Educational Psychology Review, Vol. 12, No. 2, 2000

The Smarter Sex: A Critical Review of Sex Differences in Intelligence

Diane F. Halpern[1,3] and Mary L. LaMay[2]

Although there are no sex differences in general intelligence, reliable differences are found on some tests of cognitive abilities. Many of the tasks that assess the ability to manipulate visual images in working memory show an advantage for males, whereas many of the tasks that require retrieval from long-term memory and the acquisition and use of verbal information show a female advantage. Large effects favoring males are also found on advanced tests of mathematical achievement, especially with highly select samples. Males are also overrepresented in some types of mental retardation. Effects sizes are variable and often large. These differences are generally found cross-culturally and across the life span. The nature–nurture dichotomy is rejected as an interpretive framework. In light of recent findings that environmental variables alter the biological underpinnings of intelligence and individuals actively participate in creating their environments, we prefer a psychobiosocial model for understanding sex differences in intelligence.

KEY WORDS: human sex differences; intelligence; spatial abilities; verbal abilities; human cognition.

INTRODUCTION

It is a question that has been asked in many different ways, in multiple contexts, and with a variety of political agendas in mind: "Which is the

[1]California State University, San Bernardino, California.
[2]Loma Linda University, Loma Linda, California.
[3]Address all correspondence to Diane F. Halpern, Department of Psychology, California State University, San Bernardino, 5500 University Parkway, San Bernardino, California 92407. Email: dhalpern@csusb.edu

smarter sex[4]—female or male?'' This question has a long and turbulent history entangled with beliefs about the appropriate roles of males and females and the political and economic ramifications of the way we answer this question. The question of whether males or females are smarter presupposes that it can be answered, that valid comparisons can be made, and that ''female,'' ''male,'' or ''neither'' are possible answers. The question of the smarter sex also assumes that the answer, whatever it is, will not be restricted to this time in history, to contemporary Western societies, to any specific developmental period in the life span, or to any of the other ways in which humans differ. It will not surprise psychologists or educators to learn that the answer to this question depends on the way intelligence is conceptualized and measured.

WHAT IS INTELLIGENCE?

The answer to the question of which is the smarter sex depends on how ''smart'' is defined. This question is based on the underlying idea that there is a theoretical construct that we have labeled ''intelligence'' and that the sex that has more of it, on average, is the ''smarter'' one. Although the ''smarter sex question'' is seemingly innocent and straightforward, it is, in fact, a sociopolitical minefield, with serious implications for a wide range of public policies (e.g., affirmative action, compensatory education, pay equity; Halpern, 1996).

The easiest way to answer the question of which is the smarter sex is to give a psychometrically sound test of intelligence to a large, representative sample of males and females and then determine which group gets the higher average score. This solution rests on the belief that the test of intelligence is really measuring what psychologists mean by intelligence, and that it is doing so in a way that will yield a fair assessment for males and females—two assumptions that may not be justified. There is a vast literature in which the ''experts'' and almost everyone else debates what the term *intelligence* should mean, whether it has a single underlying dimension (known as ''g'') (e.g., Jensen, 1998) or multiple dimensions (e.g., Gardner, 1983), or whether it is best conceptualized as varying along dimensions that are crystallized (influenced by education and culture) or fluid (a purer indicator of ability; Horn, 1985). There are several alternative definitions

[4]Some authors prefer to differentiate between the terms *sex* and *gender,* using *sex* to refer to biologically mediated differences and *gender* to refer to socially mediated differences. In accord with the psychobiosocial model presented, we believe that biological and social influences often are not separable; therefore, we use the more generic term *sex* to refer to differences between women and men, regardless of their origin.

of intelligence, but none of the alternatives are reflected in the most commonly used standardized intelligence tests. For example, the idea that intelligence includes the acquisition and use of practical knowledge (Sternberg, Wagner, Williams, and Horvath, 1995), bodily kinesthetics or musical abilities (Gardner, 1983), and emotional capabilities (Goleman, 1995) has gained some popularity among educators and the general public, but these conceptualizations are not assessed in the most frequently used tests of intelligence. Thus, a central issue is whether intelligence is being operationally defined in a way that is fair and reasonable to both sexes. Yet, who decides what constitutes a fair and reasonable definition of intelligence?

The tests of intelligence that are used in educational settings are based on the idea that intelligence is related to success in a traditional, academic curriculum. In Western societies, there is a handful of commonly used intelligence tests that are recognized as valid predictors/correlates of academic success: the Wechsler Adult Intelligence Scale (WAIS), the Wechsler Intelligence Scale for Children (WISC, now in its third revision), the Standford–Binét Intelligence Test, Raven's Progressive Matrices (a nonverbal test), and those tests designed to predict performance in college and graduate school, such as the Scholastic Assessment Tests (SAT) and Graduate Record Examination (GRE). The SAT, GRE, and other similar tests do not claim to be tests of intelligence because they measure "developed abilities," not innate traits of the test taker. They are included in this list because they are used as traditional intelligence tests and because they figure prominently in questions about female and male similarities and differences. Intelligence tests are validated with school achievement as assessed with graduation rates or grades in school, or with "achievement tests" in subject areas that are taught in school. For these tests, intelligence is conceptualized as knowledge of information that most people (in Western societies) would commonly encounter, such as word meanings, deducing information, and maintaining information in working memory. More advanced measures of intelligence reflect information taught in formal academic disciplines and the application of reading and mathematical skills.

