, even when the sex of the writer is unknown. These creativity researchers reviewed several other studies with similar conclusions: women write more creative poems than men do, but scores on general creativity tests show few differences (most of which favor females), suggesting that the advantage that females have in creativity is most probably limited to verbal tasks (Bauer & Kaufman, 2008).

The research evidence from a variety of sources favors female superiority on verbal tasks including reading and speaking and is largest in the high and low ends of the distribution. Despite the finding that females score higher on at least some tests of verbal ability, the overwhelming majority of critically acclaimed writers are male. Other careers and prestigious professions that require advanced verbal abilities, careers like lawyer, politician, and journalist, are also predominantly male. Adelman (1991) noted this disparity in a report for the U.S. Department of Education, called the "paradox of achievement," and laments the economic loss to the United States created by the underdevelopment of women's intellectual potential. In his own words:

> The paradox of this story—that women's educational achievements were superior to those of men, but that their rewards in the labor market were thin by comparison—is set in the context of national economic development. (p. v)

There are several possible reasons for the discrepancy for women between their abilities and their achievement. It is possible that women are not using their talents as frequently as men, or the tests are not measuring high-level creative ability, or differential criteria are being used to judge women's and men's writing. It is interesting to note that several outstanding women writers such as Dickinson and the Brontes were single women with other means of support. If ability is only a small part of eminence, then the lack of eminent female writers is not surprising.

Age Trends in Verbal Abilities

Sex differences in some verbal abilities appear early in life. According to Cole and Cole (2001), children learn to use 200 to 300 words by age 2. Between 16 months and 30 months of age, girls lead boys in the number of words they can say by about one month of development (Fenson, Dale, Reznick, Bates, Thai, & Pethick, 1994). Another study provided a somewhat higher estimate of girls' early vocabulary development, with 2-year-old girls using an average of 275 words, whereas boys use an average of 197 words (Lutchmaya, Baron-Cohen, & Raggatt, 2002). Girls also show better language skills in preschool (e.g., Blair, Granger, & Razzam, 2005). Based on a review of 24 large data sets (including several large representative samples of U.S. students, working adults, and military personnel), Willingham and Cole (1997) concluded that differences are small in the elementary school grades, with only writing, language use, and reading favoring females at fourth grade, $d > 0.2$. In the United States, by the end of high school, the largest differences, again favoring females, are found for writing (d between 0.5 and 0.6) and language usage (d between 0.4 and 0.5).

Another report on writing proficiency for children in Grades 4, 8, and 11 in 1984, 1988, and 1990 showed that girls were better writers in each of the nine comparison groups (U.S. Department of Education, 1997). More recently, the 2007 Nation's Report Card reported that females are 20 points ahead of males in writing in eighth grade (d = 0.4) and 18 points ahead in 12th grade (d = 0.36; National Assessment of Educational Progress, 2008). After a comprehensive review of the literature on writing skills, Hedges and Nowell (1995) concluded: "The large sex differences in writing . . . are alarming. These data imply that males are, on average, at a rather profound disadvantage in the performance of this basic skill" (p. 45). This conclusion is supported by data from the U.S. Department of Education (2000) that show that girls in their senior year of high school are approximately 36 months ahead of boys in writing skills.

Skillful writing is a generative activity that includes good organization of ideas, grammatically correct constructions, and accurate use of words. The conclusion that females excel in writing is bolstered by data released by the U.S. Department of Education on writing proficiency tests given at Grades 4, 8, and 11 in 1984, 1988, 1990, and most recently in 2007. The data from 2007 are graphically presented in Figure 3.10. The graph shows that girls in eighth and twelfth grades are writing better than same-age boys.

In a detailed investigation of language development among children aged 2½ to 4 years, Horgan (1975) examined the mean length of utterances (MLU, the average number of words strung together in a single utterance) for girls and boys.

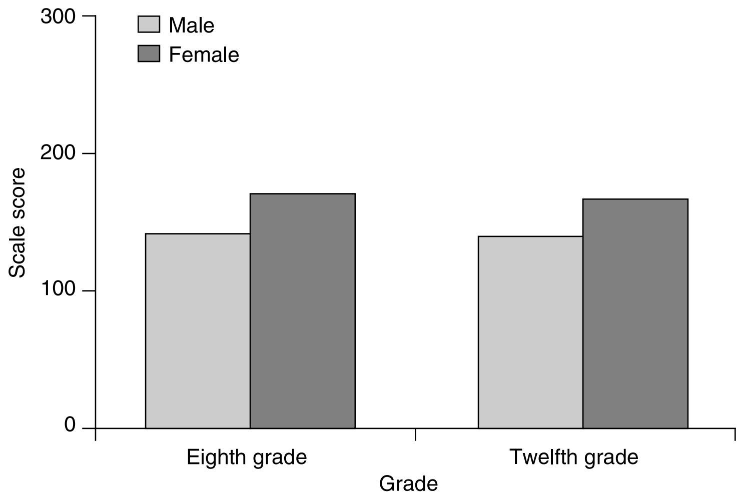
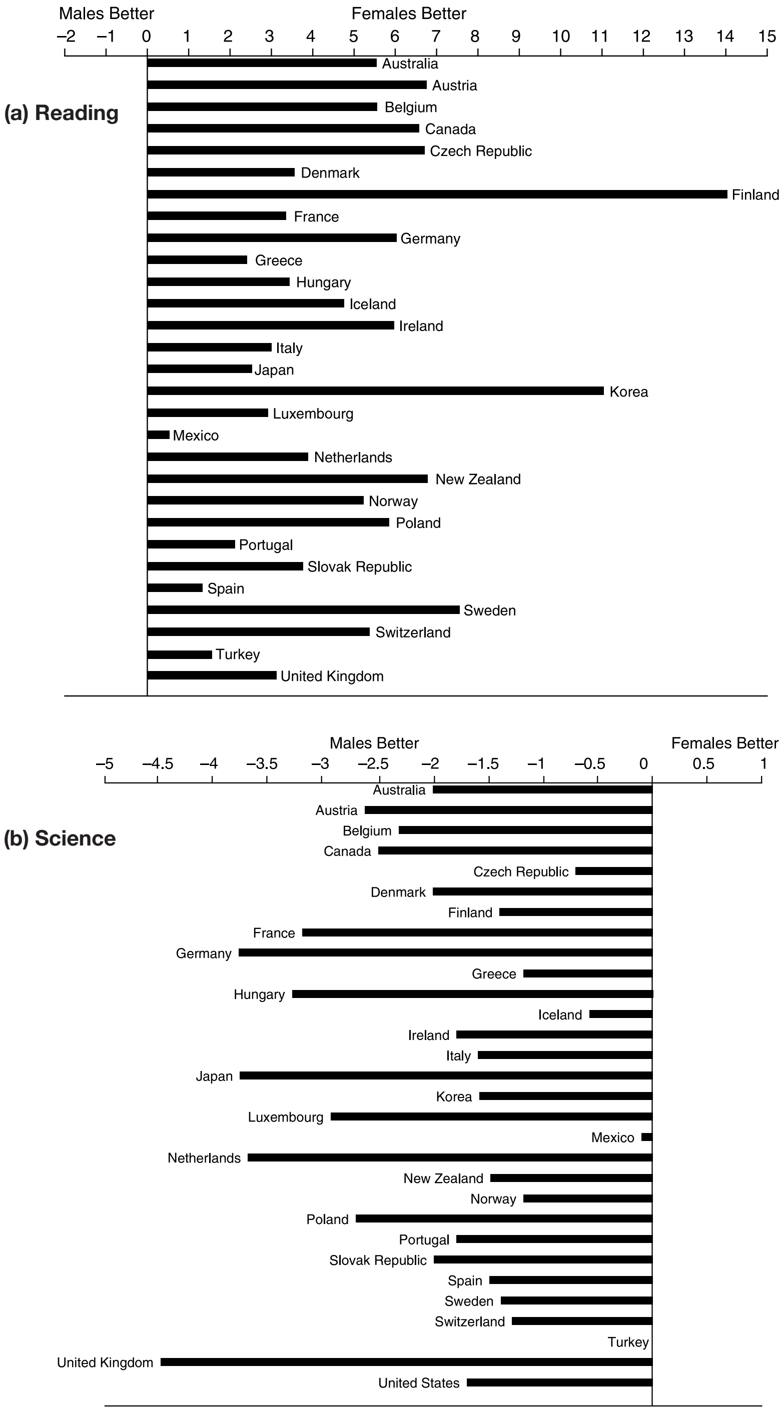


Figure 3.10 Writing proficiency scores for males and females in eighth and twelfth grade. These data correspond to effect sizes of d = 0.40 for eighth grade and d = 0.36 for twelfth grade. Data from U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (2007a).

She argued that MLU is a good indicator of linguistic maturity for preschool children who are learning their first language. Horgan reported that prior to MLUs of four words, boys and girls perform equally well; however, sex differences favoring girls occur beyond MLUs of four words (i.e., girls use longer utterances at younger ages than boys). Horgan also analyzed other indicators of linguistic maturity including: use of the passive voice (e.g., The lamp was broken); truncated passive (e.g., The window's broken.); and use of participles (verbs used as adjectives—e.g., The moving truck crashed). Girls spontaneously generated all of these advanced linguistic forms at an earlier age than males; furthermore, they made fewer errors in language usage overall. Horgan concluded: "Girls produce longer utterances at younger ages, they produce more varied constructions, and they make fewer errors" (p. 48). In a more recent study with Swedish children, Lundberg (2009) studied phonological awareness in 1,100 6-year-old children in preschool. Phonological awareness is knowledge about the relationship between letters and their sounds. If children do not understand the relationship between letters and their sounds, they are at great risk for reading failure. Lundberg found that 19% of the boys, but only 7% of the girls, were among the children with the poorest performance. By contrast, the group with the highest level of performance contained only 14% of the boys, but 29% of the girls. After training the children in phonological awareness for 8 months, there were very few children with low scores, but the highest performing group was 73% of the girls and 47% of the boys. In general, the majority of the evidence tends to support the idea that young girls are more verbally precocious than young boys. Of course, it is important to remember that these are group averages, and almost half of all boys reached the highest levels of phonological awareness after 8 months of training.

