BEHAVIORAL AND BRAIN SCIENCES (1988) 11, 169–232
Printed in the United States of America

Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes

Camilla Persson Benbow
Study of Mathematically Precocious Youth (SMPY), Department of Psychology, Iowa State University, Ames, Iowa 50011-3180

Abstract: Several hundred thousand intellectually talented 12- to 13-year-olds have been tested nationwide over the past 16 years with the mathematics and verbal sections of the Scholastic Aptitude Test (SAT). Although no sex differences in verbal ability have been found, there have been consistent sex differences favoring males in mathematical reasoning ability, as measured by the mathematics section of the SAT (SAT-M). These differences are most pronounced at the highest levels of mathematical reasoning, they are stable over time, and they are observed in other countries as well. The sex difference in mathematical reasoning ability can predict subsequent sex differences in achievement in mathematics and science and is therefore of practical importance. To date a primarily environmental explanation for the difference in ability has not received support from the numerous studies conducted over many years by the staff of Study of Mathematically Precocious Youth (SMPY) and others. We have studied some of the classical environmental hypotheses: attitudes toward mathematics, perceived usefulness of mathematics, confidence, expectations/ encouragement from parents and others, sex-typing, and differential course-taking. In addition, several physiological correlates of extremely high mathematical reasoning ability have been identified (left-handedness, allergies, myopia, and perhaps bilateral representation of cognitive functions and prenatal hormonal exposure). It is therefore proposed that the sex difference in SAT-M scores among intellectually talented students, which may be related to greater male variability, results from both environmental and biological factors.

Keywords: environment; giftedness; heredity; hormones; intelligence; laterality; mathematical ability; nature/nurture; sex differences; spatial relations

My emotions while setting out to write this *BBS* target article are well reflected in the words of Adelson (1985): "I can think of few activities more enervating emotionally than to survey the psychological literature on sex differences. I first did so about 15 years ago, soon after the birth of contemporary American feminism, and was taken aback by the gap between the actual, enfeebled state of our knowledge, and the dogmatic self-assertions of so much being written on the topic" (p. 9). Adelson continued to state that the literature on the subject has not changed much since then. A great amount of work has been done, but it has not been accompanied with much growth in understanding (Adelson 1985).

Studies of mathematical ability and achievement have consistently found sex differences favoring males (Backman 1972; Benbow & Stanley 1980; 1981; 1982a; 1983b; 1984; Bieri et al. 1958; College Board 1985; Deaux 1985; Ernest 1976; Fennema 1974; Fox 1976; Fox et al. 1983; Fox et al. 1980; Garai & Scheinfeld 1968; Glennon & Callahan 1968; Keating 1974; Maccoby & Jacklin 1974; NAEP 1975; Suydam & Weaver 1970; Very 1967; Wilson 1972). In the mid-seventies there was heightened concern about the impact of these differences on women's career options (Datta 1985). The National Institute of Education set up a grants program to fund research on the topic (NIE 1977), which generated a great amount of work in the late 1970s. The work done at that time and in the early 1980s may have raised the level of consciousness of women and thus affected their behavior in relation to mathematics, but it did not reduce sex differences in mathematical aptitude. In 1987 there are still reports of these differences in ability (e.g., Holden 1987; Jones 1987; Moore & Smith 1987). A comprehensive survey of recent achievement test data, obtained from a wide domain of types and levels of tests, revealed the extent of these sex differences in educational achievement in mathematics and the sciences among secondary school and college students (Stanley et al. 1987). Effect sizes (computed by dividing the difference between the male and female means by the standard deviation) were determined for each test. Males scored considerably higher than females on both computer science tests (.62, .50); on all 6 physics tests (.59, .56, .51, .41, .37, .29); on all 4 chemistry tests (.51, .39, .37, .33); on all 4 biology tests (.36, .29, .25, .22); on both general science tests (.40, .38); and on 12 of the 16 quantitative tests (.71, .67, .43, .41, .39, .38, .37, .34, .20, .18, .04, .04). Females scored higher (.16, .16, .15, .01) on the DAT Numerical Ability Test in four of the five school years it was given (Stanley 1987).

0140-525X/88 $5.00+.00

It is in junior high school that the sex difference in mathematics first becomes apparent. Girls excel in computation, boys on tasks requiring mathematical reasoning, and no differences are seen in the ability to apply learned concepts or algorithms (Fennema 1974). As would be predicted from this pattern, sex differences in mathematics emerge in the United States when the mathematics curriculum becomes somewhat abstract. In addition, the absolute size of the sex difference is largest among the intellectually talented. (It is also of interest, however, that more males than females are learning disabled in mathematics [Geschwind & Behan 1982].)

This article attempts to describe the sex difference in scores on the mathematics section of the Scholastic Aptitude Test (SAT-M) among the intellectually talented, the relation of this sex difference to other attributes and achievement behaviors, and the work that has been done to understand this sex difference better.

1. The Scholastic Aptitude Test (SAT)

The SAT is administered several times a year to more than a million 16- to 18-year-old applicants by Educational Testing Service in Princeton, New Jersey, under the auspices of the College Board. It is a multiple-choice test of abilities which have developed over years of study and use, rather than a test of primarily learned subject matter. It was designed for abler twelfth-grade students throughout the country, most of whom are 17 or 18 years old (Angoff 1971).

One part of the SAT is the verbal section, or the SAT-V. It has four types of questions: 25 antonyms, 20 analogies, 15 sentence completions, and 25 questions based on reading passages. The chief purpose of the SAT-V is to test the type of reading and verbal reasoning ability needed for success in college. The following is a sample item from the SAT-V:

> Bequeath:Legacy::(A) achieve:goal (B) worship:idol (C) enforce:law (D) endow:gift (E) endure:pain.

The other major part of the test is the SAT-M, which tests developed mathematical reasoning ability. There are 40 standard multiple-choice items on arithmetic, algebra, geometry, and miscellaneous areas as well as 20 quantitative comparison questions, all to be solved in 60 minutes. Some of the questions require the application of numerical, graphic, spatial, symbolic, and logical techniques to familiar situations. Only one year of algebra and some knowledge of geometry is necessary to solve the problems on the SAT-M.

The following is a sample item from the SAT-M:

> A man willed $\frac{2}{5}$ of his estate to his daughter and $\frac{1}{20}$ to each of his three sons. Another $\frac{1}{10}$ was divided equally among his five grandchildren. If the remainder of the man's estate was left to his wife, what part of his estate did his wife receive?
>
> (A) $\frac{1}{5}$ (B) $\frac{3}{10}$ (C) $\frac{7}{20}$ (D) $\frac{2}{5}$ (E) $\frac{9}{20}$

2. Use of the SAT with young students

In 1971 Professor Julian C. Stanley founded the Study of Mathematically Precocious Youth (SMPY) and pioneered the use of the SAT, especially the SAT-M, with intellectually talented 12- to 13-year-olds (Keating & Stanley 1972). His purpose in using the SAT was to identify mathematically talented individuals in order to help them educationally. A secondary objective was to study these students' development.

To identify mathematically talented 12- to 13-year-olds, a talent search strategy was developed. Students in the seventh grade (and, in the first three years of conducting talent searches, students in the eighth grade as well) were eligible to participate in a talent search if they had scored in the upper 5% in 1972, upper 2% in 1973 or 1974, and upper 3% thereafter on national norms for the mathematics part of a standardized achievement test administered by the students' schools as part of their regular testing program. The mathematics section of the Iowa Test of Basic Skills would be one acceptable screening test. All middle or junior high schools in the talent search region (i.e., the Mid-Atlantic area) were informed of the talent search (through letters to the principal, counselor, and mathematics department chair) and were asked to give an application form to all those students who qualified. Students applied for the talent search and then took the test at a national administration of the SAT in their area.

3. What is the meaning of the SAT-M score?

Although the College Board reports that the SAT measures developed mathematical reasoning ability, we do not know exactly what the SAT-M measures, especially among these young students. It was hypothesized that when the SAT-M is given to gifted 12-year-olds rather than high school students, it is a much stronger measure of mathematical reasoning. The young students have not had much experience with abstract mathematics and have not been exposed to the content of the test. Thus, they must figure out by themselves how to solve the problems (Benbow & Stanley 1981; 1983b; Stanley & Benbow 1986).

To test this hypothesis, a factor analysis of the SAT-M was performed separately for male and female 7th and 12th graders (Minor & Benbow 1986). We are presently investigating the item difficulties on that form of the SAT-M for those 7th and 12th graders. Both studies revealed statistically significant age effects but no sex differences on item difficulties. Moreover, the age differences in performance were on item types that would be predicted from the above reasoning hypothesis.

We could approach this point from another perspective. Mayer (1985) delineated four steps in mathematical problem solving: problem translation, problem integration, solution planning, and solution execution. Problem integration requires specific knowledge of problem types. When students lack a schema for a problem type, they are more prone to error. Moreover, lack of strategic knowledge may relate to failure in appropriate solution planning. The gifted seventh graders in the talent searches had not been explicitly taught the schemas or the strategic knowledge necessary to solve the problems posed on the SAT-M. Yet they were successful nevertheless. Why? It may be because they are especially good at translating the problem statements and representing relational propositions. This again seems to imply that gifted seventh graders are able to score well on the

Table 1. *Mean SAT-M and SAT-V scores of SMPY's talent search students, 1972–1979, by sex* (N = 9,927)

Talent search date	Grade	Sex	*N*	SAT-M score Mean	SAT-M score S.D.	SAT-V score Mean	SAT-V score S.D.
March 1972	7th	M	90	460	104		
	7th	F	77	423	75		
	8th	M	133	528	105		
	8th	F	96	458	88		
Jan. 1973	7th	M	135	495	85	385	71
	7th	F	88	440	66	374	74
	8th	M	286	551	85	431	89
	8th	F	158	511	63	442	83
Jan., Feb. 1974	7th	M	372	473	85		
	7th	F	222	440	68		
	8th	M	556	540	82		
	8th	F	369	503	72		
Dec. 1976	7th	M	495	455	84	370	73
	7th	F	356	421	64	368	70
Jan. 1978[a]	7th	M	1,549	448	87	375	80
	7th	F	1,249	413	71	372	78
Jan. 1979[a]	7th	M	2,046	436	87	370	76
	7th	F	1,628	404	77	370	77

[a] Some accelerated eighth graders were included.