AVERAGE IQ SCORES

The average scores of males and females, in intelligence quotients (IQ) or other units, should provide an answer to the question of which is the smarter sex. Although between-sex comparisons of IQ scores may seem like a reasonable approach to answering the question of which is the smarter sex, it actually is not because tests that purport to measure intelligence have been constructed so that there are no overall sex differences (Brody, 1992; Makintosh, 1996). Female and male scores are equalized by eliminating questions that show a large advantage for either females or males or by balancing questions that a greater proportion of females' answers correctly with questions that an equally greater proportion of males' answers correctly. Full scale IQ scores represent an average of heterogeneous subtests, and although on average there are no differences between males and females on the IQ scores obtained, there are group differences on the subtests, suggesting that females and males differ on at least some of the abilities assessed with intelligence tests. The largest differences are found in the "tails," or extreme ends, of IQ distributions, with males overrepresented in some types of mental retardation, learning disabilities, and language disorders (DeFries and Gillis, 1993; Henning-Stout and Close-Conoley, 1992).

Despite the fact that intelligence tests are deliberately constructed in ways that ensure the same average scores for males and females, some psychologists have used intelligence test scores to argue that males are more intelligent than females. The most vocal advocate of the view that males are the smarter sex is Lynn (1994, 1998), who has administered intelligence tests to samples in Ireland and Scotland and reported that males obtained scores that are 2-5 points higher, on average, than female IQ scores. He used these results to argue that males are smarter than females, an argument that he bolsters with brain size data showing the male brains are, on average, slightly larger than female brains, even after adjusting for body size. There are many flaws with Lynn's conclusions. First, tests that were deliberately written to yield equal average IQ scores for both sexes cannot then be used to support the position that there is a smarter sex. Using data from tests that are designed to yield no sex differences to argue for a difference is psychometric nonsense. Second, there were no sex differences in scores obtained with the standardization sample, and other researchers (e.g., Weiss and Prifitera, 1995) do not find even the small sex differences that Lynn reports. For example, no sex differences were found on the WAIS-R with a large representative sample from the United Kingdom, a group that should have been similar to that used by Lynn (Crawford and Allan, 1995). Small differences are inconsequential and could be caused by sampling variability, experimenter bias, or other factors. Third, even if there were average between-sex differences in IQ scores, this result could be due to bias in the test or to other factors that are unrelated to intelligence. Fourth, sex differences in brain size or structure do not imply a smarter sex. Brain structure changes in response to personal experiences (Greenough, Black, and Wallace, 1987). Thus, sex differences in brain size could result from sex-related differences in experiences. Finally, the brain size data do not support the conclusion that males are more intelligent than females. Female and male brains do, however, differ in many ways. For example, females have, on average, a higher rate of oxygenated blood flow, which could just as readily be used to argue that females have superior brains (Gur and Gur, 1990). A measure as gross as overall brain size would not be expected to reflect relative intelligence, which depends differentially on the number and complexity of neural connections, chemical components of neural transmitters, and the density of neurons. (Sex-related brain differences are discussed more fully in Halpern, 2000.)

Jensen (1998) also entered the fray. Jensen is no stranger to controversies regarding group differences in intelligence. In a paper published in 1969, Jensen argued for race differences in intelligence. In an examination of the question of the smarter sex, Jensen analyzed data from tests that were not written in ways that would ensure equal overall scores for males and females. He used five different test batteries for which he had large, representative samples that encompassed the full range of ability in the general population, a construct that is known as "g" in the literature. 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). According to Jensen, there is no evidence that, overall, either sex is the more intelligent. His conclusion that there are no overall sex differences in intelligence has been supported by Stumpf and Stanley (1996) in their analysis of scores on the Achievement Tests and Advanced Placement Tests taken by college-bound high school seniors. As noted earlier, these tests are not classified as "intelligence tests," but they are used to make decisions about academic ability, and presumably reflect intelligence, at least in part. Stumpf and Stanley found sex differences on individual tests, but overall the sex differences on these tests canceled each other.

The intense and hostile nature of this debate should serve as a reminder that questions about intelligence are rooted in sociopolitical ideologies, and that the way researchers and the rest of us interpret data is highly dependent on personal beliefs. It is easy to understand why many psychologists and others have argued against the study of any sex differences (Hare-Mustin and Marecek, 1994; Hollway, 1994). As noted elsewhere (Halpern, 1997, 2000; Hunt, 1999), we maintain that despite the dangers inherent in answering questions about group differences, censorship—even self-censorship—does not promote equality and can be far more dangerous and counterproductive than directly addressing the question. Stereotypes and prejudice are not caused by an open process of scientific inquiry; in fact, they seem to flourish in the absence of data. The scientific method is the only way psychologists can determine which of the commonly held beliefs about differences and similarities between the sexes has their basis in fact. There are numerous ways in which the scientific method is influenced by the philosophical tenor of the times and the known and unknown biases of the experimenter. Despite the fact that science is not immune to bias, it is the fairest method psychologists have for finding answers to complex questions, especially when the procedures and results are open to all segments of society for commentary and scrutiny.

ARE BIAS-FREE MEASURES OF INTELLIGENCE POSSIBLE?