Female superiority on verbal tasks may seem reminiscent of the stereotype that females talk more than males, but it is the quality of the speech produced and the ability to comprehend or decode language that is being assessed, not merely the quantity. Some readers may be thinking about the well-known finding that females talk about 3 times as much as males do. This statistic has been repeated on television (CBS and CNN), National Public Radio, *Newsweek*, in *The New York Times*, and *The Washington Post* (Mehl, Vazire, Ramirez-Esparza, Slatcher, & Pennebaker, 2007). The problem with this common knowledge is that it is wrong. To test the difference in words spoken per day, Mehl and his colleagues fitted men and women with recording devices that were automatically activated when there are speech sounds. The real numbers are that women speak an average of 16,215 words a day compared to 15,669 words a day for men. This difference was statistically significantly different, but so small ($d = 0.07$) that it is not meaningful!

The general superiority of girls in reading can be seen internationally in Figure 3.11, which shows the difference in the percentage of top performers between females and males in international tests. The differences in the percentages of top performers in science and mathematics are shown in Figures 3.11b and 3.11c.


Males Better
Females Better
–2 –1 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
(a) Reading
Australia
Austria
Belgium
Canada
Czech Republic
Denmark
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Italy
Japan
Korea
Luxembourg
Mexico
Netherlands
New Zealand
Norway
Poland
Portugal
Slovak Republic
Spain
Sweden
Switzerland
Turkey
United Kingdom
Males Better
Females Better
–5 –4.5 –4 –3.5 –3 –2.5 –2 –1.5 –1 –0.5 0 0.5 1
(b) Science
Australia
Austria
Belgium
Canada
Czech Republic
Denmark
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Italy
Japan
Korea
Luxembourg
Mexico
Netherlands
New Zealand
Norway
Poland
Portugal
Slovak Republic
Spain
Sweden
Switzerland
Turkey
United Kingdom
United States

(c) Mathematics

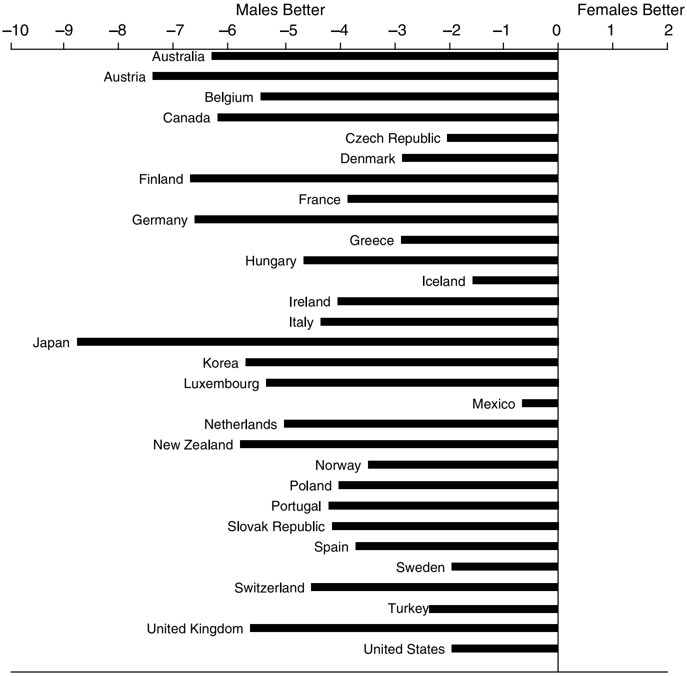


Figure 3.11 Bar chart showing the difference in the percentages of top performers between girls and boys. Data from *Education at a Glance 2009*, Organisation for Economic Co-operation and Development (2009). Data are from Table A4.1b. Note that negative numbers indicate a higher percentage for males and positive numbers indicate a higher percentage for females.

In an extensive meta-analytic review of the literature on sex differences in verbal ability, Hyde and Linn (1988) divided experiments based on the age of the subjects and type of verbal ability assessed—all tests, vocabulary tests, and tests of reading comprehension. Differences were found in the "all tests" category for children 5 years and younger ($d = 0.13$) and for adults over the age of 26 ($d = 0.20$), both favoring females. There were no notable differences as a function of sex for ages 6 through 25. The developmental pattern of vocabulary proficiency is difficult to comprehend. Hyde and Linn reported a male advantage in the 6- to 10-year-old age range ($d = -0.26$) and a female advantage in the 19- to 25-year-old age range ($d = 0.23$), with essentially no differences in the other age categories. The largest differences were in reading comprehension for children 5 years of age and younger, with females reading more proficiently than males ($d = 0.31$). Hyde and Lynn's meta-analytic review is now well over 20 years old and may have underestimated the size of the female advantage on a variety of verbal tasks, at least when compared with more recent studies, probably because of the types of verbal ability tests they reviewed.

In a 10-year study, De Frias, Nilsson, and Herlitz (2006) found that across age groups ranging in age from 35 to 80 years, the female advantage in episodic memory (remembering names, words, and activities, recognizing faces), vocabulary, and fluency tasks (naming words that begin with a particular letter such as “M” and naming words by category, such as “professions”) remained constant over the 10-year period of their study. These researchers also found that the size of the sex difference in these tests did not change with age. The authors concluded that “There was stability of sex differences across five age groups and over a 10-year period” (p. 574). These are tests of verbal memory, which seem to support the conclusions that on average, females excel at some memory tasks and some verbal tasks.

Although it seems that there is little change over a 10-year period in the size of sex differences, a different picture emerges when data are collected from a very large sample of adults that range in age from 20 to 60 years. The BBC sponsored an internet study of sex differences, which resulted in many different published studies of cognitive sex differences (Lippa, 2007). Almost half a million people took a variety of cognitive tests online and responded to demographic questions. The treasure trove of data has provided a unique opportunity to answer many questions about cognitive sex differences and results from this study appear in several places in this book. In one set of analyses, researchers examined sex differences on memory for the location of an object and category fluency (e.g., naming as many members of some category, such as animals, as possible in one minute). In one of these studies, the researchers (Maylor, Reimers, Choi, Collaer, Peters, & Silverman, 2007) found the usual advantage for females with these tasks ($d = 0.33$ for memory for location and $d = 0.18$ for fluency), but when they looked across the age range they found two important results: (a) performance declined with age; and (b) the decline was greater for males. As you will read in the next section, men also showed greater decline with age on other types of tasks.

VISUOSPATIAL ABILITIES

> Too many jokes to recount here are made about the spatial skills of women versus men. While once good-humored, they now take on a social significance that becomes lost in current social values.
>
> —Michael Gazzaniga, Richard B. Ivry, and George R. Mangun (1998, p. 507)

The term “visuospatial abilities” may not convey much meaning to people who are not cognitive psychologists. In fact, it is not an easy term to define because it is not a unitary concept. Linn and Petersen (1985) provided this definition: "*Spatial ability* generally refers to skill in representing transforming, generating, and recalling symbolic, nonlinguistic information" (p. 1482). Generally, it refers to the ability to imagine what an irregular figure would look like if it were rotated in space or the ability to discern the relationship among shapes and objects. The ability to utilize spatial relationships is an important aspect of human thought. Visuospatial skills (spatial skills that are visual in nature) are used extensively in engineering, architecture, chemistry, the building trades, and air crew selection (Lohman, 1988). After reviewing the literature on visuospatial ability, Cooper and Mumow (1985) concluded, "The spatial aptitude literature is quite clear in showing that a broadly defined spatial factor exists independent of verbal and quantitative factors and that this spatial factor is more effective than other measures of intelligence in predicting success in certain academic and industrial areas" (p. 71).

Five Categories of Visuospatial Abilities

In 1985, Linn and Peterson used factor analysis, a data analysis technique that finds commonalties in the data, on a number of tests of visuospatial abilities. The results clustered the tests into three main categories—spatial perception, mental rotation, and spatial visualization. Although numerous new tests of visuospatial ability have been used since then, and hundreds of different tests have been identified as measures of visuospatial abilities, the three categories that they identified plus two others—one that involves movement through space and another that involves the generation and maintenance of visual images—are a good organizing framework for understanding the literature in this area. If you are wondering what "generation and maintenance of visual images" means, try this demonstration: Think about a lower case letter "B." Is the round portion of the "B" to the left or to the right of the vertical line that forms the other portion of the letter? In order to answer this question, you had to generate a visual image of a lower case "b" and then maintain that image in memory while answering questions about its appearance. This is an example of generating and maintaining a visual image.

I have included memory for visuospatial information on this list of five categories of visuospatial tasks, even though it was also discussed in the section on sex differences in memory. All of these tasks involve several types of memory, so they fit into more than one cognitive category. It seems that there are at least five qualitatively different types of visuospatial ability:

1. Spatial perception, requires subjects to locate the horizontal or the vertical in a stationary display, while ignoring distracting information. Examples are the Rod and Frame Test, which requires subjects to position a rod within a tilted frame so that it is either vertical or horizontal, and the Piaget

Water-Level Test, which requires subjects to draw in the water level of a tilted glass that is half filled with water. An example of the Water-Level Test is shown in Figure 3.12.

One test that is sensitive to sex differences is the "Water-Level Test" originally devised by Piaget and Inhelder (1956). In one version of this test, the subject is shown a bottle partially filled with water and is told to notice the way the water fills the bottle. The subject is then asked to predict where the water will be when the bottle is tipped. Piaget and Inhelder believed that the relevant knowledge about the horizontal would be attained at an average age of 10 years. The Water-Level Test as originally conceptualized by the developmental psychologist Piaget was never intended to test anything about water per se. It was meant to be a task of spatial concepts—in this case the ability to use a Cartesian coordinate system to represent space—but increasingly contemporary researchers discuss it as a test about the fact that the surface of water remains horizontal despite the tilt of its container, thus the meaning of this test has drifted since it was originally devised by Piaget. It seems that girls demonstrate this principle at a later age than boys. In fact, it has been estimated that 40% of college women don't know the principle that the water level remains horizontal. This is a surprising result that has been replicated many times (Wittig & Allen, 1984). Robert and Chaperon (1989), for example, reported that 32% of college women and 15% of college men failed the

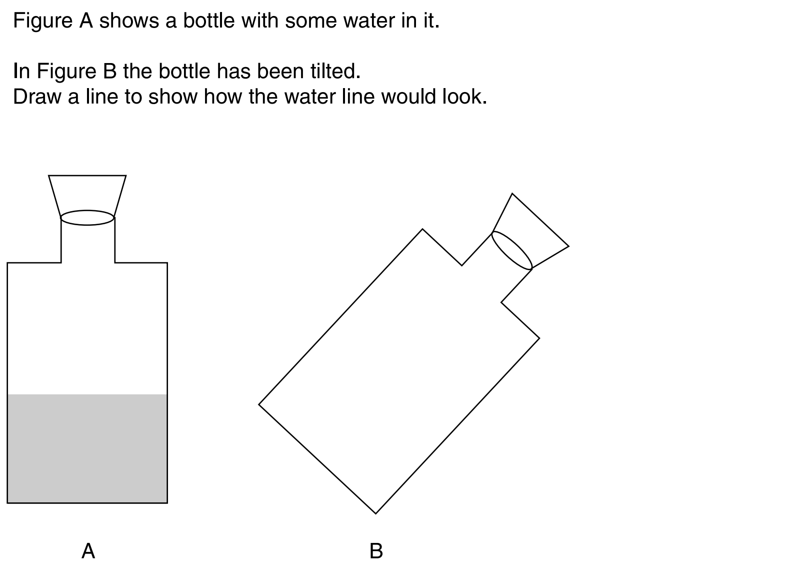


Figure 3.12 An example of the Water-Level Test. The task is to draw in the top of the water level in Bottle B, assuming that it is Bottle A tilted on its side.