Source: Adapted from Table 1 in Benbow and Stanley (1980).

SAT-M if they have well-developed mathematical reasoning ability.

Moreover, in a review of the literature on the nature of mathematical talent, all definitions studied dealt with higher-level cognitive processing requiring this mathematical reasoning ability (Benbow 1988). Computational ability was not viewed as an essential aspect of mathematical talent. I concluded then that "mathematical talent was best defined as the ability to handle long chains of reasoning" (Benbow 1988). In addition, mathematical reasoning ability is important for high-level achievement in mathematics and the sciences (Stanley & Benbow 1986). Thus, SMPY's use of the SAT to identify mathematical talent seemed justified.

Nonetheless, we do not have a theory to explain or define mathematical talent or mathematical reasoning ability, nor do we currently have data to suggest one. Moreover, we do not know the underlying cognitive abilities contributing to this mathematical reasoning. A few researchers, however, have investigated which cognitive abilities or characterisitcs correlate with mathematical talent. These are spatial ability, field independence, use of images, logicism, intuition, flexibility, the ability to recognize unproductive strategies, excellent memory, and high verbal and reasoning skills (Benbow 1988). Plans are in progress to investigate the role of these factors in mathematical reasoning and to address questions such as how mathematical information is represented or the speed with which mathematical information is accessed and/or manipulated by mathematically talented adolescents. There is already some evidence that differences in problem-solving knowledge are closely related to memory capacity and speed (Chi 1978).

The development of mathematical talent in children has not been studied much (Benbow 1988). Because a major difference between the mathematically talented and other students may be that the talented acquire skills and competences at an earlier age (i.e., they are precocious) (Robinson 1983), studies on the development of mathematical thinking in average children (e.g., Ginsburg 1983) may be relevant.

4. Sex differences on the SAT-M

Perhaps the most unexpected finding arising out of SMPY's identification of mathematical talent was the early discovery of large sex differences in SAT-M scores. When the first talent search was conducted in 1972, a large difference was found. That year almost 8% of the seventh-grade boys and 27% of the eighth-grade boys scored over 600 on the SAT-M, whereas not one girl did. The mean difference in scores was 37 points for the seventh graders and 70 points for the eighth graders. It was initially believed that the result was an artifact or, if not that, purely an effect of environmental factors. Thus, when a sex difference was found the following year in the talent serach, the staff of SMPY were rather surprised. In fact, a sex difference has since been found in every talent search conducted by SMPY (Benbow & Stanley 1980). The results for SMPY's six talent searches involving 9,927 students are shown in Table 1. Although there were no statistically significant sex differences in verbal ability, as measured by the SAT-V, in any of the six talent searches, there was a consistent sex difference on the SAT-M. On the average the boys scored about a half standard deviation higher on the SAT-M than the girls did even though, as explained above, they had been matched previously on mathematics achievement test scores.

Although systematic sampling bias seemed unlikely because of both the matching via the achievement test and the lack of difference on the SAT-V, we decided to check whether the talent searches attracted a more select group of boys than girls, especially since there were more boys than girls (57% vs. 43%) participating. Thus, the talent search girls' performance on the SAT was compared against the norms for college-bound female high school seniors, and the talent search boys' performance was checked against college-bound senior male norms. The outcome was that the girls in the SMPY talent searches comprised a more select group than the boys did (Benbow & Stanley 1980). That is, the percentile rank of the SAT-M scores of talent search girls on female norms was somewhat higher than that of the talent search boys on male norms. Thus, it seemed illogical to conclude that we had tested a more able group of boys than girls and thereby artificially produced a sex difference.

There were several limitations in the data from the SMPY talent searches. The talent searches were in the area of mathematics only. Thus, the results were only generalizable to highly motivated, mathematically talented boys and girls. Moreover, in the initial talent searches the 10,000 cases tested were insufficient to study the distribution of especially high scores. Sex differences in the proportion of high scorers on the SAT-M could have practical significance.

Then in 1980 the Center for the Advancement of Academically Talented Youth (CTY) at Johns Hopkins University began conducting talent searches not only in mathematics but also in verbal and general ability. Thus, the top 3% of seventh graders in verbal, overall, or mathematical ability on standardized national achievement tests were eligible for the talent searches. As a result, a more general sample of gifted students was being tested each year. The staff of CTY also expanded the talent searches from the Mid-Atlantic area to the Northeast. Benbow and Stanley (1983b) reported the results for CTY's first three talent searches and the preliminary results for the fourth. Moreover, data from SMPY's national search for students who represent the top 1 in 10,000 in mathematical reasoning ability by age 13 (i.e., those who score at least 700 on the SAT-M before age 13) were also presented.

Modifications of the regional talent search produced equal numbers of girls and boys being tested. Yet the mean sex difference remained constant at 30 points favoring the males among the sample of 19,883 boys and 19,937 girls (Benbow & Stanley 1983b). Although important, this was not the major thrust of the subsequent analysis. Rather, enough cases were now available to study the distribution of high scores for males and females. It was found that among 12-year-old students scoring ≥ 500 on the SAT-M (the approximate average score of college-bound 17-year-old males), there is a ratio of two males (2.1) for every female (based on 5,325 cases); at ≥ 600 on the SAT-M (77th percentile of 1984–1985 college-bound 12th-grade males) the ratio is 4.1 to 1 (based on 806 such cases); and at ≥ 700 on the SAT-M (the 94th percentile of 1984–1985 college-bound senior males) the ratio is 12.9 to 1 for the 278 cases reported in Benbow and Stanley (1983b). When in November 1983 SMPY had temporarily completed its national search for students scoring at least 700 on the SAT-M before age 13, 23 such girls and 268 such boys had been identified. Subsequently, one more male was added. Nonetheless, the ratio remained around 12 to 1.

The above ratios had been anticipated since the mean and variance of the male scores were larger than for the female scores. Yet since the SAT scores were somewhat positively skewed for males, it was not possible to predict the ratios precisely. Moreover, when low scores on the SAT-M were studied, the ratio of females to males did not increase much as scores diminished (i.e., about 1.5 girls per boy were in score ranges less than 400). Nevertheless, even if one concludes that our findings result primarily from greater male variability, one must still explain why males are more variable.

In sum, Benbow and Stanley have reported their findings based on 49,747 intellectually talented students who were tested between and including 1972 and 1982. In this large sample it is abundantly clear that *far* more boys than girls (chiefly 12 year olds) scored high on the SAT-M, even though the girls were matched with boys for ability, age, grade, and talent search entry. The talent searches from which these data were derived are still continuing and now cover the entire United States. Almost 100,000 young adolescents are being tested every year through the talent searches conducted by Johns Hopkins, Duke, and Northwestern universities and the University of Denver. Every year the sex difference in mathematical reasoning ability is apparent and of the same magnitude as reported by Benbow and Stanley (1980; 1983b).

5. Is the sex difference decreasing?

In a short article in *Psychological Reports* it was suggested that the sex difference Stanley and I had reported was actually getting smaller each year and that in a few years it would be nonexistent (Freed 1983). We showed then that this was not the case (Stanley & Benbow 1983). For example, in the Johns Hopkins' 1983 Talent Search the sex difference on the SAT-M was 31 points based on 7,561 males and 7,918 females (CTY 1983). In 1984 this sex difference had increased to 36 points for the 9,680 males and 9,543 females (CTY 1984). In 1985 the difference was 34 points for the seventh graders and 49 points for the small number of eighth graders tested (12,783 males and 12,796 females) (CTY 1985). In the 1986 Talent Search the sex difference was 32 points for the 12,085 males and 11,647 females (CTY 1986). Moreover, the ratios of high scoring boys to girls as well as the differences have remained relatively constant over the 15 years. Thus, sex differences in SAT-M scores among young adolescents are not temporary trends. They have been stable even in times of great change in attitudes toward women.

6. Geometric proofs

Senk and Usiskin (1983) suggested that perhaps the SAT-M is not a very sensitive measure of mathematical reasoning ability. They felt that the ability to formulate proofs in geometry reflects that ability much more closely. Thus, they studied boys' and girls' performance in geometric proofs after a year-long course teaching that skill. No sex differences were found. It should be noted, however, that this measure was testing the ability to develop geometric proofs after prolonged instruction. Thus, the ability measured by Senk and Usiskin is very different from SMPY's mathematical reasoning ability. Senk and Usiskin (1983) thus provided further evidence that no sex differences are found in the ability to apply already-learned knowledge (Fennema 1974).

7. Sex differences in ethnic groups, other countries, and before seventh grade

Sex differences in SAT-M scores among 12 year olds are not limited to the United States. The SAT-M was translated into German and into Mandarin Chinese. These respective translations were then administered in Hamburg, West Germany, and Shanghai, China, to intellectually talented preadolescents. In both countries there were sex differences favoring males (Durden, personal communication; Stanley et al. 1986). Stanley et al. (1986) found, for example, that there were four times as many males than females scoring above 700 on the SAT-M before age 13 in Shanghai. Among Asian Americans in SMPY's sample of such students, the ratio is also 4 to 1. Within the United States the magnitude of the sex difference in mathematics does indeed appear to vary among different ethnic groups, but it is present in every case (Jones 1987; Moore & Smith 1987). Differences were largest among Hispanics but smallest among blacks (Moore & Smith 1987). Moreover, sex differences in mathematical ability have been found in several other countries (Kelly 1979). Within each country there was a sex difference in mathematics favoring males. Girls of one country, however, scored better in some cases than boys of a different country. This and the ethnic group studies were not limited to intellectually talented students, however.

Thus, there is strong evidence for the existence of sex differences in mathematical reasoning ability by age 12. It is rather difficult to obtain such data below that age because there are no tests of mathematical reasoning ability for younger students. This is probably because the elementary curricula tend to cover mainly computation and basic arithmetic facts. Yet Dougherty et al. (1980) and NAEP (1975) have found sex differences in mathematical reasoning even among 7- and 9-year-old children by looking at selected test items. Moreover, Marshall (1983) found that the mistakes made by elementary school boys and girls on the mathematics part of the Survey of Basic Skills were different in nature. Girls were more likely than boys to make errors by misusing spatial information, using irrelevant rules, negative transfer, inappropriate key word associations, or the choice of incorrect operations. Boys, in contrast, were more likely than girls to make errors of perseverance and formula interference. Differences in the development of mathematical reasoning ability possibly could explain these findings.