Although it should be obvious that scores on intelligence tests depend on the questions asked, the content of the questions is frequently overlooked in reviews of intelligence tests. Consider, for example, the "General Information" subtest of the WAIS. Several investigators have noted that females score significantly lower than males on this subtest (Feingold, 1996). In thinking about what this result tells us about female and male intelligence, consider the question, "What 'general information' is indicative of intelligence?" Should an intelligent person in contemporary Western society be expected to know how long it is safe to keep uncooked meat in a refrigerator or how to determine if an infant has a fever? These are general knowledge questions that one might expect more females to answer correctly than males. By comparison, the ability to name at least one country that shares a border with Turkey or knowledge of how to add additional memory to a computer are general knowledge questions that one might expect more males to answer correctly. In general, females and males have somewhat different values and interests and engage in different activities, so one would expect different levels of knowledge about different subject areas. It is difficult, if not impossible, to find sex-neutral occupations or activities. Even college students, whose lives appear to be a similar mix of attending class, doing homework, working, and engaging in leisure activities, report large sex-related differences in how they spend their time. For example, Astin, Sax, Korn, and Mahoney (1995) found that female students spend much less time exercising, partying, and watching television, and much more time on housework, child care, reading for pleasure, and doing volunteer work than do male students. In this study, only 7% of the female students, compared to 37% of the male students, spent more than 1 hour per week playing video games. Given the differences in the daily lives of males and females, it is not surprising that there are questions that will be answered correctly more often by one sex or the other. This is an important caveat to keep in mind when interpreting the results of intelligence tests.

WHAT ARE THE SUBTEST DIFFERENCES BETWEEN MALES AND FEMALES?

Given that there are no differences in the average scores obtained by females and males on intelligence tests, it is more fruitful to examine scores on subtests and on other tests that are more frequently used by cognitive psychologists in investigating individual and group differences in cognitive processes. The most frequently citied difference between the sexes is in the ability to transform a visual–spatial image in working memory. Visual–spatial ability is not a unitary construct (Linn and Petersen, 1986). It seems that there are at least five qualitatively different types of visual–spatial abilities:

1. *Spatial perception* requires participants to locate the horizontal or the vertical in a stationary display while ignoring distracting information. Examples are (1) the rod and frame task, which requires participants to position a rod within a tilted frame so that the rod is either vertical or horizontal, and (2) the Piaget Water Level Task, which requires subjects to draw in the water level in a picture of a tilted glass that is half filled with water.
2. *Mental rotation* involves the ability to imagine how objects will appear when they are rotated in two- or three-dimensional space. There are timed and untimed versions of these tests. Several researchers believe that mental rotation is a measure of general spatial reasoning ability (Casey, Nuttall, Pezaris, and Benbow, 1995).
3. *Spatial visualization* refers to complex analytic multistep processing of spatial information. Tests that tap spatial visualization are the embedded figures test, in which a target figure is "hidden" in the contours of a larger figure; and paper folding, in which participants imagine the spatial result of folding a piece of paper in several directions.
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 and McPhee, 1987) and making "time of arrival" judgments about a moving object (Schiff and Oldak, 1990). 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, and Alderton, 1988).
5. *Generation and maintenance of a spatial image* requires participants

> to generate an image (either from long- or short-term memory), such as the shape of a particular letter of the alphabet, and then use the information in the image to perform a specified cognitive task.

On average, males have a large advantage on these tasks, with the possible exception of spatial visualization, where the differences in favor of males are smaller and less robust than with the other measures. In their meta-analytic review of 286 studies, Voyer, Voyer, and Bryden (1995) concluded that the mean effect size for all of the visual–spatial tasks, favoring males, is approximately one-third of a standard deviation ($d = .37$). The authors calculated a "fail-safe" value for these findings, which is the number of nonsignificant and unpublished findings that would need to exist (unpublished and therefore unavailable) to nullify their conclusion. The fail-safe value for sex differences in visual–spatial abilities is 178,205. In other words, there must be 178,205 unpublished studies that show no significant sex differences to offset the available studies that report sex differences in visual–spatial abilities.

Voyer, Voyer, and Bryden (1995) concluded that there is a nonsignificant difference between males and females on spatial visualization tasks, but that the effect size for spatial perception tasks is $d = .44$, and the effect size for mental rotation tasks is $d = .56$. By comparison, Masters and Sanders's (1993) review of mental rotation studies showed the effect size to be $d = .90$, and Resnick (1993) estimated the effect size to be in the range of $d = .74$ to .80. Spatiotemporal tasks and generation and maintenance tasks were not included in these meta-analyses. Estimates of the size of the sex differences on these task are similar in magnitude to those reported for mental rotation (Loring-Meier and Halpern, 1999).

It is interesting to note that many of the sex differences found in the laboratory are mirrored in ecologically valid settings. For example, when males and females give directions, males are more likely to use Euclidean strategies and north–south–east–west directions (Dabbs, Chan, Strong, and Milun, 1998), and males are more accurate with these relational strategies (Brown, Lahr, and Mosley, 1998). Women, however, are more likely to use landmarks and left–right directions (Dabbs, Chan, Strong, and Milun, 1998).

Many of the differences in visual–spatial abilities appear early in life, and are thus, less likely to be solely caused by differential life experiences. The male advantage in transforming information in a visual–spatial working memory is seen as early as it can be tested, perhaps by age 3 (Robinson, Abbott, Berninger, and Busse, 1996). Meinz and Salthouse (1998) posed a question that is of great interest in an aging society, "Is age kinder to females than to males?" The authors were intrigued by data from brain researchers showing that older women showed less reduction in brain volume in old age and did not begin to lose brain volume until an older age than men (Cowell, Turetsky, Gur, Grossman, Shtasel, and Gur, 1994; West, 1996). If there is a fairly direct relationship between brain volume and intelligence, then it would be expected that older women would be more intelligent than older men. Meinz and Salthouse (1998) examined data from 25 separate studies that compared men's and women's cognitive abilities in old age with that of younger adults. The older group showed many of the same overall patterns of cognitive sex differences that had been reported with younger age groups: older men scored considerably better than older women on spatial tasks. Thus, it seems that this pattern of differential visual–spatial ability persists throughout the life span.