Water-Level Test. Sex differences in the Water-Level Test have been confirmed internationally, for example, with a sample from Bombay, India (De Lisi, Parameswaran, & McGillicuddy-De Lisi, 1989). Vasta and Liben (1996) reported effects sizes that range between $d = 0.44$ and $d = 0.66$. It is difficult to understand why this should be such a formidable task for college women.

Results from the Water-Level Test are strange. Why should women (in many samples college women were used as participants) perform less well on a test of whether water remains horizontal in a tilted glass? As discussed in Chapter 6, at least part of the sex differences we find with spatial tasks can be attributed to differential learning experiences, with boys typically engaging in more spatial activities. Sex differences in the Water-Level Test are not amenable to this sort of explanation as no one believes that boys have more experiences than girls with glasses of water. In one study, Hecht and Proffitt (1995) hypothesized that experience with liquid surfaces would be associated with poorer performance on the Water-Level Test because people who work frequently with liquids in containers may have adopted a perspective that was relative to the tilt of the container—in other words they paid attention to the orientation of the container and not the level of the water surface. In a test of the hypothesis that more experience would lead to poorer performance on the Water-Level Test, Vasta, Rosenberg, Knott, and Gaze (1997) found the reverse results: subjects with more experience performed better than those with less experience. Thus, the poorer performance of females on this test remains unexplained.

Kalichman (1989) investigated the possibility that the results reflect some idiosyncrasy of the test, rather than sex differences in either the knowledge that water remains horizontal or the ability to draw an approximately horizontal line. Kalichman devised a more "ecologically valid" (i.e., more like the real world) test in which the tilted glass was held in a human hand. An example of his stimuli is shown in Figure 3.13.

Kalichman found that significantly fewer college women than college men draw an approximately horizontal line to indicate the water level in both the standard test format and in the human context format. He concluded that "sex differences on the water-level task remain robust regardless of task context" (p. 138). The Water-Level Test was used in a study of 1,704 participants ranging in age from 4 to 95 years (Tran & Formann, 2008). They used depictions of eight round bottles tilted at nine different degrees. The authors found that performance was best for adults between the ages of 16 and 60, with considerably lower performance at younger and old ages. They also reported a sex difference, which they claim was significant from adulthood into old age, but it was significant at young ages only for certain degrees of tilt. But a careful look at their data suggests that there was a floor effect at the younger ages, meaning that both girls and boys were performing so poorly at these ages that it would not be possible to tell if there was a sex difference. (See Chapter 2 for a discussion of floor and ceiling effects and how they can mislead researchers into

Figure 3.13 The Water-Level Test embedded in an ecologically valid (i.e., real-world) context. The glass is half-filled with water. Draw a horizontal line across the glass to indicate the top of the water line. Reprinted from Kalichman (1989) with permission from the author and Taylor & Francis.

concluding that there are no differences.) Many psychologists have studied sex differences on the Water-Level Test, perhaps because it is surprising. Correct performance on this task requires that participants understand that the surface level of water remains horizontal regardless of the tilt of the glass. As Vasta and Liben (1996) concluded in their review of this task, the puzzle is far from solved.

2. Mental rotation, includes the ability to imagine how objects will appear when they are rotated in 2- or 3-dimensional space. There are timed and untimed versions of these tests. Several researchers believe that mental rotation is a measure of a general spatial reasoning ability (Casey, Nuttall, Pezaris, & Benbow, 1995). In a recent study of college students at a highly selective school for science and engineering, the sex difference on a test of mental rotation was close to d = 0.60 (Miller, Halpern, & Saeta, 2010). A classic example of mental rotation is shown in Figure 3.14. The task is to determine which (if any) of the figures on the right can be rotated in space so that they are the same as the figure on the left.

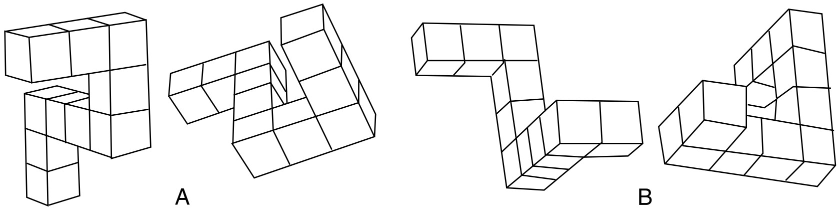


Figure 3.14 A classic mental rotation problem. For each pair of figures, determine if they can be rotated so that the two figures in each pair are identical.

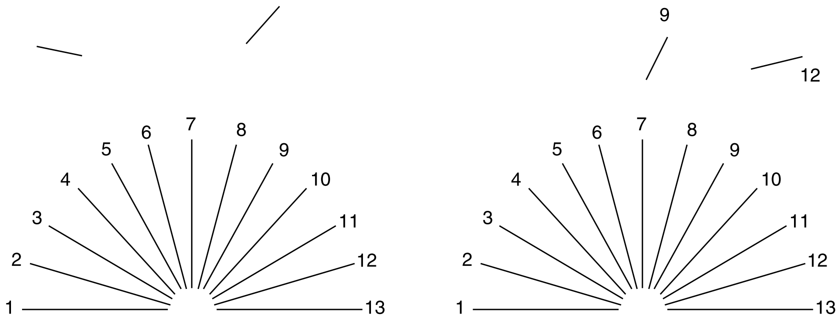


Figure 3.15 Judgment of line orientation task. Participants match each of the target line segments in the top half of each item with a numbered line from the bottom half. Answers are shown for the right-hand figure. Figure from Collaer and Nelson (2002). Copyright © 2002, with permission from Elsevier.

Sex differences in mental rotation have been studied for over 25 years and findings have been summarized in several meta-analytic reviews. A recent review of the sex differences literature on mental rotation found that male performance exceeds that of females across all age ranges, with the size of the between-sex difference ranging between $d = 0.52$ and $d = 1.49$, which increases slightly across the life span (Geiser, Lehmann, & Eid, 2008).

Another visuospatial task that shows very large sex differences and requires memory for different orientations is the judgment of line and angle orientation task (Collaer & Nelson, 2002). Look at Figure 3.15. There is a "fan" of lines at different orientations. In this task, participants examine the two lines above the fan and then indicate which of the lines in the fan matches the degree of tilt of these two lines. The answers for the lines in the figure on the right are provided. The first line matches line 9 in its orientation and the second line matches line 12 in its orientation. This task shows very large sex differences, typically at $d = 0.85$, with males showing better ability at matching the tilting line.

In a combined study of mental rotation and line angle judgments with more than 90,000 women and 111,000 men from 53 countries, men outperformed women in every country, but there were large between-country effects (Lippa, Collaer, & Peters, 2010). The many possible reasons for cognitive sex differences are discussed in following chapters; however, I note here a critical finding. Across nations, higher levels of gender equality and economic development were significantly associated with larger sex differences favoring males on both visuospatial tasks. The data from this massive cross-national study are shown in Figure 3.16. I discuss these findings in more detail in Chapter 7, where cultural influences on cognitive sex differences are considered.

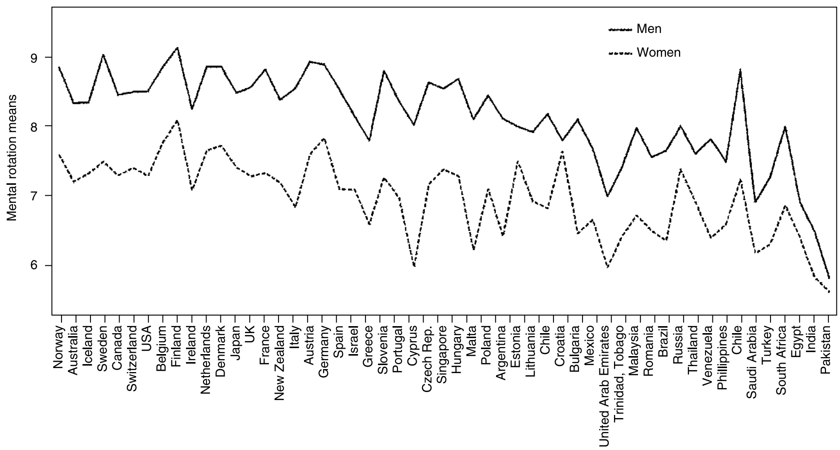


Figure 3.16 Men outperform women on mental rotation and judgment of line orientation tasks in 53 countries, but differences are largest in countries that are higher on gender equity and economic development measures. From Lippa, Collaer, and Peters (2010). Reproduced with kind permission from Springer Science + Business Media.

3. *Spatial visualization* refers to complex and analytic, multistep processing of spatial information. Tests that tap spatial visualization are the Embedded Figures Test, paper folding, hidden figures, and spatial relations test. In general, sex differences tend to be small on tests of spatial visualization, but when they are found, they tend to favor males (Miller, Halpern, & Saeta, 2010). A sample item from the paper folding test is shown in Figure 3.17.