Clearly then, there are substantial sex differences in mathematical reasoning ability among intellectually talented students. We do not know how these findings may relate to students of average ability. Differences may be smaller at that level. Nevertheless, most of the concern about the lack of participation of females in higher levels of mathematics (e.g., Ernest 1976) has focused on intellectually able girls, rather than those of average ability. Average students tend not to major in the sciences at the academically difficult colleges that produce most top-level scientists (Davis 1965; Werts 1967). Accordingly, the students Stanley and I studied are an appropriate opportunistic sample of the population.

8. Consequences of sex differences in mathematical reasoning ability

It might be thought that the sex differences reported by Benbow and Stanley (1980; 1983b) are inconsequential and have no lasting influence. The longitudinal study conducted by SMPY to investigate the development of intellectually talented students, especially mathematically gifted ones, provides an opportunity to test this hypothesis. Results have already been published on the post-high school follow-up of students in SMPY's first three talent searches who as seventh or eighth graders had scored at least 370 on the SAT-V or 390 on the SAT-M, the average scores of high school females in the early 1970s (Benbow 1981; Benbow & Minor 1986; Benbow & Stanley 1982a; 1984). These students have now been followed up one year after expected college graduation. Moreover, a much more sophisticated follow-up study at the end of high school was conducted on a subset of participants in the last three SMPY talent searches.

The results of already published work will now be summarized together with some preliminary results of the other surveys. Instead of a list of means and standard deviations for each variable by sex, the effect sizes (i.e., difference between means divided by the standard deviation) were computed and reported. These effect sizes were then classified as small ($\geq .2$), medium ($\geq .5$), and large ($\geq .8$) by Cohen's (1977) criteria.

First, the high school achievements of participants in the first three talent searches were investigated. It was found that the sex difference in SAT-M scores persisted and was related to subsequent sex differences in achievement in mathematics and science (Benbow & Minor 1986; Benbow & Stanley 1982a; 1984). Both the SAT-M and SAT-V scores of males improved by at least 10 points more than those of females during this time. The difference on the SAT-M increased from 40 points at the time of the talent search to 50 points at the end of high school. The latter 50 point difference is almost exactly the sex difference reported for college-bound seniors (College Board 1985). Rather surprising was the greater gains of males than females on the SAT-V as well. If there had been a greater verbal emphasis for or by females than males during the high school years, which was detrimental to their mathematical aptitude and/or achievement, then at least females' verbal abilities should have benefited.

When talent search students were studied in seventh or eighth grade, only negligible sex differences (as determined by the small or nonexistent effect sizes) had been found in attitudes toward mathematics and science, in perceptions of relative standing within the mathematics class, and in how their mathematical knowledge had been acquired (Benbow & Stanley 1982b). In high school, mathematically talented youth, whether male or female, tended to have favorable attitudes toward mathematics and science and to participate in mathematics and the sciences at a level much higher than average (Benbow & Minor 1986; Benbow & Stanley 1982a). There were small sex differences favoring boys in course-selection and favoring girls in course grades. Indications of more substantial sex differences favoring males were found, however, in participation in high school calculus (67% vs. 40%) and physics (76% vs. 58%), participation and performance on high school and college-level mathematics and science achievement tests (effect sizes were generally greater than .4), and intention to major in the more quantitatively oriented fields of engineering and physics (32% vs. 15%). There were no sex differences in intention to major in mathematics (15% of males and 17% of females). No substantial sex differences in attitudes toward mathematics and the sciences, except physics (effect size, .5, medium), were detected at the end of high school. Although overall attitude toward science did relate somewhat to participation in science (range of *r*: .20–.37), attitudes toward mathematics did not relate to participation in it. Overall, for the mathematics and science areas, the best predictor of high school participation and achievement was the seventh or eighth grade SAT-M score, not sex, background, or attitudinal variables.

Several years after the completion of the above follow-up study, a second, more detailed after-high school follow-up was initiated. A much more select (in ability) group of seventh graders from SMPY's last three talent searches (1976, 1978, and 1979) was surveyed. At present only preliminary results on 508 students are available. Judging from them, however, basic support was found for the conclusions drawn from the initial longitudinal study (i.e., Benbow & Minor 1986; Benbow & Stanley 1982a). Achievement in mathematics and science during high school for both the males and females was especially high and much higher than the earlier group of talent search participants who had been studied in high school. Nonetheless, sex differences emerged, but none in course-taking in mathematics or science or in achievement scores in biology at any level (which favored females slightly but not significantly). Many more of the females in this cohort took calculus in high school compared to the previous cohort (72% vs. 40%), and only a small sex difference was apparent.

Yet in this cohort as well, males attained much higher scores than females on the high school and college-level achievement tests in mathematics, chemistry, and physics, even though many more boys than girls took these tests. Although only two effect sizes were below .5, the range of effect sizes for the sex differences was from .36 to .72 (except for a test where only 31 scores for the males and 5 for the females had been reported). Moreover, there were large sex differences in SAT-M scores favoring the males (effect sizes were 1.13 in the talent search and .82 at the end of high school). Although the sex difference in SAT-M scores did not increase during the high school years for this group, this result may be spurious because the SAT-M did not have a high enough ceiling for the males at the end of high school (i.e., the mean SAT-M score for males was about 750, close to the highest possible score). In terms of attitudes, only small differences were seen in mathematics and none in the natural sciences, but fairly large ones (effect size, .51) favoring males in the physical sciences. It accordingly seems that the earlier sex differences described above (i.e., Benbow & Minor 1986; Benbow & Stanley 1982a) were duplicated in the later study, although slight variations existed and the later cohort demonstrated higher achievement levels.

Preliminary data are now also available on the first cohort of students (described above) five years beyond high school graduation (Benbow 1987a). It was previously documented that the SMPY students attended academically strong colleges or universities (Benbow 1983). Whereas 90% had initially entered college, 84% had now received their bachelor degrees, an extraordinarily high percentage. Moreover, SMPY students completed college with quite outstanding academic records, female grades being somewhat higher. In addition, approximately 47% of the SMPY students are furthering or have furthered their education beyond college. In this case, however, more SMPY males than females were attending graduate school, especially at the doctoral level (38% vs. 24%). In terms of mathematics and science achievement, 37% of the SMPY females and 59% of the SMPY males had majored in those areas. This represented a large decrease among the females (i.e., 50% to 37%; $p < .01$). Furthermore, among those students who continued their education beyond the bachelor's degree, 22% of the females and 41% of the males were enrolled in mathematics or science departments. These statistically significant ($p < .01$) gender differences in science majors were primarily due to the fact that fewer SMPY females than males chose engineering, computer science, and the physical sciences. There were no gender differences, however, in the proportion majoring in mathematics or biology. Finally, when these students' long-range career goals were classified according to type and area, almost 40% of the SMPY males and 26% of the SMPY females had career goals in mathematics and the sciences. Somewhat more surprisingly, however, SMPY males and females differed significantly in only one respect in what they wished to do (which was broadly defined) within their chosen area. SMPY males were almost twice as likely as the SMPY females to choose research careers ($p < .01$). These preliminary data illustrate that sex differences in mathematics and science achievement persist during college and the post-college years among intellectually talented students (Benbow 1987a).

9. Socialization and sex differences

In the above sections it was shown that there is a large sex difference in the scores on the SAT-M and in achievement in mathematics and the sciences among the intellectually talented population from which this nation's future scientists are likely to be drawn. Below, several possible socialization or environmental hypotheses for these differences will be described. These classical hypotheses were developed and have received some support for explaining sex differences among children of average ability. Yet we do not know whether they apply to intellectually talented students. Thus, the usefulness of these hypotheses for explaining sex differences in the SMPY population will be evaluated in a separate section. This review, however, is not intended to be exhaustive or to provide in-depth critical analysis that would consider factors such as experimental design.

9.1. Females have a lower liking for or more negative attitudes toward mathematics than do males (Brush 1980; Carey 1958; Dutton 1956; Fennema & Sherman 1977; Fox 1976; Hilton & Berglund 1974; Husen 1967; Keeves 1973; Tobias 1978). In addition, more women than men suffer from math anxiety (Brush 1978; 1980; Dutton 1956; Pedro et al. 1981; Tobias 1978; Tobias & Weissbrod 1980). Such attitudes have been shown to correlate with performance and confidence in, perceived usefulness of, and intention to take further courses in mathematics (Aiken 1972; Aiken & Dreger 1961; Brush 1978; 1980; Carey 1958; Dreger & Aiken 1957; Dutton 1956; Fennema 1976; Fennema & Sherman 1977; Fox 1982; Haven 1971; Hungerman 1967; Paulsen & Johnson 1983; Pedro et al. 1981; Sherman 1980; Sherman & Fennema 1977).

Two studies cite evidence, however, that females have more positive attitudes toward mathematics than their male counterparts (Paulsen & Johnson 1983; Stright 1960), and others (Besag 1987; Brush 1978; Dreger & Aiken 1957) find no sex difference in math anxiety. Several other studies have noted that male and female students from the second grade through high school report equal liking of mathematics (Battle 1966; Dutton & Blum 1968; Ernest 1976; Hungerman 1967; Parsons 1983). Moreover, if attitudes toward mathematics have an influence on mathematical achievement, it is stronger for men than women (Ethington & Wolfle 1986).

The general evidence seems to indicate that females do have somewhat more negative attitudes toward mathematics than do males. Furthermore, it is probable that these attitudes are correlated with variables related to mathematical performance. Among the high-ability population studied by Benbow and Stanley (1980; 1983b), however, no sex differences in attitudes toward mathematics have been found; nor are attitudes correlated with either concurrent or subsequent mathematical achievement, as will be described below.

9.2. As early as the seventh and eighth grades, females feel that mathematics is less useful or important to future career goals than do males (Brush 1980; Dornbusch 1974; Fennema & Sherman 1977; Fox 1975; Hilton & Berglund 1974; Parsons 1983; Sherman 1980). The perceived usefulness of mathematics has been cited as related to liking for and performance in mathematics, intentions to take future mathematics courses, perception of parental evaluations of mathematical ability, and self-concepts about mathematical ability (Brush 1980; Fennema & Sherman 1977; Fox 1977; Haven 1971; Parsons 1983; Pedro et al. 1981). Yet Brush (1980) found no relation between perception of future value of mathematics and intention to take future mathematics courses. Again, no sex differences or predictive value has been found in the SMPY population for this variable.