Other areas that have often shown a male advantage include some tests of mathematical ability. The male advantage on advanced tests of mathematical reasoning is probably related to the use of visual–spatial strategies in solving mathematical problems. In one of the most careful studies to date on the relationship between visual–spatial abilities and quantitative abilities, Casey, Nuttall, Pezaris, and Benbow (1995) administered a paper-and-pencil test of mental rotation (the Vandenberg Mental Rotation Test) to 760 males and females for whom they also had scores on the Scholastic Assessment Test-Verbal (SAT-V) and Scholastic Assessment Test-Mathematics (SAT-M). They found that scores on the mental rotation test were related to SAT-M scores for all of the study's participants, except for a group of average to lower-than-average-ability students. These results suggest that there is little or no difference in mathematical performance between male and female participants in the general population—among those of approximately average ability—but males are more likely to outperform females in samples selected for above-average mathematical performance, in part because of their excellent visual–spatial skills (Brody, 1992). The largest male advantage on the SAT-M was found for geometry and word problems, both of which depend on the ability to generate and manipulate the information in a mental representation.

Sex differences, in general, are more prevalent in both extremes of the ability distributions (Willingham and Cole, 1997). Differences in mathematical ability in above-average math students have provided an inequality in the number of females accepted to the most selective universities. It is well documented that on average, males obtain higher scores on standardized tests that are designed to predict grades in college or graduate school (SAT and GRE), yet females, on average, get higher grades (Stricker, Rock, and Burton, 1993; Willingham and Cole, 1997), suggesting that grades and the test scores are not measuring the same construct.

Females show higher levels of ability on a variety of memory tasks. In an examination of sex differences in memory, Stumpf and Jackson (1994) analyzed a battery of tests that examined different aspects of memory. Their subjects were medical school applicants in Germany. They found that women were substantially better on these tests of memory than men (effect size, $d = .56$, over one-half of a standard deviation unit). Stumpf (1995) explained that memory is usually not studied in the context of sex 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. He believes that the size of the female advantage on memory tasks has been underestimated because the tasks that researchers use are unreliable, and because memory is a multidimensional construct. In a later study, Stumpf and Eliot (1995) examined academically talented students in middle and high schools in the United States. In this study, they also found an advantage for females, this time on tests of visual memory.

In a study of healthy adolescents between 16 and 18 years of age, Geffen, Moar, O'Hanlon, Clark, and Geffen (1990) found that girls recalled significantly more words from word lists than did boys. Jensen's (1998) extensive review of multiple tests showed that females scored higher on tests of short-term memory (memories that are approximately 1 to 2 minutes old), with an effect size between $d = .20$ and .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), $d = .54$, and larger working memories, $d = .35$ (Huang, 1993). These differences in some types of memory can be found throughout the life span. Larrabee and Crook (1993) reported than among older adults, females are better than males on several different verbal learning/remembering tasks, name–face associations, the grocery-list selective reminding task (which is exactly what its name implies), and first–last name associates learning (also exactly what its name implies).

Females may 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. Birenbaum, Kelly, and Levi-Keren (1994) reported that females excel at associative memory tasks, once again confirming that, overall, females have better memories than males. Another way of conceptualizing differences in types of memory is to divide memories into episodic—memory for events in one's own life—and semantic—general memory for facts—such as the multiplication tables, historical events that were not experienced personally (e.g., the Revolutionary War), and general word knowledge. Herlitz, Nilsson, and Baeckman (1997) found that females have better episodic memory than do males.

Females also excel in many measures of language usage, including spelling (Stanley, Benbow, Brody, Dauber, and Lupkowski, 1992), reading comprehension (Hedges and Newell, 1995), and writing (U.S. Department of Education, 1997). Of all of the cognitive sex differences, differences in verbal ability are among the first to appear developmentally. Girls aged 1-5 years old are more proficient in language skills than are their male counterparts (McGuiness, 1976; Smolak, 1986). There is also some evidence that girls may talk about 1 month earlier than boys and produce longer utterances than boys (e.g., Gazzaniga, Ivry, and Mangun, 1998; Moore, 1967; Shucard, and Shucard, Thomas, 1987). There are significant sex differences in the rate of vocabulary growth during the toddler years. On average, there is a 13-word difference in vocabulary size between girls and boys at 16 months of age, which grows to a 51-word difference at 20 months and a 115-word difference at 24 months (Huttenlocher, Haight, Bryk, Saltzer, and Lyons, 1991). These researchers found that the differential rate in vocabulary growth was unrelated to how much mothers spoke to their children: mothers spoke as much to their boy babies as they did to their girl babies. Huttenlocher *et al.* (1991) concluded that "gender differences in early vocabulary growth seem to reflect early capacity differences" (p. 245). The female advantage on writing tests led to the inclusion of a writing component on such high stakes tests as the PSAT (a preliminary scholastic assessment test that is usually taken in 11th grade) as a means of increasing gender equity on this test. A test of writing skills will be used as part of the GRE, starting in 1999, to improve overall gender equity on this test as well. Skillful writing is a generative activity that includes good organization of ideas, grammatically correct constructions, and accurate use of words. By almost any standard, it conforms to most people's notions of intelligence, despite the fact that it has been excluded from other intelligence tests.