4. *Spatiotemporal ability* involves judgments about and responses to dynamic (i.e., moving) visual displays. There are several different tasks that involve information that is moving, such as having subjects press a key when a target is coincident with a stationary line (Smith & McPhee, 1987) and making "time

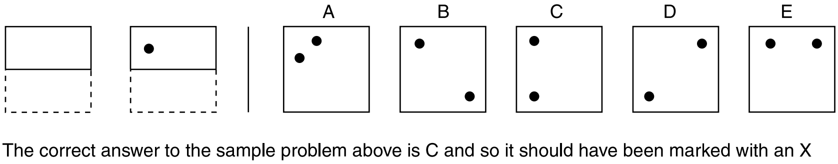


Figure 3.17 Imagine that the paper on the left is folded as shown and then has a hole punched in it. Which of the stimuli on the right shows what the paper will look like when it is unfolded?

of arrival" judgments about a moving object (Schiff & Oldak, 1990). A schematic diagram of a time of arrival task is shown in Figure 3.18. Investigators have concluded that the ability to reason about dynamic visual displays is correlated with, but different from, the abilities used in reasoning about static displays (Hunt, Pellegrino, Frick, Farr, & Alderton, 1988).

Robust sex differences favoring males are found when the task involves movement-related judgments such as judging velocity (Law, Pellegrino & Hunt, 1993). Hancock (2011) believes that sex differences in time perception are important in determining sex differences in visuospatial abilities. If, for example, women are less accurate than men in estimating time intervals, their poorer performance on dynamic visuospatial tasks may be caused by discrepancies in their estimation of the time that it will take a ball to pass behind an opaque screen. This is an interesting hypothesis, but thus far, we have no direct tests of this relationship.

Although it is difficult to isolate any single factor that might be responsible for these results, judgments concerning dynamic visual displays must involve

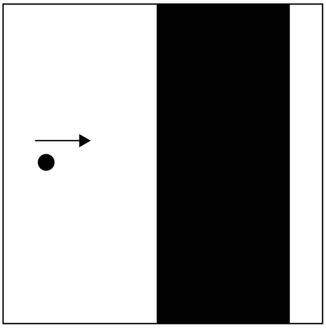


Figure 3.18 A moving ball is obscured by the solid area on a computer screen. Press a computer key when you expect it to be visible on the other side of the solid area. Reaction times are recorded.

time estimation in some way (i.e., when will the moving object reach a destination). In an effort to identify which components of a dynamic spatial task show sex differences, researchers compared performance factors, which can be defined as the components that make up a more complex task (Contreras, Rubio, Pena, Colom, & Santacreu, 2007). The researchers tested over 2,500 applicants for a training course in air traffic control. They used a computerized test of dynamic spatial ability called the Spatial Orientation Dynamic Task (SODT). It looks like a computer game in which the player can move two different colored dots using directional arrows. An example of a trial from the SODT is shown in Figure 3.19. The goal is to guide the dots toward a target. The researchers measured reaction times (how long it took the players to respond by pressing an arrow key to change the direction of a moving dot), response frequency (how many times the players pressed the arrows on each trial), and invested time (time from the first press to the last press). The main

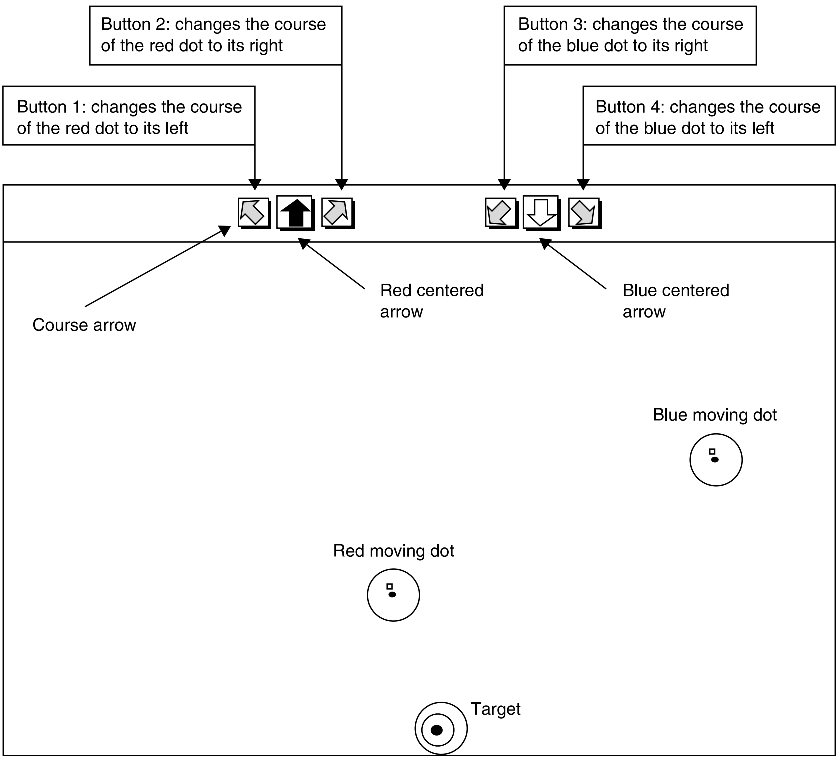


Figure 3.19 A computerized game set-up that was used to study sex differences in dynamic spatial ability. From Contreras, Rubio, Pena, Colom, and Santacreu (2007). Reproduced with kind permission from Springer Science + Business Media.

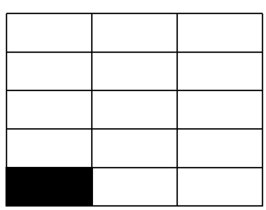
measure was the distance between the dot and the target. They found large effect sizes, favoring males. Thus, like earlier research, this new computerized measure of dynamic spatial ability showed large sex differences.

5. *Generation and maintenance of a spatial image* requires participants to generate an image (either from long-term or short-term memory) and then use the information in the image to perform a task. An example of a generation and maintenance task is shown in Figure 3.20. Performance factors in visuospatial imagery were investigated by Loring-Meier and Halpern (1999) with a set of four tasks. These tasks were developed by Dror and Kosslyn (1994) for use in a study on age-related differences in visual imagery. In one of the tasks, participants had to generate an image of a capital letter and then decide if the letter would cover a portion of a rectangular frame. A second task required participants to create a visual image of a geometric figure that had just been displayed (i.e., without retrieval from long-term memory) and then make a similar spatial judgment. A third task required participants to scan an image that they retrieved from long-term memory, and a fourth task required the mental rotation of an image. In all four tasks, the male participants were significantly faster than the female participants, with no differences in accuracy (all *d*s between 0.63 and 0.77). The faster response times for the males could be reflecting an actual difference in the time it took to perform the cognitive tasks, but it also could be reflecting greater confidence on this task. There are cognitive tasks on which females, generally, show faster responding, so these results are not simply a matter of motor speed needed to respond or a general reflection of confidence or cautiousness. Readers are asked to keep all of these possible explanations in mind as they review the theories and research presented in later chapters.

Given the large variety of tests that have been used to measure visuospatial ability, it is not surprising that sex differences depend on the type of test used. Not coincidentally, this is an area replete with contradictory findings because of the multidimensional complexity of visuospatial abilities. Caplan,

(1) Image a capital letter "b" on the grid at right.

(2)



(3) Does the capital letter "b" cover the solid square?

(4) Press one computer key for "yes" and a different one for "no." (Reaction times measured)

Figure 3.20 Generation and maintenance of an image.

MacPherson, and Tobin (1985) questioned the legitimacy of the assumption that the construct "spatial abilities" exists. They believe that the entire notion suffers from a "definitional dilemma." As noted in a response to Caplan, MacPherson, and Tobin (Halpern, 1986), much of the confusion in this area is attributable to the types of spatial ability tests used. Numerous researchers have attempted to define "spatial abilities" and "spatial thinking skills" (e.g., Chatterjee, 2008; Hegarty, 2010; National Research Council, 2006). Regardless of which definition is used, sex differences in spatial tasks are among the largest sex differences.

Visuospatial Knowledge and Memory

Spatial abilities are important in many areas of math and science. Consider for example, the pulley system shown in Figure 3.21. In order to understand how the system works, individuals must be able to visualize movement from static displays.

Caplan, MacPherson, and Tobin (1985) noted that the types of tasks that are used to assess spatial ability are fairly abstract and that a much more valid test would involve finding one's way in a real-world environment. This is certainly

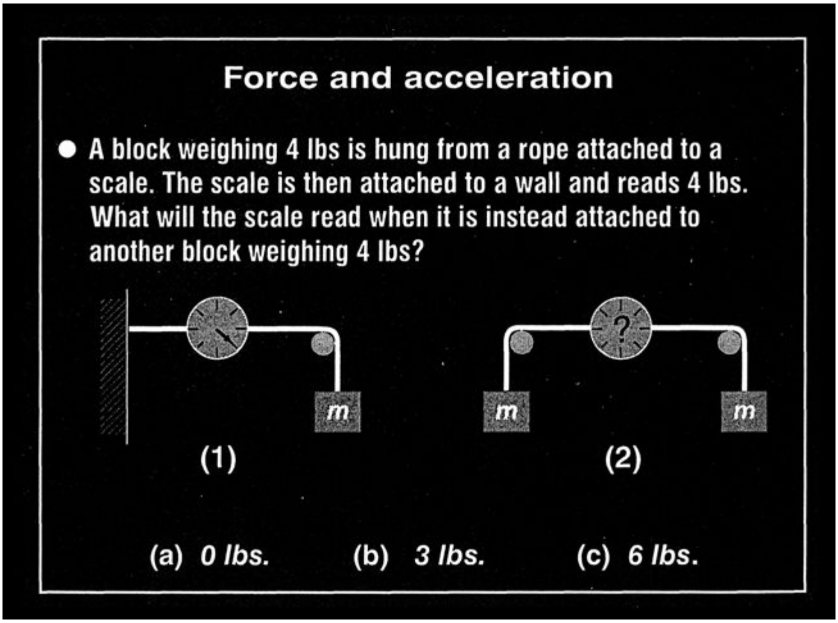


Figure 3.21 In this diagram of a pulley system, the reader needs to be able to imagine the simultaneous movement of multiple pulleys to understand the underlying physics principles.

a sensible suggestion, even though spatial ability tests that are conducted outside of the laboratory are more difficult to administer. There is always the problem that some subjects will have greater knowledge of a given geographical area. Furthermore, subjects could rely quite well on verbal strategies when they have to maneuver through a real-world space (e.g., turn at the green house). The few studies that have investigated route knowledge and "way finding" tend to support laboratory findings. In general, men learn a route from a 2-dimensional map in fewer trials and with fewer errors than a matched group of females (Galea & Kimura, 1993). Beatty (2002) wondered if some part of the sex difference in knowledge of geography could be due to differential experience with driving. In general, men drive more often than women, and it is the driver who is responsible for knowing routes. He tested this possibility with samples of teens who were too young to drive and older adults where the discrepancy in the amount of driving between women and men is the greatest. He found that driving experience was not responsible for the sex difference in geographical knowledge. Instead, it seems that women are more likely to attend to landmarks and men are more likely to use directional cues and estimate distances.