9.3. Females have less confidence in their mathematical ability than do males (Brush 1980; Ernest 1976; Fennema & Sherman 1977; 1978; Fox 1977; 1982; Parsons 1983; Parsons, Kaczala & Meece 1982; Pedro et al. 1981; Robitaille 1977; Sherman 1980), even though mathematical achievement in the classroom is consistently higher among females than males. Although sex differences do not exist in elementary school, by the early adolescent years, even among high-ability and highly motivated students, boys have greater confidence in their own overall mathematical ability and expected future performance in mathematics classes (Fox 1982; Heller & Parsons 1981; Mura 1987; Parsons 1983; Stein & Smithells 1969).

Confidence has been shown to correlate negatively with math anxiety and positively with mathematical achievement, value placed on mathematics, and intention to take optional mathematics courses (Fennema & Sherman 1977; 1978; Parsons 1983; Pedro et al. 1981; Sherman 1980).

The relationship of confidence to performance is strong and consistent for males, yet neither as consistent nor as strong for females (Parsons, Kaczala & Meece 1982). Females are more likely than males to cite lack of ability as the explanation for why they received a poor grade in mathematics; they are also more likely to cite lack of ability to explain a poor grade than they are to cite superior ability to explain a good grade (Dornbusch 1974; Parsons 1983; Pedro et al. 1981). Results have not been entirely consistent, however. Stein (1971) found no sex difference in expectancies for mathematical achievement. Another study found no relation between expectancies and intention to take further mathematics courses (Parsons 1983).

Overall, research indicates that females have less confidence in their mathematical ability than do males and that this lower self-concept may relate to females' lower levels of participation and performance in mathematics. Moreover, two studies in the SMPY population have revealed less self-confidence among SMPY females than males (Fox et al. 1982; 1985), even though no sex differences are found in this population in mathematics course-taking or in the intention to major in mathematics in college.

9.4. Females and males stereotype mathematics as a masculine discipline, thereby making females less motivated to achieve in mathematics (Dwyer 1974; Fennema & Sherman 1977; Maccoby 1966; Mokros & Koff 1978; Sherman 1980; Stein 1971; Stein & Smithells 1969). Although sex-typing of mathematics is often stronger among males than females, there could be an indirect effect of sex-typing as females strive to be accepted by their male peers (Fennema & Sherman 1977; 1978; Fox 1982; Fox et al. 1979; Sherman 1980; Sherman & Fennema 1977).

Sex-typing has been found to relate to mathematics scores and intention to take further mathematics courses (Dwyer 1974; Fennema & Sherman 1978; Paulsen & Johnson 1983; Sherman 1980; Sherman & Fennema

1977). For females, perception of mathematics as a male domain is linked to lower confidence in mathematical ability and to lower mathematical performance (Fennema & Sherman 1977; Sherman 1980).

Moreover, Ernest (1976) found that after the sixth grade children were more likely to receive help with homework in mathematics and science from their fathers than from their mothers. This is even more remarkable, because in general more help with schoolwork was received from mothers than fathers. Children also reported that their fathers had more positive attitudes toward them as learners of mathematics than did their mothers (Fennema & Sherman 1978). Fox (1977; 1982) noted that most advanced mathematics courses are taught by men and that there is a scarcity of female role models in mathematics and the sciences.

One study (Fennema & Sherman 1978) found sex differences in student perceptions of their mothers' and fathers' attitudes toward them as mathematics learners only in groups that also exhibited sex differences favoring males in mathematics test scores. In groups demonstrating equal mathematical ability by sex, no sex difference was found for this variable or for the perceived usefulness of mathematics.

Not congruent with this hypothesis is one researcher citing mathematics as the least clearly sex-typed subject area (Stein 1971). In addition, Brush (1980) found that mathematics was not significantly sex-typed and that sex-typing had no relation to further mathematical study; another investigation found a significant relationship only for males between sex-typing mathematics and plans for further study (Pedro et al. 1981). Boswell (1985) did not find consistent support for a relationship between sex-role stereotyping and women's participation in mathematics.

It appears then that if mathematics is sex-typed, directly or indirectly, it is considered masculine. The relationship of sex-typing to mathematical performance variables is unclear, however. Moreover, in a study conducted on the SMPY population, this explanation for sex differences was not given support (Raymond & Benbow 1986).

9.5. Differential expectations and encouragement. Significant others, such as parents and teachers, have different expectations for male and female mathematical achievement and encourage males more than females to achieve in mathematics (Fox 1977; Fox & Richmond 1979; Parsons, Adler & Kaczala 1982; Tobias 1978). Expectations and encouragement have been found to relate to children's expectations for performance in current and future mathematics courses, the perceived value of and attitudes toward mathematics, and enrollment in higher-level mathematics (Fennema & Sherman 1977; Haven 1971; Parsons 1983). In a collection of studies appearing in one book (i.e., Chipman et al. 1985), the most consistent finding was the important role of parents in influencing their daughters' participation in mathematics. Furthermore, Pedersen et al. (1986) reported that parental attitudes and student career interests were significant predictors of mathematical achievement over and above variables commonly used for prediction. It is interesting that sex could not account for any variance in mathematical achievement in that study. The above relationships between parental attitudes and performance are found even though students' self-concepts affect their perception of their parents' and teachers' estimates of their ability rather than these perceptions' affecting their self-concepts (Parsons 1983).

Another study found that, in general, the school performance of females was better than that of males, and parents of girls did not rate their daughters' mathematical ability lower than did parents of boys. Girls' parents, however, did feel that their daughters had to work harder in mathematics and that it was more difficult for them than for boys (Parsons, Adler & Kaczala 1982). Paternal responses were more sex-differentiated than were maternal responses. Furthermore, fathers of sons rated advanced mathematics as important, whereas fathers of daughters felt that English and American History were more important. Although fathers may treat their daughters and sons differently with respect to encouragement and expectations, mothers have been found to behave similarly toward children of both sexes (Johnson 1963).

Finally, children perceive their parents' and teachers' evaluations directly as well as indirectly: Females felt that their parents rated their mathematical ability lower than males felt their parents did (Fennema & Sherman 1977; Fox 1975; Parsons 1983). Sixth-grade females felt that teachers preferred good reading students, whereas their male counterparts stated that teachers liked good mathematics students better (Mokros & Koff 1978).

Again the results are not entirely consistent. Two studies found that females and males reported equal parental evaluation of themselves as mathematics learners (Parsons 1983; Sherman 1980), and one study reported that female mathematicians recalled greater past encouragement and greater past discouragement than did male mathematicians (Luchins & Luchins 1980). Another study found only a minimal sex difference in student reports of parental evaluations, and no relation between this and future mathematics course-taking (Brush 1980). Aiken and Dreger (1961) reported no relation between parental encouragement and the child's attitudes toward mathematics, and in Fox et al.'s (1982) sample, parents of both males and females reported encouraging their gifted child's mathematical self-confidence and enjoyment. In addition, a survey of entering college freshmen found no reported differences in mothers' versus fathers' expectations for and encouragement of mathematical achievement, although differences in reports by male and female students were not investigated (Poffenberger & Norton 1959).

Nevertheless, the general indication is that there are some differences in encouragement from significant others, especially parents, for males and females in mathematics, although the magnitude of these differences and their effect on the children's attitudes or achievement may not be great. Among the high-ability SMPY population, however, no such differences or effects have been found.

9.6. Differential course-taking by boys and girls in mathematics. Fennema and Sherman (1977) postulated that sex differences in mathematical aptitude occur because boys take more mathematics courses in high school than do girls. Pallas and Alexander (1983) reported finding some support for this hypothesis (but see Benbow & Stanley 1983a), as did Jones (1987). Armstrong (1985) and Wise (1985), however, presented data revealing that sex differences in mathematics participation in high school are very small. Ethington and Wolfle (1984) found that mathematical background could not account for sex differences in mathematical achievement test scores. Moreover, Boli et al. (1985) found no differences in undergraduate mathematics course participation between high-ability males and high-ability females. It seems then that the difference in mathematics course participation in high school has become almost nonexistent in recent years. Yet sex differences in mathematical aptitude have remained. Obviously, course participation does not control for other variables occurring in the classroom. Peterson and Fennema (1985) attempted to study classroom factors and found that student engagement and nonengagement in mathematical activities are related to students' mathematical achievement. Yet the global variable of student engagement/nonengagement in mathematics did not explain sex-related differences in mathematical achievement (Peterson & Fennema 1985).

Thus, the preponderance of evidence seems to indicate that the differential course-taking hypothesis is not a valid explanation for sex differences in mathematical aptitude.

9.7. Career and achievement motivation. Females are less intrinsically motivated than males and exhibit learned helplessness, both of which have negative implications for female achievement (Dweck et al. 1978; Nicholls 1980). Farmer (1987) reviewed the literature and presented a mutivariate model for explaining gender differences in career and achievement motivation. She concluded, as did Eccles et al. (1984), that the strength of this motivation for women is not unlike that for men. Yet "the pattern and type of factor influencing motivation for men and women differ significantly. The effect of parental and teacher support on motivation is stronger for women than men" (Farmer 1987, p. 5).

10. Environmental hypotheses and the SMPY population

The above hypotheses were formulated for and tested on average-ability populations. They may hold to varying extents for such groups, although causality cannot be truly demonstrated. Yet we do not know whether they can help explain sex differences among the intellectually talented. The validity of the first six of these environmental hypotheses has been evaluated for SMPY's high-ability population studied by Benbow and Stanley (1980; 1983b); the results of that work will be summarized below. Work is in progress to evaluate aspects of the last hypothesis.