A PSYCHOBIOSOCIAL MODEL AS AN ALTERNATIVE TO THE NATURE–NURTURE DICHOTOMY

How can psychologists understand the reasons why males and females show different average patterns of performance on tests of cognitive abilities? The age-old distinction between nature (i.e., biological factors) and nurture (i.e., environmental factors) has proved as unproductive in arguments over which sex has the better brain as it has over which sex has the better genitals. A psychobiosocial model offers an alternative conceptualization to the debate. The psychobiosocial model is based on the idea that some variables are both biological and social and, therefore, cannot be classified into one of these two dichotomous categories. For example, learning is a socially mediated event with a biological basis. Individuals are predisposed to learn some topics and skills more readily than others. This predisposition is determined by prior learning experiences (i.e., what is already known and the ease with which the prior learning occurred) and the neurochemical processes that allow learning to occur (e.g., release of neurotransmitters). Modern neuroscience has provided us with the opportunity to identify the areas of the brain that are active during the performance of a cognitive task, which has allowed psychologists to view learning from the inside as well as from the outside of the head (Posner and Raichle, 1994). We can now trace the way in which specific learning experiences alter brain structures, such as dendritic branching and cell size, thus clouding the distinction between environmental and biological events (Ungerleider, 1995).

A variety of data collected throughout the 1990s shows that gonadal hormones, secreted by the ovaries in women and testes in men, have demonstrable effects on the cognitive abilities of women and men (e.g., Kampen and Sherwin, 1996). Conclusions about hormone effects on intelligence are still tentative because the experimental research necessarily involves nonhuman mammals (for ethical reasons) and because many of the inferences for normal humans come from cases where hormone levels are abnormally high or low due to a disease or genetic defect. Converging evidence from a variety of sources supports the idea that prenatal hormone levels affect patterns of cognition in sex-typical ways (e.g., visual–spatial abilities better than verbal abilities for males) and that normal fluctuations in daily and monthly hormone levels in healthy adults correlates with performance on some cognitive tasks (e.g., Hampson and Kimura, 1988).

Estrogen and other gonadal hormones also play an important role in maintaining intelligence in old age. As the average age of the population in the United States has increased, so has the number of women who are taking estrogen replacement therapy (ERT) to prevent bone loss, fractures, and other age-related problems (Smith *et al.,* 1999). Another benefit of ERT that may prove more important than its role in preventing osteoporosis is its beneficial effect on the cognitive abilities of aging women (Schmidt *et al.,* 1996). When compared with control groups of women who did not receive ERT, women with replaced estrogen made fewer errors on tests of short-term visual memory, visual perception, and construction, and showed stable performance over time, whereas the control group declined in their performance (Resnik, Metter, and Zonderman, 1997). At least a dozen studies have found that ERT also reduced the incidence of Alzheimer's disease and markedly reduced the symptoms of Alzheimer's disease in women with mild to moderate dementia (e.g., Henderson *et al.,* 1994; Jacobs *et al.,* 1998; Mortel and Meyer, 1995; Ohkura *et al.,* 1995; Paganini-Hill, Buckwalter, Logan, and Henderson, 1993). Animal studies and human autopsies have shown that ERT has observable effects on the brain. It spurs neuronal growth and increases the speed of communication among the brain's neurons (Simpkins, Singh, and Bishop, 1994). These results may appear to fall solidly under rubric of "biological effects," and may thus seem to contradict the idea that all effects are both biological and environmental. However, there are environmental factors that cause variations in the amount and type of estrogen that is available. Some forms of plant estrogens, such as soybeans, may have similar beneficial effects, although it is still too early to know if this hypothesized relationship will hold up under scientific scrutiny. Cultures with diets that are high in soy products may show how an environmental intervention like dietary choices can have profound effects on biology of intelligence.

Psychologists now know that beliefs about group differences in abilities can affect performance on cognitive tests, usually without the conscious awareness of the test taker (Steele, 1997). When there is a negative stereotype about the performance of one's own group, individuals perform more poorly when the stereotype is "activated" than when it is not. Stereotypes can be activated by such subtle cues as having test takers provide demographic data about their group membership (i.e., male or female). Although psychologists' understanding about the activation of stereotypes is still incomplete, there is a variety of data from other sources suggesting that beliefs about group differences can operate automatically and unconsciously (e.g., Banaji and Hardin, 1996). Thus, even the most careful attempts to manipulate or eliminate the effects of stereotypes on cognitive performance may unknowingly activate the very stereotypes one tries to avoid.

The existence of stereotypes may also seem to contradict the basic premise of the psychobiosocial model—that variables cannot be divided into distinct categories of environmental and biological effects. However, Swim (1994) found that many stereotypes reflect real (i.e., statistically associated) differences, making it impossible to know if cognitive sex differences drive or magnify the stereotypes or if the stereotypes drive or magnify the size of any between-sex differences. This again demonstrates the futility of separating biological and environmental effects.