Data from the National Geography Bee tell a very compelling story. Liben (1995) estimated that 6 million school children in the United States participate in this competition. She describes "a shocking gender disparity among winners at every level" (p. 8). In 1993, of the 18,000 school winners, approximately 14,000 were boys; of the 57 state winners (including U.S. territories) 55 were boys; and in most years, all 10 finalists were boys, despite the fact that girls and boys participate at almost equal rates. Liben found that geography is not a stereotypically male domain (unlike other fields like "being a plumber" or "fixing cars"). She reported that the boys were more interested in geography and liked it more than the girls did. Furthermore, these huge sex ratios are not a fluke that is unique to samples from the United States. They are similar to those found with the International Assessment of Educational Progress that samples students from many countries. It is now over 15 years since Liben's highly publicized studies showing huge sex differences in winners of the geography bee, and there are still very few girls who make it to the final rounds of competition. A geographer, Eric Clausen, recently sued the National Geographic Society (Kolpack, 2011; Turley, 2011), alleging that in the 19-year history of the National Geography Bee, only two national winners have been girls. Clausen noted that in 2009, only 2 out of 54 state winners were girls. But, as the court and others have responded, the fact that there are very few girls winning the National Geography Bee does not mean that the competition is discriminatory. Lynn Liben, the leading developmental expert on children's understanding of space, responded that "From what I can tell at this point, the bottom line answer is that the same kinds of experiences, skills, interests and so on that lead boys to do well on the bee, also lead girls to do well on the bee. But boys have had more of those experiences" (Kolpack, 2011, para. 10).

A key question concerns possible sex differences in the ability to navigate through real space, which is sometimes called "way finding." Wolbers and Hegarty (2010) reviewed the research literature on what determines navigational abilities. According to their review, there are multiple demonstrations that males have the advantage when learning from virtual maze tasks (i.e., mazes presented on the computer) and when learning from navigating through the world. Women typically report that they use landmarks (Turn left at the bank), whereas men more often report using cardinal directions (Turn north) and distance (Turn after 3 miles). As discussed in later chapters, these differences are found in other nonhuman mammals.

Even if women, in general, use different navigation cues, there are no data to support the notion that females are less able drivers than males. In fact, all of the data suggest that the opposite is true—women have far fewer automobile accidents and auto citations than men. According to a report on insurance risk prepared by the Social Issues Research Centre (2004, p. 4): "In all studies and analyses, without exception, men have been shown to have a higher rate of crashes than women. This gender difference is most marked in the population under the age of 25 years, but is also evident among older drivers." Differences seem to be due to a greater male propensity to speed. These and similar findings about increased accidents of all sorts for males lead to this conclusion by the World Health Organization (2002): "Masculinity may be hazardous to health."

As noted earlier in the section on memory, Voyer, Postma, Brake, and Imperato-McGinley (2007) conducted a meta-analytic review of 123 different findings from 36 studies on memory for location. They analyzed separately for object identity memory, which is memory for objects that were shown, usually on a paper display, and memory for location, which is memory for where on the display different objects were displayed. They found that across studies, females performed significantly better than males on both tasks ($d = 0.23$ for object memory and $d = 0.27$ for location memory), after the age of 13. It is difficult to know if differences might occur prior to age 13 because there were few studies that included young children.

Most major reviews of the literature have concluded that males are more variable in their visuospatial performance than females (e.g., Willingham & Cole's, 1997, review of hundreds of tests, many with spatial ability components). Hedges and Nowell (1995) conducted a meta-analysis of many types of tests and also concluded that males are more variable than females in their spatial ability. The finding of greater variability in male performance on spatial tasks is theoretically important because one hypothesis about the cause of the sex difference is that many females do not use a spatial-imagery strategy to solve problems that are spatial (e.g., geometry problems). Perhaps some try to visualize an answer and others try to use verbal labels. If the sex difference in spatial ability were caused by the fact that more women than men use inappropriate strategies, then the females should show more variable performance than the males. Given the opposite finding, it seems unlikely that females use a greater variety of strategies with these tasks.

Age Trends in Visuospatial Abilities

> The male advantage in spatial abilities is evident throughout the life span.
> —Elizabeth J. Meinz and Timothy A. Salthouse (1998, p. 56)

Many of the differences in visuospatial abilities appear early in life. How early? Moore and Johnson (2008) tested sex differences in mental rotation skills in 5-month-old infants. Readers may be wondering how anyone gets infants to respond to a mental rotation task. The researchers used a research paradigm that is commonly used with infants. It is called a habituation task. The underlying idea is that infants will look longer at novel stimuli than they will at familiar stimuli. In the jargon of psychology, infants habituate to familiar stimuli and stop looking at them. Using this paradigm, Moore and Johnson evaluated the hypothesis that infants can mentally rotate visual stimuli through 3-dimensional space and investigated possible sex differences in performance. They reasoned that if infants recognized that a familiar object was the same except for its orientation, then they should look at it for a shorter amount of time than they would for the mirror image of the same object, which would be a novel stimulus for the infants. They tested 20 female and 20 male 5-month-old infants (plus another 5 that were not included in their data analysis because of "fussiness" and "sleepiness"—a common problem when doing research with infants). The male infants looked longer at the novel objects than they did at the familiar ones, with no difference in looking times for the female infants. The effect size for this sex difference in looking time was $d = 0.66$, which is a fairly large effect.

Amazingly, Quinn and Liben (2008) used a similar paradigm with 3- to 4-month-old infants. Their stimuli were somewhat different, but like Moore and Johnson (2008), they found that male infants showed a novelty preference that indicated that they could tell the difference between the mirror image of a familiar stimulus and the same stimulus in a novel orientation (rotated in space to a tilt that had not been seen before). In an earlier version of this book, I wrote that sex differences in visuospatial processing can be found as early as it can be tested, which at the time was age 3. It is now clear that these sex differences occur very early in life and can be found with 3-month-old infants.

In the section on verbal abilities, I reviewed a 10-year follow-up study conducted by De Frias et al. (2006) in which they tested groups of people ranging in age from 35 to 80 years old and then retested them 10 years later. They concluded that the female superiority on many verbal tasks remained the same over the intervening decade. In the same study, they also assessed visuospatial abilities using the block design test which is part of the WAIS. They found that the male advantage on measures of visuospatial abilities did not change over the 10-year period between test administrations.

Meinz and Salthouse (1998) posed a question that is of great interest in an aging society, "Is age kinder to females than to males?" They examined data from 25 separate studies that compared men's and women's cognitive abilities in old age with those of younger adults. For the older group, they found the same overall pattern of cognitive sex differences that has been reported with younger age groups: older women are faster than older men on speeded perceptual tests and are slightly better on verbal fluency tasks (in this study it was nonsignificant). Older men scored considerably better than older women on visuospatial tasks and somewhat better on working memory tasks (which may be associated with the type of memory tasks they used). It is comforting to know that although the older adults declined in most cognitive abilities, especially visuospatial ability, they showed increases in knowledge into old age and no change in verbal fluency. When Jansen and Heil (2010) examined the aging question using mental rotation tasks, they found the typical large effect size favoring males for young adults aged 20 to 30 years (d = 1.07), with smaller effect sizes for the sex differences at ages 40 to 50 and 60 to 70 (d = 0.53 and d = 0.59, respectively). The decline in the effect size for older adults probably reflects two processes. First, performance declined for both women and men with age, and the overall low rate of correct responses could be causing a floor effect, and second, there is some evidence that these abilities decline more rapidly in old age for men. Overall, it does seem that age is kinder to women, at least when some cognitive abilities are concerned.

The conclusion that age is kinder for women was replicated in the massive study sponsored by the BBC that collected data on the internet. Researchers found that men performed better than women on an internet version of the mental rotation task (d = 0.49) and the judgment of line orientation task (d = 0.57), and although performance declined with age from ages 20 to 60, the decline was greater for the men (Maylor et al., 2007). The data from Maylor et al. are shown in Figure 3.22.

Cognitive Styles

There has been considerable interest in recent years in the notion that males and females may have different cognitive styles. The term "cognitive styles" does not have an intuitive meaning. In general, it refers to individual differences in modes of perceiving, remembering, and thinking (Kogan, 1973). It was a popular area of interest during the 1960s and 1970s. It is used by some psychologists in conjunction with the concept of psychological differentiation (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962). An individual who is highly differentiated can separate herself or himself from the environment and can separate items from each other in the environment. According to the theory of psychological differentiation, we all differ in terms of how well we can