10.1. Attitudes toward and usefulness of mathematics. No substantial differences have been found since at least 1976 in attitudes toward mathematics and the sciences of these high-ability preadolescent boys and girls (Benbow & Stanley 1982b). Fox et al. (1985) used the Fennema-Sherman Mathematics Attitudes Scales with a subgroup of talent search participants and found no differences between males and females. Differences are predicted by the socialization hypotheses. Furthermore, when these talent search students were studied five years later (i.e., after high school), few sex differences in attitudes toward mathematics were found. SMPY boys and girls reported roughly equal liking for mathematics, biology, chemistry, and science at that time. Although the differences between means were statistically significant due to the large sample size, they were not substantial (i.e., effect sizes were small or less), except perhaps toward physics (Benbow & Minor 1986; Benbow & Stanley 1982a; 1984). In the replication of those initial studies, unpublished preliminary results were consistent with that finding.

The absence of sex differences in attitudes toward mathematics is further exemplified in that slightly more girls than boys were planning to major specifically in the mathematical sciences in college and equal numbers actually did. Moreover, SMPY females received better grades in their high school mathematics classes than did SMPY males. Reported attitudes toward mathematics also had little relation to subsequent achievement in mathematics. For example, attitudes toward mathematics at approximately the ages of 13 and 18 could not predict the number of semesters of mathematics taken, the SAT-M score in high school, or the high school mathematics achievement test score (Benbow & Stanley 1982a).

Tobin (1986) studied scores on the Study of Values (SV) and career choices in the seventh grade and at the end of high school for a subgroup of 185 high-scoring participants in SMPY's 1976 Talent Search. It was concluded that the values of mathematically gifted youngsters, as measured by the SV, are not stable between the ages of 13 and 19 because the Aesthetic value increases over time and thereby forces corresponding decreases in all other values due to the ipsative nature of the SV. It was also found that boys' career choices are stable, whereas girls' career choices are much less so. Moreover, mathematically talented males tended to choose mathematically oriented careers in contrast to such females who chose careers that required minimum mathematical training. No prediction could be made, however, from values, aptitude, or sex as to career intentions six years later. Thus, measured values at age 12 seem unable to account for sex differences in mathematical career aspirations many years later or for the sex difference in aptitude. This study's methodological problems limit interpretability, however.

Relevant in this regard is the observation made by Benbow and Stanley (1982a) that high-aptitude girls may participate in mathematics less than high-aptitude boys, not because they like it less, but perhaps because they like the verbal areas much more than boys do. Arjmand et al. (in preparation) have found support for this hypothesis. Moreover, Ethington and Wolfle (1986) reported similar results. They found that attitudes toward mathematics were more negatively influenced by verbal abilities in women than in men.

Probably very few SMPY students suffer from math anxiety. A student with math anxiety would not enter a mathematics competition such as SMPY's except under duress. In addition, these students are in the top 3% of intellectual ability with a demonstrated aptitude for mathematics and above-average performance in their mathematics classes. Thus, the math anxiety hypothesis does not seem appropriate for this population. Moreover, no sex differences have been found in the perceived usefulness of mathematics among the SMPY students.

It seems then that the two environmental hypotheses pertaining to attitudes and perceived usefulness of mathematics have not been able to account for the sex differences in mathematical aptitude and achievement in the SMPY population. Essentially no differences in the critical variables were seen.

10.2. Females have less confidence than males. Fox et al. (1982; 1985) studied SMPY students' degree of self-confidence in mathematics. They found that the SMPY males were more self-confident than the SMPY females. This is difficult to interpret because there were no differences in course-taking in high school mathematics, slightly more females than males were planning to major in mathematics in college, and essentially equal numbers actually did. Nonetheless, the self-confidence hypothesis cannot be ruled out as a partial explanation of sex differences in the SMPY population.

10.3. Sex-typing and differential encouragement. Sex-typing and differential encouragement of boys and girls may help explain some sex differences. Fox et al. (1982) investigated the family backgrounds of SMPY talent search participants. They found few differences between the male and female participants. In particular, indications of more training or encouragement of boys were not detected.

The staff of SMPY has carried out a similar investigation. The relationship of sex differences in mathematics to differential patterns of parental involvement in their children's education in quantitative and verbal areas and to children's sex-typing was investigated both among young adolescents who were extremely talented mathematically or verbally (.01%) and among a comparison group of at least above-average ability students (Raymond & Benbow 1986). Patterns of parental support or encouragement in verbal or quantitative areas did not vary as a function of the child's gender. Fathers tended to be viewed as being more involved with their children's mathematical activities, and mothers with their verbal activities. These perceptions by the talented children themselves were weak, however. Moreover, fathers were not perceived to be more involved with mathematically talented children than with verbally talented ones, and the reverse was not found for mothers. In addition, the children did not appear to be strongly sex-typed and the extent of their sex-typing did not relate to perceived parental behaviors or to SAT scores (Raymond & Benbow 1986). The sex-typing measure used was not optimal, however.

The above study used the students' own responses about how they viewed their parents' behaviors in the various areas studied. The study was replicated using the parents' responses to the very same questions their children had answered (Raymond & Benbow, submitted). Essentially the same findings were obtained. We accordingly concluded that these aspects of socialization did not relate to current sex differences in mathematical reasoning ability, which have been shown to relate to later sex differences in achievement in quantitative areas.

Zimmerman (1984) completed a more thorough investigation of sex roles and mathematical achievement. A sample of high-scoring students in SMPY's 1976 Talent Search were given the Bem Sex Role Inventory in the seventh grade and again at the end of high school. These students had also completed SMPY's comprehensive after-high school questionnaire. Sex-role scores on the Bem were fairly stable over the 5- to 6-year period. Yet the test was unable to predict future mathematical course-taking and achievement in high school for mathematically talented students. The existence of a relationship between masculine identification and mathematical achievement was not supported.

Benbow (1986a) compared early toy preferences of extremely mathematically or verbally talented boys and girls and a comparison group (i.e., the same students that were studied by Raymond and Benbow [1986]) in an effort to obtain support for the hypothesis that because boys and girls play with different types of toys, they later develop sex differences in abilities, such as mathematical reasoning. No support was provided. Very few substantial differences in toy preferences were found among the subgroups of extremely talented boys and girls, and the pattern of results was not consistent.

Hence, there is so far no evidence that differential encouragement or sex-typing of males and females causes the sex difference in mathematical reasoning ability in the SMPY population.

10.4. Differential course-taking. In the very first article on sex differences published by Benbow and Stanley (1980), the differential course-taking hypothesis could not explain the sex differences in SAT-M scores among intellectually talented students. The initial sex difference on the SAT-M was found in the seventh grade before differential course-taking had begun, as is evident from the normal curriculum available to seventh graders and from the students' reports (Benbow & Stanley 1982b). Moreover, an equal percentage of SMPY boys and girls took mathematics in high school until the 12th grade, when the SATs are normally taken. In the first cohort to be studied longitudinally, SMPY boys did take about one more semester of high school mathematics than the girls did. This difference, however, was due to the larger number of males than females taking calculus. Because calculus was taken in the 12th grade after the high school SAT-M was taken, and because calculus items do not appear on the SAT-M, this is not a likely explanation of the sex difference in high school SAT-M scores. Furthermore, the best predictor of the high school SAT-M score was the talent search SAT-M, not the number of semesters of mathematics taken in high school, which accounted for little additional variance (Benbow 1981). Clearly, the differential course-taking hypothesis does not explain the ability difference found in this population in the seventh grade or at the end of high school. Conversely, however, the students who took calculus in high school had significantly higher *initial* mathematical and verbal reasoning abilities than students not electing to take this course (Benbow & Stanley 1982a).

Brody (1985) studied the effects on SAT-M scores of intensive course work in precalulus mathematics, the type of mathematics covered by the SAT. The subjects were seventh graders with high scores on the SAT who had enrolled during the summer before eighth grade in an intensive three-week course in either writing, languages, mathematics, or science. A control group consisted of students who met the selection criteria for the program but chose not to participate. At the end of the course, the SAT was again administered. No effects on SAT-M scores of intensive course work in precalculus mathematics were found. Moreover, the course work did not contribute differentially to the development of mathematical reasoning ability in males and females. These are further indications that formal course work in precalculus mathematics is unlikely to be a cause of the sex difference in mathematical reasoning ability.

Although sex differences in formal course work in mathematics may not relate to the sex difference in SAT-M scores, it may close out women from certain career options if they do not take enough mathematics in high school. To increase the participation of women in high school mathematics and especially calculus, an intervention program was implemented by SMPY in 1973, a time when sex differences in mathematics participation were much larger than now (Fox 1976). Moderately gifted seventh grade SMPY girls were invited to an accelerated mathematics program in algebra during the summer of 1973. The program, in addition to emphasizing algebra, catered to the social needs of girls, provided interaction with female role models who had careers in the mathematical sciences, and encouraged girls to study mathematics for a number of years. The girls who successfully completed the program (i.e., those who were placed in Algebra II the following fall) did take more advanced mathematics in high school and college (Fox et al. 1983). That, however, was the only major difference in academic achievement between this group of girls and an equally able group of girls not invited to attend the program. The other comparison group of equally able males achieved more academically than the girls attending the program No other positive effects were found for the program. Clearly, an early intervention strategy can improve the participation of girls in higher-level mathematics, but the girls must be successful in such a program. Moreover, exposure to role models and receiving encouragement were not sufficient in themselves to enhance female participation in mathematics later on.

We conclude that many years of research by the staff of SMPY and by others using the SMPY data base have not turned up results that provide support for the various socialization hypotheses. The reason for not finding any substantial differences in the socializing experiences of our high-aptitude boys and girls may be that it is not possible to detect subtle social influences affecting a child from birth on. Nevertheless, it is not entirely clear how differences in socialization experiences of boys and girls could affect the mathematical reasoning ability of girls so adversely and significantly, yet at the same time have no detectable effect on their reported attitudes toward mathematics, their taking of mathematics courses during the pre-SAT years, and their mathematics course grades. It is important to emphasize, however, that our results are limited to highly able students. The various socialization processes may have more impact on achievement and aptitude in mathematics in average-ability girls (Meece et al. 1982).

A possible reason why the various classical socialization hypotheses do not help explain the sex differences in mathematics among the intellectually talented is that those hypotheses were formulated to explain differences between means. SMPY's sex differences data may be the result of greater male variability.