A psychobiosocial model is preferable over a purely biological or purely environmental model because it allows for multiple processes to operate simultaneously, with environmental inputs altering the biological bases for cognition; which in turn changes the nature of what the individual selects from the environment; which further changes biology, attitudes, and the way individuals make choices. Talented children may receive encouragement and rewards to develop their talent, which provides incentive for them to work at developing it to the fullest, resulting in an increased self-esteem. If males, for example, were predisposed to excel in math for biological reasons, then males would be more likely than females to seek mathematically related experiences, which would improve their skill level, which in turn would make them more motivated to study and persist at difficult mathematical problems. Societal expectations would fuel this cycle, in that males would receive more encouragement, rewards, and opportunities. Neurochemical processes that underlie complex cognition would respond in ways that demonstrate greater efficiency and reduced anxiety and stress. In this way, biological and environmental variables would blur, making the distinction between nature and nurture artificial and meaningless.

THE ROLE OF SCHOOLS: CHANGE OR SHORTCHANGE?

In a report issued by the American Association of University Women (AAUW; 1992), the authors discuss the many ways that "schools shortchange girls." They note many behaviors that teachers engage in, without conscious awareness, that negatively affect the girls in their classes. For example, the authors note that boys get more of the teacher's attention (usually for disruptive behavior), boys call out the answers more often than do girls, and teachers give boys more "thinking time" when they call on them in class and provide them with more encouragement in terms of positive comments. In an angry response to this report, Kleinfeld (1998) wrote, "the idea that the 'schools shortchange girls' is wrong and dangerously wrong" (p. 1). She characterizes the AAUW report as "false political propaganda" (p. 3). Kleinfeld argues that the fact that girls get higher (average) grades in school in all subjects, obtain higher scores on most achievement tests in the early elementary school years and on selected tests in later years, and are less likely to be diagnosed as dyslexic or as a "slow reader," shows that girls are achieving at a higher rate than boys. According to Kleinfeld, schools are not "shortchanging" girls. She goes on to urge that more attention be paid to the achievement of boys, and in the United States, particularly minority boys, because they are dropping out of school at an alarming rate and failing to achieve even basic work skills. Kleinfeld acknowledges that females are seriously underrepresented in some academic disciplines, including physics, chemistry, computer science, and engineering, but she dismisses the importance of this disparity because only a small proportion of the population is employed in these scientific and technical areas, compared to the larger number of boys who are not completing high school.

THE TAKE-HOME MESSAGE

Kleinfeld's (1998) angry response to the AAUW report (1992) brings the question of the smarter sex back into focus, as it raises questions about which differences are real, which are important, and which require social action. Thus, the question of which is the smarter sex is a question about both science and values. Perhaps the most important "take-home message" is to remember that the data reported here are based on group averages, and that no individual is average. As work by Steele (1997) has shown, the danger in reporting sex differences is that they become self-fulfilling prophecies and thus increase the size of differences in the future. The data presented in this article do not support the notion of a smarter sex. There are some cognitive areas in which females, on average, excel, and others where males, on average, excel. Furthermore, these differences are not immutable. The differences that are reported at this time in history could change in unexpected ways in even the near future. Finally, psychologists now know that the brain is altered in response to the environment, even into very old age, so that it no longer makes sense—if, indeed, it ever did—to dichotomize variables into categories of nature or nurture. There are exciting new possibilities for the enhancement of intelligence and for its extension into very old age. New research with ERT and other gonadal hormones offers great hope for an intelligent, aging population. The workforce of the 21st century needs educated citizens who can perform a diverse array of complex tasks. The most important question for the 21st century is how can educators help everyone develop his or her intelligence to the fullest extent possible?

REFERENCES

American Association of University Women. (1992). *How Schools Shortchange Girls: The AAUW Report.* Marlowe, New York.

Astin, A., Sax, L., Korn, W., and Mahoney, K. (1995). *The American Freshman: National Norms for Fall 1995.* Higher Education Research Institute, Los Angeles.

Banaji, M. R., and Hardin, C. D. (1996). Automatic stereotyping. *Psych. Sci.* 7: 136-141.

Birenbaum, M., Kelly, A. E., and Levi-Keren, M. (1994). Stimulus features and sex differences in mental rotation test performance. *Intelligence* 19: 51-64.

Brody, N. (1992). *Intelligence* (2nd ed.). Academic Press, New York.

Brown, L. N., Lahr, C. J., and Mosley, J. L. (1998). Age and gender-related differences in strategy use for route information: A map-present direction-giving paradigm. *Environ. Behav.* 30: 123-143.

Casey, M. B., Nuttall, R., Pezaris, E. and Benbow, C. (1995). The influence of spatial ability of gender differences in mathematics college entrance test scores across diverse samples. *Dev. Psych.* 31: 697-705.

Cowell, P. E., Turetsky, B. I., Gur, R. C., Grossman, R. I., Shtasel, D. L., and Gur, R.