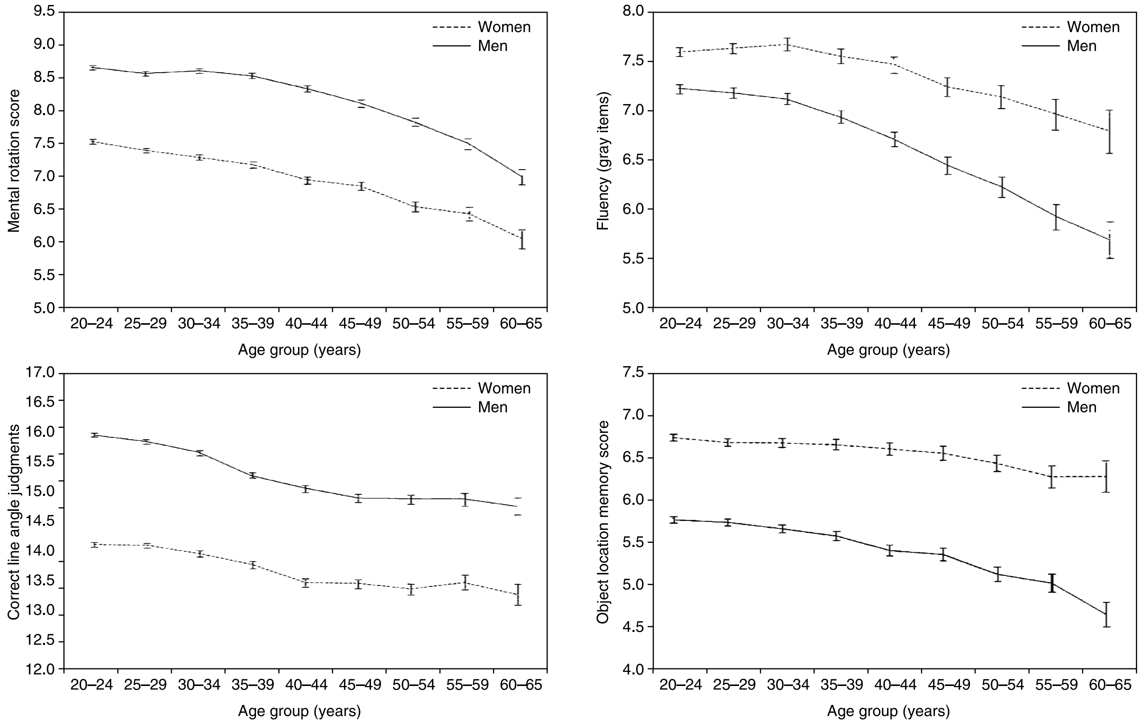


Figure 3.22 Data from BBC internet study showing that at every adult age, women outperform men on a measure of verbal fluency (naming words that belong in categories) and memory for object location and men outperform women on mental rotation and judgment of line orientation tasks. From Maylor et al. (2007). Reprinted with kind permission of Springer Science + Business Media.

separate items in the environment. There are several dimensions or aspects of psychological differentiation. One dimension along which the sexes are said to differ is in field articulation or “field dependence and independence.” These terms were coined by Witkin and have been used to characterize the degree to which subjects are influenced by objects in their visual field.

One way of assessing field dependence and independence is with the Rod and Frame Test. In this test, subjects are seated in a darkened room and are presented with a luminous rectangle (the frame) that has a luminous rod positioned inside of it. The rectangle is rotated to different orientations by the experimenter. The task for the subject is to position the rod so that it is vertical. Figure 3.23 shows a schematic drawing of two rod and frame combinations with which participants could be presented. Some subjects’ judgments of true vertical for the rod are influenced by the tilt of the frame surrounding the rod. They are labeled “field-dependent.” Other subjects’ judgments of true vertical for the rod are not influenced by the tilt of the frame surrounding the rod. They are labeled “field-independent.” In general, sex and age differences are found with the Rod and Frame Test (although differences are not unanimously reported). The usual findings are that children are more field dependent than adults, and females are more field dependent than males.

Measures of field dependence and independence obtained with the Rod and Frame Test are highly correlated with measures obtained with a test known as the Embedded Figures Test. In the Embedded Figures Test, subjects are shown a simple geometric form and then must maintain it in memory and pick it out from a more complex form. Sample items similar to those found in the Embedded Figures Test are shown in Figure 3.24.

Both the Embedded Figures Test and Rod and Frame Test require the subject to segregate a geometric form from its context (the form is either an embedded multi-sided figure or the rod), and in both tests females are more influenced by the context than males. Field dependence was historically hypothesized to reflect personalities that are conforming, submissive to authority, into comfortable ruts, and passive (Elliot, 1961). Women’s field dependence was described

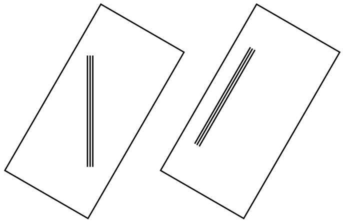


Figure 3.23 A schematic diagram of the Rod and Frame Test. The instructions are, “Align a rod within these frames so that it is vertical” (ignoring the tilt of the screen).

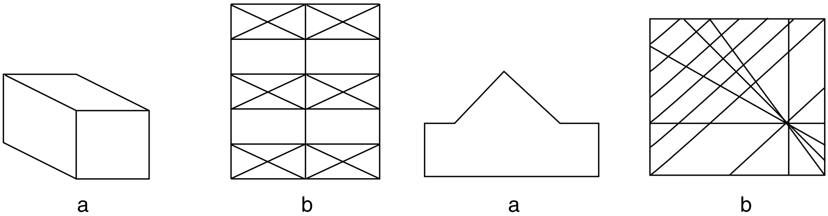


Figure 3.24 Embedded Figures Test. Is Figure (a) part of Figure (b)?

as “accepting the field more passively than men” (Sherman, 1967, p. 290). On the basis of these test results, women’s cognitive style was described as “global,” “conforming,” and “child-like.” According to Witkin et al. (1962), it was similar to the undifferentiated thought processes found in “primitive” cultures. The field independence associated with male performance was described, by contrast, as reflecting a cognitive style that is “analytic” and “self-reliant.” (The value-laden bias in these descriptive terms should be too obvious to require comment.) Witkin et al. believed that because women are unable to maintain a “sense of separate identity” (p. 218), they were less skilled at certain types of problem solving, more likely to conform to group pressure, and more concerned with the facial expressions of others. Thus, different cognitive styles were ascribed to men and women on the basis of their performance on these two tests.

It would appear, however, that spatial tests of field dependence and independence are not indicative of cognitive styles, that is, they are unrelated to passivity or submissiveness, notwithstanding the claims of Witkin (1950; Witkin et al., 1954) and others, but merely reflect sex differences in visuospatial abilities. Several researchers have argued that sex differences in field independence are an artifact of sex differences in visuospatial ability because both the Rod and Frame Test and the Embedded Figures Test have a strong spatial component (Sherman, 1967). This is yet another example of the important distinction between experimental results and the explanations that we “invent” for them. Somehow a test of visuospatial ability came to be used as an indicator of personality traits and the inferiority of women.

A similar and more modern example is the idea that women are biologically predisposed to empathize and men are biologically predisposed to synthesize (Baron-Cohen, Knickmeyer, & Belmonte, 2005). This theory is described in more detail in Chapter 5. As readers might expect given the distinction between field-dependent and field-independent styles, the style associated with being female is purported to be less compatible with careers in science and engineering than that associated with being male. This distinction has been pilloried in recent books on using science to support a sexist agenda, including books by Fine (2010) and Jordan-Young (2010).

QUANTITATIVE ABILITIES

> The underrepresentation of women in mathematics related careers, long an issue of equity and justice, has serious economic implications as the United States faces a shortage of scientists, engineers, and mathematically trained workers.
>
> —Penelope H. Dunham (1998, para. 1)

Plake, Loyd, and Hoover (1981) summarized findings of sex-related differences in quantitative (mathematical) ability this way: "There is little doubt that females score differently from males on mathematical tests" (p. 780). As you can probably guess, "differently" is a euphemism for poorer, but is this widespread belief that males outperform females in quantitative skills supported by data? The short answer is both yes and no.

It seems that quantitative abilities, like spatial and verbal abilities, are a heterogeneous concept. There are several different aspects of quantitative abilities, and there is good evidence that sex differences are manifested in only some of them. Examples of the types of tasks that are used to assess quantitative ability are shown in Figure 3.25.

Janet Hyde and her colleagues (Else-Quest, Hyde, & Linn, 2010; Hyde, 2005; Hyde, Lindberg, Linn, Ellis, & Williams, 2008) have published a series of carefully conducted meta-analyses and, in all of their papers, they find support for the "gender similarity hypothesis," which, as its name states, is the idea that males and females are fundamentally similar in most (but not all) measures. In one study (Hyde et al. 2008), the authors used data on mathematical achievement from several states in the United States. They compared the average math scores for girls and boys from Grades 2 through 11, and found virtually no difference in any of the grades (*d*s ranged from +0.06 to −0.02). But, as the authors noted, the tests were more heavily weighted with items that required lower level thinking (i.e., recall and skill/concept) rather than higher level thinking (i.e., strategic thinking and extended thinking). Thus, based on these data, there is little difference in the average performance of girls and boys in mathematics in Grades 2 through 11, at least when the problems are not very difficult. One major criticism of this study is that sex differences in mathematics become progressively larger as the sample becomes more selective and the type of math skill becomes more advanced. To test this possibility, the authors examined the ratio of males to females among the highest-scoring students. For White students scoring at the 95th percentile, the ratio of boys to girls was 1.45; for White students scoring at the 99th percentile (top 1% of all students), the ratio of boys to girls was

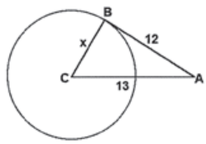
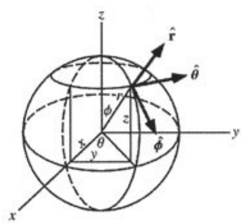
Algebra	Solve the set of equations for x and y. $4x + 7y = 10$ $5x + 3y = -6$
Geometry	Find the value of x given that the segment AB is tangent to circle C at B. 
Geometry – advanced and highly spatial (introductory course in multivariable calculus in sophomore year for majors in many different STEM fields)	Express the Cartesian coordinates x, y, and z using the spherical coordinates φ, θ, and r shown below.  Figure is from http://mathworld.wolfram.coin/SphericalCoordinates.html
Computation	276 $\times\ 16$ $?$
Word problem	If Fred can paint a room in two hours and Sally can paint the same room in three hours, how long will it take them to paint the room if they work together?
Calculus (introductory college calculus course or advanced high school calculus course)	$\int x^2 \cos(x)\, dx = ?$
Mathematical proof (introductory mathematical analysis course for mathematics majors in their sophomore or junior college year)	Suppose a and c are real numbers, $c > 0$, and f is defined on [–1, 1] by $f(x) = \begin{cases} \|x\|^a \sin(\|x\|)^{-c}) & \text{(if } x \neq 0\text{)}, \\ 0 & \text{(if } x = 0\text{)}. \end{cases}$ Prove that f is continuous if and only if $a > 0$. This problem is from Rudin (1976, p.115) Rudin, W. (1976). *Principles of Mathematical Analysis, Third Edition*. McGraw-Hill: New York.
Differential equations (in quantum mechanics) (introductory quantum mechanics course for physics majors in their sophomore or junior college year)	Consider a particle of mass m constrained to the 1-D potential of $V(x) = 0$ for $0 < x < L$, but $V(x) = 8$ otherwise. Solve Schrödinger's equation $i\hbar \frac{\partial(x,t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right](x,t)$ for the generalized eigenstate (x, t) of the energy operator $\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$

Figure 3.25 Sample questions used to assess quantitative ability. I thank David I. Miller at University of California, Berkeley, for this figure.