11. Mathematical ability and spatial ability

Relationships between mathematical aptitude and spatial ability have been reported for many years (e.g., Harris 1978; Maccoby & Jacklin 1974; McGee 1979; Sherman 1967; 1977; Smith 1964). Since there is a well-documented sex difference in spatial ability favoring males (e.g., Maccoby & Jacklin 1974), it has been proposed that sex differences in spatial ability can account for the sex difference in mathematical aptitude (e.g., Sherman 1967). Support for this hypothesis has been mixed, however. Armstrong (1981) did not find that sex differences in mathematical achievement could be accounted for by sex differences in spatial ability. Ethington and Wolfle (1984) found that sex differences in mathematical ability remain even after controlling for differences in spatial ability. Fennema and Tartre (1985) found that students who differed in spatial ability did not differ in their ability to find correct problem solutions, although spatial ability did relate to the use of spatial processes in problem solving. Both Fennema and Tartre (1985) and Ethington and Wolfle (1984) did find, however, that low spatial ability may be more debilitating to mathematical problem solving in females.

There have been similar findings with the SMPY population and scores on the SAT-M. Becker (1978) found that the three-way interaction of spatial ability, sex, and item performance on the SAT-M was not significant for SMPY students in the seventh grade. Her conclusion was that among SMPY students there were no differences in performance from item to item on the SAT-M according to sex and spatial ability. Spatial ability was found to be related to superior performance on the SAT-M as a whole, however. Becker's results could have been confounded by the spatial ability test used, which had a large verbal component. Thus, the girls may have solved that test using a verbal strategy (Benbow 1978; McCall 1955; Sherman 1974). Furthermore, the SAT-M has few items with a spatial component, which may also have proved to be a problem in testing this hypothesis. Alternatively, mathematically precocious girls may require higher spatial ability than mathematically precocious boys in order to perform as well on the SAT-M (Cohn 1977). The latter explanation fits well with the findings of Ethington and Wolfle (1984) and Fennema and Tartre (1985).

On the other hand, Sherman (1977) cited two studies by Fennema and Sherman which found evidence that sex differences in mathematical ability could be attributed in part to sex differences in spatial ability. Furthermore, Burnett et al. (1979) found among a college sample that the sex difference on the SAT-M was no longer significant after controlling for spatial ability. Finally, McGee (1979) concluded that "sex differences in various aspects of perceptual-cognitive functioning (e.g., mathematics and field independence) are a secondary consequence of differences with respect to spatial visualization and spatial orientation abilities" (p. 909).

The above discrepancies among studies may be due to the fact that sex differences in both abilities are highly task-specific. Results would therefore vary depending upon the exact measures utilized. Thus, it cannot be ruled out that mathematical reasoning ability and spatial ability are somehow related. This may also be true for the SMPY population for a further reason. In an earlier study, the most precocious students that SMPY had identified were tested with a battery of specific mental ability measures (Benbow et al. 1983). Two factors were found able to account for these students' high performance: a verbal and a spatial factor. This implicates spatial ability in the high-level test performance of these mathematically talented students. Moreover, Benbow and Benbow (1984) argued that although mathematical reasoning ability may not be directly related, these two mental abilities may rely on similar, as yet undetermined, cognitive processes that might be mediated by the right hemisphere of the brain. If the association between spatial and mathematical reasoning ability does indeed exist, the tremendous number of research findings available for the sex difference in spatial ability may be relevant in understanding the sex difference in mathematical reasoning ability as well (Benbow & Benbow 1984).

12. Physiological correlates of mathematical reasoning ability

Because the staff of SMPY and others have been unable to find support for a primarily environmental explanation of the sex difference in mathematical reasoning ability in SMPY's high-ability population, and because of the potential association between mathematical reasoning and spatial ability (and hence the latter's possible biological basis), we began to investigate possible biological factors. Not much research on the biological correlates of extremely high mathematical reasoning ability has been conducted, but three physiological correlates of that ability have been identified to date. These are left-handedness, symptomatic atopic disease (allergies), and myopia (Benbow 1986b). The first two may be related to bihemispheric representation of cognitive functions or to the influence of fetal testosterone. If so, these may be two additional physiological correlates of that ability. We propose that all the correlates may be relevant to understanding the sex difference in mathematical reasoning ability.

12.1. Left-handedness and laterality. No overall differences in general intellectual functioning between left- and right-handers have been found (Hardyck et al. 1976). This does not imply, however, that differences in specific abilities cannot exist. Until recently it was assumed that right-handedness was advantageous (Harris 1980). New evidence, however, indicates that left-handers may be superior on those tasks mediated by the right hemisphere of the brain, such as spatial and musical performance (Burnett et al. 1982; Deutsch 1980). Consistent with these data, higher frequencies of left-handedness have been found among university mathematics teachers and students, music students, artists, astronauts, and architects (Annett & Kilshaw 1982; Deutsch 1980; Mebert & Michel 1980). Moreover, dyslexics, who are often left-handed, have superior talents in certain areas of nonverbal skill, such as art, architecture, engineering, and athletics (Geschwind 1982).

The right hemisphere is traditionally considered specialized for nonverbal tasks and the left for verbal, although these differences may not be qualitative but quantitative. Mathematical reasoning ability, especially in contrast to computational ability, may be more strongly under the influence of the right hemisphere (Benbow 1988; Gardner 1983; Troup et al. 1983; Warrington 1982). Thus, it would seem reasonable to predict that left-handers have an advantage on such tasks, and we began to investigate this hypothesis among our most talented mathematical reasoners. This occurred partly in response to an article published by Geschwind and Behan (1982), which is described below.

The Edinburgh Handedness Inventory (Oldfield 1971) was administered to 340 young adolescents who had scored extremely high (at least in the top 1 in 10,000) on the SAT-M or the SAT-V to their parents, and to a comparison group ($N = 201$) of above-average ability students (about the top 5% in ability). Self-reported handedness was obtained from the siblings of the index cases. In comparison to population norms for this inventory (i.e., 7.2%; Geschwind & Behan 1982), the extremely precocious students were more than twice as likely to be left-handed (Benbow 1986b). Moreover, they were more frequently left-handed than their parents, siblings, and the comparison group (i.e., 15.1% for the extremely precocious vs. around 10% for the others). Indeed, the comparison group and the parents (42% and 53%, respectively) were about twice as likely as the extremely precocious (23%) to report using their right hand to perform all of the ten tasks on the handedness inventory (Benbow 1986b). Finally, there was a sex difference in handedness among the extremely precocious students. More males than females (16.4% vs. 11.4%; $p < .05$) were left-handed (Benbow 1986b). Yet the incidence of left-handedness was somewhat elevated among the extremely talented females too.

A counterintuitive result in the Benbow (1986b) study was that students with high verbal reasoning ability also exhibited an increased frequency of left-handedness. Indeed the males who were extremely talented verbally exhibited the highest frequency of left-handedness (23.5%). This may have occurred because left-handers comprise a very mixed group in terms of cerebral dominance, or because of a possible interaction among handedness, sex, and reasoning ability (Harshman et al. 1983). I should like to offer another possible explanation for the seemingly anomalous result, however.

The verbal ability measured has been described as verbal *reasoning*. This verbal ability may have a considerable right-hemisphere contribution in addition to the traditional left-hemisphere involvement. Verbal reasoning ability involves comprehension, analogical reasoning, and the understanding of difficult words, which are just those types of cognitive processes posited to describe the verbal contributions of the right hemisphere (e.g., Caramazza et al. 1976; Gardner et al. 1983; Read 1981; Villardita 1987). Zaidel (1978), for example, points out that the linguistic structure of the right hemisphere appears to hinge on pronounced semantic competences that enable it to have a rich lexicon. (For a review of the language functions of the right hemisphere, see Millar and Whitaker, 1983.) Thus, we propose that verbal reasoning may be more under the influence of right-hemisphere processing than language production or syntactical aspects of verbal ability. This reasoning is similar to that used to describe the differing contributions of the hemispheres in mathematics. The left hemisphere is involved in computation (e.g., Zolog 1983), while the right is involved in understanding numerical relations and concepts (Gardner 1983; Troup et al. 1983).

It is interesting to note that learning disabilities are also more frequently found among left-handers than right-handers (Geschwind & Behan 1982). Since there was not an especially increased frequency of left-handedness in the above-average ability comparison group used in Benbow (1986b), it may be that left-handedness is simply associated with extremes in ability at either end of the scale or with greater variability (Benbow 1987b).

The handedness findings may have some further implications. A variety of clinical evidence has suggested that left- and mixed-handers and right-handers with left-handed relatives (perhaps only among males) form a more heterogeneous group than right-handers in terms of their cerebral lateralization for language, and perhaps for spatial functions as well. These groups have been found more frequently to have bilateral or diffuse representation of cognitive functions (Bradshaw & Nettleton 1983; McKeever et al. 1983; O'Boyle & Hoff, in press). Thus, it may be that bilateral representation of cognitive functions (rather than greater specialization of the hemispheres) is associated with extremely high mathematical and verbal reasoning abilities. Burnett et al. (1982) proposed a similar hypothesis for spatial ability. A pilot study in which letter-matching or rotation tasks were presented with a tachistoscope to each visual field of extremely precocious students did yield data consistent with this hypothesis (Benbow & Benbow 1987). The response times for the letter-matching task (Posner et al. 1969) were somewhat higher when the problem was flashed to the left rather than the right visual field. Although the lack of a control group limits the interpretability of that experiment, Levy (personal communication, 1986) obtained similar data in support of the above hypothesis but had a control group. Thus, bilateral or diffuse representation of cognitive functions and/or a strong right hemisphere may be another possible biological correlate of extremely high mathematical reasoning ability.

McGlone (1980), as well as others, have argued that there are sex differences in human brain asymmetries, which may help explain sex differences in various cognitive abilities. The male brain may be more asymmetrically organized than the female brain, both for verbal and nonverbal functions, which suggests that sex differences in cerebral lateralization might underlie sex differences in spatial ability (Harris 1978; McGlone 1980). Lewis and Kamptner (1987) recently found support for this hypothesis and found evidence indicating that sex differences in functional lateralization may be present for only certain visuospatial processes (rotational and perhaps manipulospatial). Our data might be difficult to reconcile with that hypothesis, however.