(1994). Sex differences in aging of the human frontal and temporal lobes. *J. Neurosci. 14,* 4748-4755.
Crawford, J. R., and Allen, K. M. (1995). The WAIS-R (UK): Basic psychometric properties in an adult UK population. *Br. J. Clin. Psych.* 34: 237-250.
Dabbs, J. M., Chan, E. L., Strong, R. A., and Milun, R. (1998). Spatial-ability, navigation strategy, and geographic knowledge among men and women. *Evol. Hum. Behav.* 19: 89-98.
DeFries, J. C., and Gillis, J. J. (1993). Genetics of reading disability. In Plomin, R. and McClearn, G. E. (Eds.), *Nature, Nurture and Psychology.* American Psychological Association. Washington, DC. pp. 121-145.
Eals, M., and Silverman, I. (1994). The hunter–gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. *Ethol. Sociobiol.* 15: 95-105.
Feingold. A. (1996). Cognitive gender differences: Where are they and why are they there? *Learn. Indiv. Diff.* 8: 25-32.
Gardner, H. (1983). *Frames of Mind.* Basic, New York.
Gazzaniga, M. S., Ivry, R. B., and Mangun, G. R. (1998). *Cognitive Neuroscience: The Biology of the Mind.* W. W. Norton, New York.
Geffen, G., Moar, K. J., O'Hanlon, A. P., Clark, C. R., and Geffen, L. B. (1990). Performance measures of 16- to 86-year-old males and females on the Auditory Verbal Learning Test. *Clin. Neuropsych.* 4: 45-63.
Goleman, D. (1995). *Emotional Intelligence.* Bantam, New York.
Greenough, W. T., Black, J. E., and Wallace, C. S. (1987). Experience and brain development. *Child Dev.* 58: 539-559.
Gur, R. E., and Gur, R. C. (1990). Gender differences in cerebral blood flow. *Schizoph. Bull.* 16: 247-254.
Halpern, D. F. (1996). Public policy implications of sex differences in cognitive abilities. *Psych. Pub. Pol. Law* 2: 561-574.
Halpern, D. F. (1997). Sex differences in intelligence: Implications for education. *Am. Psych.* 52: 1091-1102.
Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd ed.). Erlbaum, Mahwah, NJ.
Hampson, E., and Kimura, D. (1988). Reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. *Behav. Neurosci.* 102: 456-495.
Hare-Mustin, R. T., and Marecek, J. (1994). Asking the right questions: Feminist psychology and sex differences. *Fem. Psych.* 4: 531-537.
Hedges, L. V., and Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269: 41-45.
Henderson, V. W., Paganini, H. A., Emanuel, C. K., Dunn, M. E., and Buckwalter, J. G. (1994). Estrogen replacement therapy in older women: Comparisons between Alzheimer's disease cases and nondemented control subjects. *Arch. Neurol.* 51: 896-900.
Henning-Stout, M., and Close-Conoley, J. (1992). Gender: A subtle influence in the culture of the school. In Medway, F. J. and Cafferty, T. P. (Eds.), *School Psychology: A Social Psychological Perspective* (pp. 113-135). Hillsdale, NJ: Erlbaum.
Herlitz, A., Nilsson, L-G., and Baeckman, L. (1997). Gender differences in episodic memory. *Mem. Cog.* 25: 801-811.
Hollway, W. (1994). Beyond sex differences: A project for feminist psychology. *Fem. Psych.* 4: 538-546.
Horn, J. L. (1985). *Psychometric Studies of Aging and Intelligence.* Raven Press, New York.
Huang, J. (1993). An investigation of gender differences in cognitive abilities among Chinese high school students. *Person. Indiv. Diff.* 15: 717-719.
Hunt, E., Pellegrino, J. W., Frick. R. W., Farr, S. A., and Alderton, D. (1988). The ability to reason about movement in the visual field. *Intelligence* 12: 77-100.
Hunt, M. (1999). *The New Know-Nothings.* Transaction, New Brunswick.
Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., and Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Dev. Psych.* 27: 236-248.
Jacobs, D. M. Tang, M. X., Stern, Y., Sano, M., Marder, K., Bell, K. L., Schofield, P., Dooneief,

G., Gurland, B., and Mayeux, R. (1998). Cognitive function in nondemented older women who took estrogen after menopause. *Neurology* 50: 368-373.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harv. Edu. Rev.* 39: 1-123.

Jensen, A. R. (1998). *The g Factor: The Science of Mental Ability.* Praeger Publishers, Westport, CT.

Kampen, D. L. and Sherwin, B. B. (1996). Estradiol is related to visual memory in healthy young men. *Behav. Neurosci.* 110: 613-617.

Kleinfeld, J. (1998). *The Myth that Schools Shortchange Girls: Social Science in the Service of Deception.* The Women's Freedom Network, Washington, DC.

Larrabee, G. J., and Crook, T. H., III. (1993). Do men show more rapid age-associated decline in simulated everyday verbal memory than do women? *Psych. Aging* 8: 68-71.

Linn, M. C., and Petersen, A. C. (1986). A meta-analysis of gender differences in spatial ability. Implications for mathematics and science achievement. In Hyde, J. S. and Linn, M. C. (Eds.). *The Psychology of Gender: Advances through Meta-Analysis.* The Johns Hopkins University Press, Baltimore, pp. 67-101.

Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Person. Indiv. Diff.* 17: 257-271.

Loring-Meier S., and Halpern, D. F. (1999). Sex differences in visual–spatial working memory: Components of cognitive processing. *Psychonom. Bull. Rev.* 6: 464-471.

Lynn, R. (1998). Has the black–white intelligence difference in the United States been narrowing over time? *Person. Indiv. Diff.* 25: 801-804.

Makintosh, N. J. (1996). Sex differences in IQ. *J. Biosoc. Sci.* 28: 559-571.

Masters, M. S., and Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behav. Gen.* 23: 337-341.

Meinz, E. J., and Salthouse, T. A. (1998). Is age kinder to females than to males? *Psychonom. Bull. Rev.* 5: 56-70.

McGuiness, D. (1976). Sex differences in the organization of perception and cognition. In Lloyd, B. and Archer, J. (Eds.). *Exploring Sex Differences.* Academic Press., New York. pp. 123-156.

Moore, T. (1967). Language and intelligence: A longitudinal study of the first eight years. *Hum. Dev.* 10: 88-106.