2.06. The authors made the same comparison for Asian/Pacific Islander students and failed to find the same predominance of boys to girls among the highest scorers (ratio of boys to girls was 1.09 at the 95th percentile and 0.91 at the 99th percentile). Critics have noted that there were relatively few Asian/Pacific Islander students at the highest levels and that since the tests were at a relatively low level of difficulty, the data from the high achieving tail of the distribution are misleading because sex differences emerge on the more difficult test items.

Many researchers have argued that sex differences in mathematics reflect opportunity inequalities and economic opportunities (e.g., Else-Quest, Hyde, & Linn, 2010). To test this hypothesis, researchers analyzed large international data sets along with measures of gender equity and economic development. As you may recall, similar analyses were conducted with the international data regarding female and male differences on mental rotation and judgments of line orientation. The authors of the international math study found that there was only a small advantage for males, which was in contrast to the finding that males were much more confident in their math abilities than girls were.

Guiso, Monte, Sapienza, and Zingales (2008) analyzed cross-national data on math and reading scores. They found that there is considerable variability among countries, but on average, girls score lower than boys (the mean score for girls is 2% lower than the mean score for boys) in math, but when gender equality of the country was considered, these differences virtually disappear. It is interesting to note that when they performed these analyses on reading scores, the advantage for girls (which was 6.6% higher than the mean score for boys) increased. The authors conclude that "in more gender equal societies, girls perform as well as boys in mathematics and much better than them in reading" (p. 1165). I return to these data in Chapter 7, where I discuss cultural influences on cognitive sex differences.

Recall that the section on sex differences in the tails of distributions showed that there are approximately 3 to 4 males for every female who achieves the highest scores among gifted students for mathematics (Wai, Cacchio et al., 2010). It seems reasonable to conclude that the average differences between females and males in math are small, but as the samples become more select, that is as ability levels increase, males outscore females. This conclusion is supported with data from both of the most commonly used standardized exams for college entrance—the SATs and ACTs. Sex differences on math tests are shown in Figures 3.26 and 3.27. One reason why the average difference on the SAT-M is as large as it is may relate to the gender make-up of test-takers. Recall from an earlier section in this chapter that many more women take the SATs than men, which should result in a lower mean score because more women of lower ability are taking the SATs than men of lower ability. The differences in number of women and men who take the SATs mean that any conclusions about sex differences based on average SAT scores should be made with extreme caution.

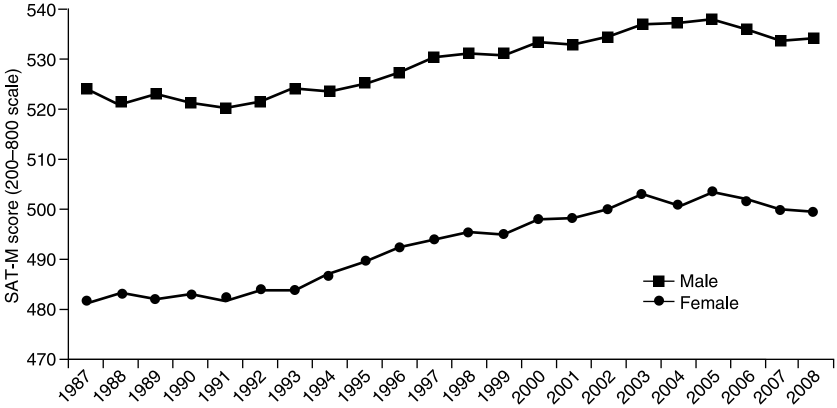


Figure 3.26 SAT mean scores in mathematics for females and males from 1987 to 2008. From Hill, Corbett, and St. Rose (2010). Copyright © 2006 The College Board. Reprinted with permission www.collegeboard.com. Data for 2007 and 2008 from College Board, www.collegeboard.com

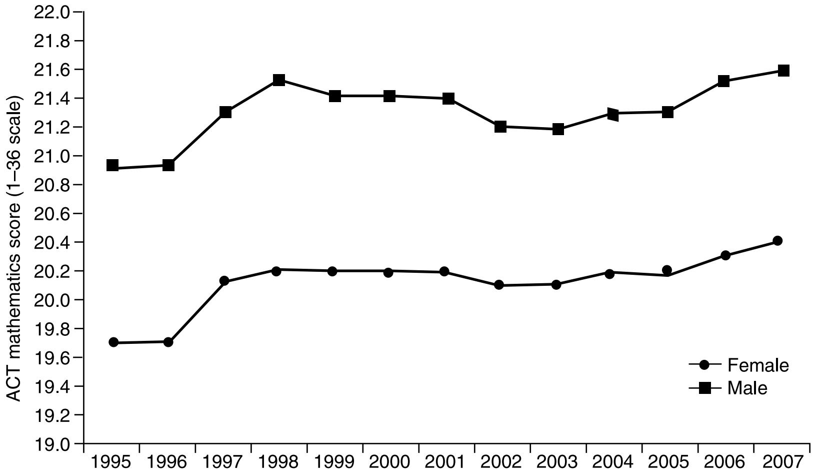


Figure 3.27 ACT mean scores in mathematics for females and males from 1995 to 2007. Unpublished data provided by the ACT Statistical Research Department. Reprinted with permission of ACT, Inc.

Age Trends in Quantitative Abilities

Sex differences in quantitative abilities vary throughout the life span. For example, among young children (ages 4 to 10 years) girls and boys perform similarly on tests of primary mathematical reasoning abilities (Spelke, 2005).

During or shortly after elementary school, however, when quantitative tests become more complex and more visuospatial in nature, sex differences emerge and continue to grow thereafter (Beilstein & Wilson, 2000). By the end of their secondary schooling (12th grade), males demonstrate significantly higher achievement than females in the areas of number properties and operations as well as measurement and geometry (Rampey, Dion, & Donahue, 2009). This trend has remained steady since the early 1970s. Interestingly, females get higher grades than males in school in all subjects, including math, at all grade levels (Kimball, 1989; Snyder, Dillow, & Hoffman, 2009; Willingham & Cole, 1997) and do slightly better on international tests of algebra (U.S. Department of Education, 2005). But, when males and females are compared on tests that reflect content learned in school, such as state-wide assessment tests, the differences disappear. Although it should be noted that these tests tend to evaluate lower level skills and leave open the possibility of sex differences if higher order skills were assessed (Hyde, Lindberg, Linn, Ellis, & Williams, 2008). Math differences favoring males are larger and more commonly found on tests that are not directly tied to the curriculum, such as the SATs, which may reflect novel problem solving skills. On average, males taking the SATs have consistently scored about a third of a standard deviation higher than girls over the past 25 years (data from College Board, 2007; for a review see Halpern, Benbow, Geary, Gur, Hyde, & Gernsbacher, 2007a, b).

The overrepresentation of boys among the most gifted in mathematics can be detected at very young ages. In a study of mathematically precocious young children, Robinson, Abbott, Berninger, and Busse (1996) found that more young boys were referred for giftedness in mathematical reasoning than young girls, despite special attempts to include girls. They administered a test battery to 143 preschool girls, 167 preschool boys, 201 girls in kindergarten, and 248 boys in kindergarten who were identified as possibly having the potential for mathematical giftedness. In this select group, they found sex differences favoring the boys in tests of number knowledge, number series, numeration, problem solving, calculation, word problems, counting span, an arithmetic screening test, and a test of visuospatial span. These children were also administered three different verbal tests. Although the sex differences on the three verbal tests were not statistically significant, for all six comparisons (two age groups on three verbal tests), the girls scored higher than the boys. These authors reported sex differences in every analysis—more boys were nominated for the mathematically gifted program, a greater proportion of the boys qualified for admission to the program, and the boys scored higher than the girls on 8 of the 11 subtests. It is important to note here that sex differences are more commonly found in highly select groups, so these conclusions are not generalizable to children whose performance is closer to average.

THINKING ABOUT THE MAGNITUDE OF DIFFERENCES

> Test performance may have real, quantifiable educational and social implications. Small mean differences combined with modest differences in variance can have a great effect on the number of individuals who excel.
>
> —Michael Beller and Naomi Gafni (1996, p. 375)

Although the preponderance of the experimental evidence points to some sex differences in verbal, visuospatial, and quantitative ability, practical questions about the size or magnitude of these differences have not been easy to resolve. Are the differences trivial and of no practical significance or do they represent meaningful ability differences between the sexes? Even if we were to conclude that there are large between-sex differences with respect to a cognitive ability, it is very important to remember that most research analyzes group average results that cannot be applied to any individual.

A serious and common misunderstanding about sex differences is to conclude that "women are like this—men are like that" and then decide to treat each sex differently based on average differences. The focus on differences does tend to obscure the many more ways males and females are similar, and the way we think about group differences could prevent us from considering people as individuals. There is considerable overlap between the female and male distributions in every ability.

All of the cognitive sex differences have been replicated numerous times and are statistically significant, which means that they are unlikely to have occurred by chance, but are they of any practical significance? Can they be used to explain why we have so few female mathematicians or engineers? Can they help us predict a male's or female's ability to perform a task? Can they be used to justify discrimination? Are they merely curiosities whose only value is to keep psychologists (and publishers) busy? Answers to these questions are hotly debated and have important implications for modern society.

On an intuitive level, effect size is a quantification of the size of the average between-sex difference on a particular test or set of tests. Unfortunately, the numbers we use to express effect size are not intuitive. Differences like the finding that men tend to outscore women by an average of 40 points on the quantitative portion of the SATs have an immediate meaning to anyone who is familiar with the scoring system for the SATs. Unfortunately, sex differences in abilities are measured with many different tests and a common measure of the average difference is needed to make comparisons across many studies. The effect size statistic is used to convey the size of the differences when many different tests are used. (Readers for whom this is a new concept are referred back to Chapter 2 where statistical concepts are discussed in more depth. It is also possible to follow the gist of the following discussion without understanding the fine points of some of the statistical concepts that are discussed.)