In this connection, it is interesting to note that Witelson (1985) has found the corpus callosum, the main fiber tract connecting the two cerebral hemispheres, to be larger in left- and mixed-handers than in right-handers. Thus, the "greater bihemispheric representation of cognitive functions in left- and mixed-handers may be associated with greater anatomical connection between the hemispheres" (Witelson 1985). Moreover, de Lacoste-Utamsing and Holloway (1982) have reported differences in the corpus callosum according to sex, although Witelson (1985), with a larger sample, was unable to replicate the finding. Yet Witelson's data did suggest the possibility of a complex sex factor which, in conjunction with hand preference, may be related to the size of some part of the corpus callosum's posterior half. Thus, it is not clear whether differences in the size of the corpus callosum could have any relation to sex differences in ability.

12.2. Allergies. Because Geschwind and Behan (1982) had found that left-handers suffer much more frequently from immune disorders than right-handers, we too investigated the frequency of allergies among extremely precocious students. An allergy questionnaire, designed by Franklin Adkinson at the Johns Hopkins School of Medicine, was mailed to and completed by the most precocious students' parents. This questionnaire classified possible allergies of family members by frequency, severity, duration, and type. As found for left-handedness, students with extremely high mathematical or verbal reasoning ability were at least twice as likely to have allergies as members of the general population (Benbow 1986b). Moreover, such students were more often reported by their parents to have allergies (i.e., 53%; no significant sex differences) than the parents themselves (44%), their siblings (35%), or a comparison group of above-average ability students (35%). The most frequently reported allergy was hayfever. [See also Gualtieri & Hicks "An Immunoreactive Theory of Selective Male Affliction" *BBS* 8(3) 1985.]

12.3. Hormonal influences. Geschwind and Behan (1982) postulated that left-handedness and immune disorders, such as allergies, are related to exposure to high levels of testosterone in fetal life or to high fetal sensitivity to testosterone. They suggested that testosterone slows the development of the left hemisphere and thereby enhances the development of the right hemisphere, simultaneously affecting immune development through the thymus gland. Coren et al. (1986) found that left-handers tend to be somewhat delayed in achieving sexual maturation, which may be an indication that sinistrality is associated with hormonal differences. Thus, since left-handedness and allergies were characteristics exhibited by extremely high mathematical or verbal reasoners, prenatal exposure to high levels of testosterone may be another possible correlate of extreme levels of those abilities. Since testosterone is the male hormone, which is secreted by prenatal testes, this suggests a connection between it and the sex difference in mathematical reasoning ability.

This link seems potentially reasonable since male and female hormones (androgens and estrogens, respectively) have frequently been implicated in the production of sex-related differences in spatial ability. (Spatial ability is also considered to be a function more efficiently carried out by the right hemisphere [for a review, see Bradshaw & Nettleton 1983] and, as discussed above, to be possibly related to mathematical performance.) Moreover, androgens have permanent organizing effects on the structure of the brain (DeVries et al. 1984). Studies of the performance and intelligence of individuals exposed to abnormal levels of hormones during development have produced results consistent with the hypothesis that androgens before or at puberty are importantly related to the development of spatial ability, and other behaviors as well (Hier & Crowley 1982; Nyborg 1984; Reinisch 1984). Since high androgen levels are associated with low spatial scores among males and with high spatial scores among females (Broverman et al. 1964; Petersen 1976), there may be an optimal androgen-estrogen ratio for the development of high spatial ability. A related hypothesis is that the onset of physical maturation is associated with the development of cognitive abilities (Waber 1976). Waber et al. (1985), however, have recently failed to find support for this hypothesis.

It has been shown further that progesterone exposure enhances numerical ability (Reinisch et al. 1979). Levy and Gur (1980) proposed that high levels of fetal sex hormones promote the maturational rate and cognitive capacity of the right hemisphere. It is of interest that first-borns compared to later-born siblings are exposed to higher levels of hormones prenatally (Maccoby et al. 1979). Most of the extremely precocious students investigated by Benbow (1986b) were first-borns (Benbow & Benbow 1987).

Although prenatal testosterone could specifically enhance the development of the right hemisphere of the brain, it may also work to make individuals more variable (Benbow 1987b). This has been suggested for physical maturation (Waber et al. 1985). If it can be established that an association exists between prenatal testosterone exposure and left-handedness, intellectual talent, learning disabilities, male variability, and allergies, then this would support such a hypothesis. The testosterone/ variability connection may also be relevant for a better understanding of the sex difference in SAT scores in the SMPY population.

12.4. Myopia. Because myopia has frequently been correlated with higher general intelligence (Karlsson 1973; 1975; Sofaer & Emery 1981), we also investigated the frequency of this trait among extremely high-level mathematical and verbal reasoners. Such high-ability students were about four times as likely to be myopic as were average high school students in the United States and were again much more frequently myopic than their siblings and a comparison group (Benbow 1986b). Verbally precocious students compared to mathematically precocious students and females compared to males in our sample were significantly more likely to be myopic. Cohn et al. (in press) studied measured refractive error of extremely gifted students and their siblings. Their data also revealed a relationship between myopia and extremely high ability.

Although a hereditary component, not related to the left-handedness and allergy findings, has been implicated in causing myopia (Ashton 1983; Bartsocas & Kastrantas 1981; Basu & Jindal 1983; Dunphy 1970; Karlsson 1973; 1975), the basis for the relationship of myopia to extremely high mathematical and verbal reasoning ability, as well as high general intelligence, is obscure. Extreme environmental factors have been shown to produce myopia (Green 1980). Yet Cohn et al. (in press) have ruled out reading or other types of "nearwork" as the mediating factor. Although hormonal explanations for the emergence of myopia have also been offered, they have failed to receive wide scientific support. Finally, Schachter et al. (1987) just recently presented some intriguing speculation regarding the possible role of melanin in the proper development of central nervous system (CNS) pathways involving vision. Moreover, they proposed that "[i]f a similar function is subserved by melanin in other CNS locations, hypopigmented humans (i.e. those with blond hair, blue eyes, and fair skin) might be more susceptible to those intrauterine influences that slow neuronal and/or axonal migration, e.g. testosterone" (p. 274). Nonetheless, the myopia result should perhaps be viewed as a heuristic finding.

In sum, the above physiological correlates, especially the possibility of prenatal testosterone exposure, lend credence to the view that sex differences in extremely high mathematical reasoning ability may be, in part, physiologically determined (Benbow & Stanley 1980). Of course, some of the above discussion on physiological correlates is speculative.

13. Summary

In conclusion, it is clear after the testing of several hundred thousand intellectually talented 12- to 13-year-old students nationwide over a 15-year period that there are consistent sex differences favoring males in mathematical reasoning ability (or more specifically in SAT-M scores). These differences are pronounced at the highest levels of that ability. They are found in other countries as well, even in countries where the cultures are radically different. The sex difference in mathematical reasoning ability can predict subsequent sex differences in achievement in mathematics and the sciences. Thus, it has practical importance. Further long-term implications of the sex difference is being investigated longitudinally by SMPY at Iowa State University.

Several environmental hypotheses about possible causes for this sex difference in aptitude have been explored. To date a primarily environmental explanation for the difference has not received support from the numerous studies conducted over many years by the staff of SMPY and by others. This and the identification of several physiological correlates of extremely high mathematical reasoning ability lend credence to the view that these sex differences may partly be biologically induced. Because there are well-documented differences in the socialization as well as in the biology of boys and girls, it is proposed that a combination of both of these factors causes the sex difference in mathematical reasoning ability. The relative degree of each, however, cannot be measured with precision.

Even though biological factors seem to be involved in determining the sex difference in mathematical reasoning ability, this does not imply that efforts at remediation cannot make a difference. They probably can and ought to be tried. Thus, practically speaking, one must be an environmentalist. Yet, in order to make remediation effective, one must know the extent of the sex difference.

It would not and does not help females if differences are swept under the rug. The exact nature of the problem needs to be determined in order to develop appropriate solutions. The intent of SMPY's research on sex differences was to address this issue. We believe that our investigation has provided some answers, while at the same time raising more questions.

ACKNOWLEDGMENTS

I should like to thank Linda Brody, Veronica Dark, Meg Gerrard, Michael O'Boyle, Ronald Peters, and Julian Stanley for helpful comments and suggestions and Lois S. Sandhofer for assistance in gathering the references. Support was provided by the Department of Education and the National Science Foundation (MDR-8470387).

In this paper I review much of my own work, and, thus, some parts are adapted from previously published writings, primarily Benbow (1987b) and Raymond and Benbow (1986).

Open Peer Commentary

Commentaries submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

The effects of selection and variability in studies of gender differences

Betsy Jane Becker[a] and Larry V. Hedges[b]

[a]College of Education, Michigan State University, East Lansing, Mich. 48824 and [b]Department of Education, University of Chicago, Chicago, Ill. 60637

Benbow's examination of data on gender differences among mathematically talented youth is interesting and comprehensive. However, as Benbow has noted, the implications of these data for the general population and for all talented youth are far from clear, primarily because they are based on a highly selected population.

The effects of selecting subjects from the extreme tail of a score distribution can be explored with the aid of statistical theory, and some simple calculations show how difficult it can be to draw inferences about the unrestricted population. In particular, it is difficult to infer most characteristics of the general population (such as population means and variances) from extremely selected samples because many different population structures can give rise to rather similar data in the tails.

To illustrate this point, we pose a series of plausible population models that could produce data like those presented in Benbow's Table 1. Suppose that the SMPY sample had been selected by taking every student who scored above a cutoff score on a test like the Scholastic Aptitude Test – mathematics section (SAT-M). (This is equivalent to assuming that the in-grade tests actually used by SMPY for talent-search screening are linearly related to that test.) Assume also that the scores on this test for the whole population were normally distributed, and that the test had sufficient range on the low end to adequately accommodate the full range of scores from an unrestricted population of seventh graders. For the purpose of illustration, we scale this test to produce scores in a range similar to that of the SAT-M, and we select 400 as the cutoff score. These assumptions imply that the SMPY girls' and boys' scores have truncated normal distributions.

We make two further assumptions for our investigation. Note that in Benbow's Table 1 the girls' SAT-M scores have smaller variances than the boys' in every case. Consequently, we assume that the variance of the unrestricted population of girls' scores is somewhat smaller than that of the unrestricted population of boys' scores. Though Benbow noted the greater variability of the males' SAT-M scores several times, she has not focused on it in her search for explanations of the gender differences in the SMPY data.