Mortel, K. F., and Meyer, J. S. (1995). Lack of menopausal estrogen replacement therapy and the risk of dementia. *J. Neuropsych. Clin.* 2: 334-337.

Ohkura, T., Isse, K., Akazawa, K., Hamamoto, M., Yaoi, Y., and Hagino, N. (1995). Long-term estrogen replacement therapy in female patients with dementia of the Alzheimer type: 7 case reports. *Dementia* 6: 99-107.

Paganini-Hill, A., Buckwalter, J. G., Logan, C. G., and Henderson, V. W. (1993). Estrogen replacement and Alzheimer's disease in women. *Soc. Neurosci. Abst.* 19: 1046.

Posner, M. I., and Raichle, M. E. (1994). *Images of Mind.* New York: Freeman.

Resnick, S. M. (1993). Sex differences in mental rotations: An effect of time limits? *Brain Cog.* 21: 71-73.

Resnick, S. M., Metter, E. J., and Zonderman, A. B. (1997). Estrogen replacement therapy and longitudinal decline in visual memory: A possible protective effect? *Neurology* 49: 1491-1497.

Robinson, N. M., Abbott, R. D., Berninger, V. W., and Busse, J. (1996). The structure of abilities in math-precocious young children: Gender similarities and differences. *J. Edu. Psych.* 88: 341-352.

Schiff, W., and Oldak, R. (1990). Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. *J. Exper. Psych. Hum. Percep. Perform.* 16: 303-316.

Schmidt, R., Fazekas, F., Reinhart, B., Kapeller, P., Fazekas, G., Offenbacher, H., Eber, B., Schumacher, M., and Freidl, W. (1996). Estrogen replacement therapy in older women: A neuropsychological and brain MRI study. *J. Am. Geri. Soc.* 44: 1307-1313.

Shucard, D. W., Shucard, J. L., and Thomas, D. G. (1987). Sex differences in electrophysiological activity in infancy. Possible implications for language development. In Phillips, S. U., Steele, S., and Tanz, C. (Eds.), *Language, Gender, and Sex in Comparative Perspectives.* Cambridge University Press, Cambridge, pp. 278-295.

Simpkins, J. W., Singh, M., and Bishop, J. (1994). The potential role for estrogen replacement therapy in the treatment of the cognitive decline and neurodegeneration associated with Alzheimer's disease. *Neurobiol. Aging* 15: S195-S197.

Smith, C. A., McCleary, C. A., Murdock, G. A., Wilshire, T. W., Buckwalter, D. K., Bretsky, P., Marmol, L., Gorsuch, R. L., and Buckwalter, J. G. (1999). Lifetime estrogen exposure and cognitive performance in elderly women. *Brain Cog.* 39: 203-218.

Smith, G. A., and McPhee, K. A. (1987). Performance on a coincidence timing task correlates with intelligence. *Intelligence* 11: 161-167.

Smolak, L. (1986). *Infancy,* Prentice-Hall, Englewood Cliffs, NJ.

Stanley, J. C., Benbow, C. P., Brody, L. E., Dauber, S., and Lupkowski, A. (1992). Gender differences in eighty-six nationally standardized aptitude and achievement tests. In Colangelo, N., Assouline, S. G., and Ambroson, D. L. (Eds.), *Talent Development, Vol. 1: Proceedings from the 1991 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development.* Trillium Press, Unionville, NY, pp. 42-65.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *Am. Psych.* 52: 613-629.

Sternberg, R. J., Wagner, R. K., Williams, W. M., and Horvath, J. A. (1995). Testing common sense. *Am. Psych.* 50: 912-927.

Stricker, L. J., Rock, D. A., and Burton, N. W. (1993). Sex differences in predictions of college grades from Scholastic Aptitude scores. *J. Edu. Psych.* 85: 710-718.

Stumpf, H. (1995). Gender differences in performance on tests of cognitive abilities: Experimental design issues and empirical results. *Learn. Indiv. Diff.* 7: 275-287.

Stumpf, H., and Eliot, J. (1995). Gender-related differences in spatial ability and the k factor of general spatial ability in a population of academically talented students. *Person. Indiv. Diff.* 19: 33-45.

Stumpf, H., and Jackson, D. N. (1994). Gender-related differences in cognitive abilities: Evidence from a medical school admissions testing program. *Person. Indiv. Diff.* 17: 335-344.

Stumpf, H., and Stanley, J. C. (1996). Gender-related differences on the College Board's Advanced Placement and Achievement Tests, 1982-1992. *J. Edu. Psych.* 88: 353-364.

Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *J. Person. Soc. Psych.* 66: 21-36.

U. S. Department of Education. (1997). *National Assessment of Educational Progress* (*Indicator 32: Writing proficiency: Prepared by the Educational Testing Service*). Washington, DC. (World Wide Web: http//www.ed/gov/nces)

Ungerleider, L. G. (1995). Functional brain imaging studies of cortical mechanisms for memory. *Science* 270: 769-775.

Voyer, D., Voyer, S., and Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psych. Bull.* 117: 250-270.

Weiss, L. G., and Prifitera, A. (1995). An evaluation of differential prediction of WIAT achievement scores from WISC-III FSIQ across ethnic and gender groups. *J. School Psych.* 33: 297-304.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psych. Bull.* 120: 272-292.

Willingham, W. W., and Cole, N. S. (1997). *Gender and Fair Assessment.* Erlbaum, Hillsdale, NJ.