There are few guidelines for determining if the size of a sex difference with respect to a cognitive ability is large enough to be important. Cohen (1977) provided an arbitrary statistical definition of small, medium, and large effect sizes using standard deviation units (0.20 SD is small, 0.50 SD is medium, and 0.80 SD is large). There is, however, no good reason to accept his effect size markers except for the fact that they provide a common ruler for comparing differences. It is important to realize that effect size should not be confused with importance. A small effect could still be important, depending on how importance is defined and who defines it. Percentage of explained variance statistics (e.g., omega squared, R^2, eta squared) are useful in this regard, but they still leave us with the question of how much explained variance is large enough to be important. If sex explained 5% of the variance in the data, is this a large or small number? In another context, like medicine, 5% of explained variance attributable to a treatment could mean many lives would be saved. Thus, the question of whether 1% or 5% or 50% of explained variance is important depends on both the context and value judgments. Value judgments never lend themselves to statistical analysis, and thus, precise answers to the question of how large does a difference have to be to be important will remain debatable.

In one of the most lucid discussions on how to interpret effect sizes, Rosenthal and Rubin (1982) attempted to shed light on the question of how large an effect size must be in order to be of practical importance. As described in Chapter 2, they used a statistical test known as the binomial effect size display (BESD). They calculated that when sex explains only 4% of the variance in test scores, this translates into distributions in which 60% of the higher scoring sex is above the median and only 40% of the lower scoring sex is above the median. They argue that outcome rates of 60% versus 40% are important because they can be used to predict performance on ability tests in these areas. Here is how they explained their reasoning:

> We do not agree that gender is a poor predictor of one's performance on ability tests. If obtaining a particular job required scoring above the median on a test that correlated .20 with being female, then for every 100 females and 100 males that applied, 60 of the women, but only 40 of the men would be job eligible. (p. 711)

They also looked at the consistencies among effect sizes across 12 studies of verbal ability, seven studies of visuospatial ability, seven studies of quantitative ability, and 14 studies of field articulation (field independence and dependence). They concluded that effect sizes differed from study to study, supporting the idea that the magnitude of the sex difference in any area depends on the type of test used.

UNDERLYING COGNITIVE PROCESSES

Examining sex differences for cognitive abilities is only one way of conceptualizing how females and males may differ in their intellectual processes. The division of abilities into verbal, visuospatial, and quantitative has been useful, and as discussed in the next two chapters, each of these abilities has distinct biological correlates. But, there are other ways of investigating the thinking process. One such way is to consider what the subject does when he or she is engaged in a particular task. This alternate approach can be thought of as examining the underlying cognitive processes.

Look carefully at Table 3.1. I have listed the types of tasks on which females tend to excel and the types of tasks on which males tend to excel. One approach is to consider these two types of tasks as representing different underlying cognitive processes. The tasks at which females excel include language production, reading, writing, generating synonyms, word fluency, memory for words, objects and locations, and algebra. (Algebra may have a more language-like structure than other types of mathematics.) These are high-level tasks that require rapid access to and retrieval of information that is stored in memory and the use of language in manipulating and creating information. On the other hand, consider those tasks at which males tend to excel—mathematical problem solving, verbal analogies, mental rotation, spatial perception, and

Table 3.1 Possible sex differences in underlying cognitive processes

Tasks at which females excel:

- Generating synonyms (associational fluency)
- Language production and word fluency
- Computation
- Anagrams
- Memory for words, objects, personal experiences, and locations
- Reading comprehension and writing

Underlying cognitive processes: rapid access to and retrieval of information in memory

Tasks at which males excel:

- Verbal analogies
- Mathematical problem solving
- Mental rotation and spatial perception
- Spatiotemporal tasks (dynamic visual displays)
- Generating and using information in visual images
- Mechanical reasoning and some science-related topics

Underlying cognitive processes: maintaining and manipulating a mental representation in visual–spatial working memory using information in dynamic visual displays (spatiotemporal tasks) and visual images. These sorts of tasks require the ability to maintain and manipulate mental representations. Thus, it may prove meaningful to differentiate cognitive tasks on the basis of the type of cognitive process that each requires. When we adopt this framework, we can account for sex differences that do not divide neatly under the tripartite cognitive abilities rubric (verbal, mathematical, and visuospatial) and we can incorporate female superiority on some visuospatial and mathematical tasks and male superiority on some verbal tasks.

ARE SEX DIFFERENCES DECREASING?

> Contrary to the findings of small scale studies, these average differences do not appear to be decreasing but are relatively stable across the 32-year period investigated.
>
> —Larry V. Hedges and Amy Nowell (1995, p. 45)

The next four chapters will describe theories that have been proposed to explain why sex differences are sometimes found. If these differences were created by sex-differentiated psychosocial variables like sex roles and different rewards for males and females, then we would expect to see some decline in the magnitude of the differences as the impact of sex roles diminishes for a variety of reasons including the fact that women have increasingly greater access to economic, educational, and political opportunities over time. Thus, the question of whether sex differences in cognitive abilities are decreasing is important. In order to conclude that sex differences are decreasing, we need to have comparable samples of subjects that have taken the same cognitive abilities tests in different time periods. There are few samples that meet these stringent requirements.

Several experimenters have examined effect sizes as a function of the date that the study was published. The underlying rationale for investigating results as a function of their date of publication is that more recent studies should, in general, show smaller sex differences than studies published many years ago, if sex differences really have been decreasing. The problem with this approach is that a great many other variables have also changed during the intervening years. In response to concerns that publication practices tend to be biased toward studies that report significant differences, many more journals and paper presentations now report nonsignificant results, thus changing the nature of the studies that can be included in meta-analyses. (In other words, a study that fails to find significant differences is more likely to be published than in the past.) The more recent tendency to publish nonsignificant results would cause effect sizes to decrease as a function of publication date.

The nature of samples has also changed with time. Women surpassed men in college enrollments in the United States in 1982. Currently women comprise approximately 56% to 57% of college enrollments, with higher proportions in community colleges than in 4-year colleges (Peter & Horn, 2005). Because a larger percentage of all females are now attending college than the percentage of males, a more select group of college men is probably being sampled than college women. The nature of many of the tests has also changed. The Educational Testing Service, which authors the SATs, has come under severe criticism for the disparities in female and male scores. Accordingly, they have responded in the last few years by scrutinizing every test question for sex-related bias in content or use of pronouns. The Educational Testing Service now trains all of its test committees on ways to avoid bias in the questions that are used in their examinations. Many of the other tests that show the greatest sex differences have been developed within the last few years (e.g., paper and pencil mental rotation tests, word fluency, and consonant–vowel matching tests) and therefore cannot be compared with comparable older studies to see if the effect sizes are diminishing. The SATs were substantially changed in 2005, which makes comparisons from earlier years more difficult.

The data on right-tail ratios clearly show a decrease in the proportion of males to females among highest scorers in math and verbal skills (Wai, Cacchio et al., 2010). As noted earlier, the ratios were reduced from 13:1 to between 3:1 and 4:1 for mathematics, with the latter value remaining constant over the last 20 years. So, when considering the highest scores, it does seem that there has been a reduction in the size of sex differences, at least on some standardized tests of mathematics. Differences are clearly small for average performers in mathematics, which may also be a reduction in the size of sex difference, but as the international study by Guiso et al. (2008) showed, we can expect the female advantage in reading to grow larger as more societies achieve gender equality and continued economic development. Like every conclusion in this text, the answers are not simple.

Very large sex differences are found on some visuospatial tasks, most notably mental rotation and judgments of line orientation. The effect sizes have remained between $d = 0.8$ and $d > 1.0$ for several decades, so it does not appear that the female–male difference in mental rotation is changing. Finally, for judgments of line orientation, there are not enough data from earlier decades to determine trends over time.

SIMILARITIES

Although the focus of this chapter has been the identification of cognitive abilities that show sex differences, the flip-side of this issue is at least as important—those areas of cognition in which similarities are found. I have focused on differences because the logic of hypothesis testing only allows conclusions about differences. Despite this limitation, it is important to note that the number of areas in which sex differences are even moderate in size is small. Males and females are overwhelmingly alike in their cognitive abilities. It is important not to lose sight of this fact as we consider theories that have been posited to explain the differences and similarities in cognitive sex differences. Also, please keep in mind a point that I repeat in several places in this book: even in the relatively few areas in which differences are found, these conclusions are based on data gathered from a large number of subjects. They cannot be applied to any single individual because the within-sex variability is so large.

CHAPTER SUMMARY

Although sex differences have not been found in general intelligence, there are some types of cognitive abilities that vary, on the average, as a function of sex. There are some sex-related differences in the earliest stages of information processing—perception and attention—but the effect of these early stage differences on later cognitive processes is unknown and we cannot conclude that they are responsible for differences in cognitive abilities. Males comprise a disproportionate share of the extremely low ability end of the verbal abilities distribution, with males overwhelmingly categorized as stutterers, dyslexics, and low IQ. By contrast, females excel at general and mixed verbal ability tests, speech production, writing, memory for words, objects, and locations, (some) perceptual motor skills, and associational fluency. These differences appear as soon as speech and language usage begin.

There are few differences in quantitative abilities for most of the population—that is, the middle range of the ability distribution, but there are 3 to 4 times more males scoring at the highest levels on standardized tests of mathematics that are designed for use in college and beyond. Similarly, there are disproportionately more females at the high ability end on writing tests and (to a lesser extent) on tests of verbal reasoning. There are at least five types of visuospatial ability that have been identified: spatial perception, mental rotation, spatial visualization, spatiotemporal ability, and generation and maintenance of visual images. Sex differences favoring males are found on all of them except spatial visualization, which typically does not show sex differences, but when sex differences are found, they favor males. The effect sizes for mental rotation and judgments of line orientation are among the largest found in the literature and can be found developmentally—in infants as young as 3 months old for mental rotation. The effect size for visuospatial abilities has remained unchanged for many decades. An analysis of the underlying cognitive processes was proposed, with males performing especially well on tasks that involve maintaining and manipulating mental representations and females performing especially well on tasks that require rapid access to and retrieval of information from memory, especially when the information is verbal. It is important to keep in mind that the list of cognitive differences is relatively small and that cognitive similarities between the sexes are greater than the differences.

Chapter 4

Biological Hypotheses Part I: Genes and Hormones

CONTENTS

> After sexism is stripped away, there will still be something different—something grounded in biology.
>
> —Michael Konner (1988, p. 35)

Perhaps this chapter and the next should come with a warning similar to the ones found on cigarette advertisements:

WARNING: Some of the research and theories described in this chapter may be disturbing to your basic belief systems.