Finally, since SMPY subjects elected to participate in the talent searches, the SMPY selection process also involved a certain amount of *self*-selection. Self-selection is problematic because it is difficult to characterize and it is not modeled in our computations. Detailed information about the selection process, such as characteristics of the students who were nominally qualified but elected not to participate, is not available for the SMPY talent searches. In the absence of information about self-selection, we can do little more than suggest that it may prove a plausible rival explanation of some of the SMPY findings.

Setting the boys' standard deviation in the unrestricted population at 200 and the girls' standard deviation at 150, we used standard results for the truncated normal distribution (Johnson & Kotz 1970, pp. 81–82, equations 79 and 80) to compute the means (expected values) and variances of the selected scores

Table 1 (Becker and Hedges). *Means, standard deviations, and sex ratios of male and female scores above 400 as a function of the mean and standard deviation of the unselected population*

	Unselected population		Selected population		Sex ratio[a] (M/F) for scores			
Sex	Mean	Standard deviation	Mean	Standard deviation	>400	>500	>600	>700
M	250	200	516	96	—	—	—	—
F	200	150	470	61	2.48	4.64	10.46	28.49
F	225	150	474	64	1.86	3.16	6.45	15.86
F	250	150	479	67	1.43	2.21	4.08	9.06
F	275	150	484	70	1.12	1.58	2.65	5.31
F	300	150	485	71	0.90	1.16	1.76	3.19

[a]The ratio is the proportion of males scoring above each given level, divided by the analogous proportion of females.

(those above the cutoff of 400). The table shows the results of such an analysis when the mean of the boys' scores in the unrestricted population is fixed at 250 (shown on line 1) and the girls' mean in the unrestricted population varies from 200 to 300.

Note that the means, variances, and sex ratios for proportions of subjects exceeding selected scores are well within the range of values in Benbow's Table 1. This table illustrates the remarkable fact that even when the girls' mean in the unrestricted population exceeds that of the boys by a substantial amount (50 points), the boys' mean is larger in the selected population and the sex ratios among high-scoring students favor males.

Although Benbow does not seriously consider differences in variability between males and females as an explanation of her results, it is possible to account for the sex differences she has reported on the SAT-M among talented youth as a consequence of selection and the fact that the males' SAT-M scores are more variable. While it is true that differences in variability also require an explanation, the nature of such an explanation might differ from that primarily oriented toward explaining a mean difference. For example, an explanation for differences in variability might focus on variances rather than means of environmental variables that affect the sexes. These findings suggest many possibilities for further inquiry, and we hope that Benbow, and others, will investigate them fully.

The plasticity of the human brain and human potential

Ruth Bleier
Department of Neurophysiology, University of Wisconsin, Madison, Wis. 53706

Benbow concludes her target article by proposing that a combination of socialization and biological factors causes the sex differences she finds in mathematical reasoning ability. Because her published articles and public statements convey her commitment to biological ("endogenous") explanations and because, in this paper, she fails to find convincing sociological evidence while accepting as valid a number of tenuous biological hypotheses and inconclusive studies, I direct my critical comments to some of these biological conjectures.

One of the biological factors Benbow proposes as the basis for sex differences in high SAT math scorers is that of sex differences in hemispheric lateralization of cognitive functions, especially spatial functions. Yet leading workers in the field (e.g., Hugh Fairweather 1976; Marcel Kinsbourne 1974, among others) do not believe that sex differences in hemispheric lateralization exist or have been demonstrated, as was also indicated by a majority of the commentaries for the frequently cited review article that appeared in this journal (McGlone 1980), which was also cited by Benbow. In his recent critical review of the five major paradigms used to investigate cerebral lateralization in children, Hahn (1987) concluded, "Some studies found the male brain to be more asymmetrically organized than the female brain, whereas other studies found the female brain to be more asymmetrically organized than the male brain. But in most cases, the data showed that sex differences do not exist" (p. 389).

Benbow suggests that bilateral or diffuse hemispheric representation of cognitive functions, as has been found in some studies of left-handers, may be associated with extremely high mathematical abilities and thus may account for the predominance of males in her population. Although she states that she has identified left-handedness as one of the physiological correlates of extremely high mathematical reasoning, in fact only 15% of her group of males is left-handed. That still leaves 85% of the talented group who are right-handed. Right-handedness, not left, is correlated with high math ability.

Benbow states in her abstract that prenatal hormonal exposure, in particular to testosterone, is another physiological correlate of extremely high math reasoning ability. Yet no study exists showing such a correlation. One clinical study cited by Benbow claimed to find low spatial ability in androgen-deficient men (Hier & Crowley 1982). For this study to be relevant to mathematical superiority in males, one must then assume that over-androgenized males would have superior spatial ability and, for the next step, that there is a known, rather than conjectured, correlation between spatial ability and mathematical reasoning ability. It is perhaps a further indication of the opaqueness of this area of theorizing that another study cited by Benbow found high androgen levels to be associated with low spatial scores among males and with high spatial scores among females (Broverman et al. 1964). Benbow offers as further supporting evidence for the testosterone hypothesis Geschwind's (Geschwind & Behan 1982) *speculation* that fetal testosterone may inhibit the development of the left hemisphere (resulting in right-hemispheric dominance) and thus account for the statistical association he found among left-handedness, dyslexia, migraine, and asthma in boys. [See also Gualtieri & Hicks: "An Immunoreactive Theory of Selective Male Affliction" *BBS* 8(3) 1985.] But the majority of his subjects were not left-handed, nor did he study mathematical ability, though he did go on to claim that his theory could account for "superior right-hemispheric talents, such as artistic, musical, or mathematical talent" (Kolata 1983). But there are more serious flaws in Goschwind's theory: (1) there is no such known effect of testosterone on the developing brain; (2) there is no explanation for how circulating testosterone would affect only one hemisphere; (3) in support of his conjecture Geschwind cited a study by Chi et al. (1977) that found two convolutions of the human fetal brain appearing one to two weeks earlier on the right side than on the left; yet Geschwind failed to mention that Chi et al. found *no* sex differences in this temporary asymmetry in the 507 fetal brains they examined; (4) finally, a recent critical article (Satz & Soper 1986) questions, on methodological and statistical grounds, the associations claimed in the Geschwind study. Even if this theory of testosterone inhibition of the left hemisphere were valid, it is not evident how it lends support to Benbow's belief concerning either bilaterality or right dominance of hemispheric function in math superiority, especially in view of the fact that 85% of her subjects are right-handed.

Finally, a recent review of a variety of studies comparing mathematical abilities of students in the United States with those in China and Japan (Steen 1987) emphasizes the vast superiority of Chinese and Japanese students from kindergarten through twelfth grade. Are we to postulate an Asian gene for math ability or an Asian hormonal profile or an Asian pattern of hemispheric lateralization? Asians would not. The same review points out that "Americans more than any other people attribute success in mathematics to innate ability rather than to hard work. Students, parents, and teachers the world over, except in the United States, believe that everyone can learn mathematics if only they work hard at it" (p. 302). But if the vast majority of Japanese and Chinese students have a consistently different experience of and attitude toward the learning and use of mathematics from earliest childhood, as the most mathematically talented in this country undoubtedly do, it is possible that their brains *are* somehow differently organized structurally and/or functionally than those of mathematically average students *because* of such early developmental experiences.

It is clear from both experimental work in animals and clinical findings in humans that the developing brain is enormously sensitive to environmental influences and that functional interactions with the external environment are critical for establishing normal synaptic connections (Goldman & Rakic 1979; Trune 1982; Webster & Webster 1977; Wiesel 1982). [See also Goldman-Rakic's accompanying commentary, this issue.] That is, brain structure (biology) itself reflects and incorporates the environmental, cultural, and learning influences to which it is exposed before and after birth. When every developing child is necessarily exposed to a different range of environmental influences and to vastly different learning, psychic, and social experiences that affect neuronal structure and function as well as cognitive and psychological characteristics, it seems evident that belief in group (gender, race) differences in cognitive abilities based on group biological differences is not supportable either by neurobiological principles or by social scientific experience. Despite a century-long effort by a number of eminent scientists to find sex differences in brain structure, solid evidence for biological constraints on any group's intellectual potential (today it is women; in the 1960s it was blacks) has yet to be discovered.

Boys and girls and mathematics: What is the difference?

Lois Bloom

Department of Developmental and Educational Psychology, Teachers College, Columbia University, New York, N.Y. 10027

The long and careful research program that is reviewed here has resulted in the heavily documented finding that among academically talented children on the threshold of the most important educational choices of their lives, females do less well than males on tests of mathematical ability. My comments on this set of findings fall into two categories: Feminist and scientist.

Do the sex differences in performance on the SAT-M matter? On the one hand, maybe not. The young adolescent girls who score less well than boys are part of a history of social change in which the balance between the sexes in many areas of life is shifting. The data cited from the follow-up on the most select group of children in the most recent talent search suggest such change. Although no change has been observed in the discrepancy between boys' and girls' SAT-M scores in the last 15 years, in the balance of social history 15 years may just not be long enough. On the other hand, the answer to the question "Does it matter (that boys score better than girls on the SAT-M)?" is *yes*, because the differences reviewed here are going to do two things. For one, they are going to make it even harder for women to open the doors to advanced courses and graduate programs in mathematics. But in addition, and more insidiously, such data could work as self-fulfilling prophecy. Young women who might otherwise have followed their inclination and pursued an interest in math (and Benbow reports that women actually perform better in math courses than men, SAT-M scores notwithstanding) could be discouraged by the sort of evidence assembled here. The problem is in the effects that this apparent confirmation of the long tradition of sex differences in math ability could have, for both the women who might now be even less inclined to fight the odds, and society which can only benefit from women's participation in science and industry.

What do the research findings mean? The research presented here raises issues that transcend sex differences in math ability. Three of these are the issues of product versus process, the relation between correlation and cause, and the relation between verbal and mathematical reasoning. First, all of the research reviewed here has been directed at the product of performance on standardized psychometric tests. The claim is that the SAT-M "taps" mathematical reasoning, and the conclusion is that females are therefore less adept than males at mathematical reasoning because they score less well on the test. However, no research has looked at sex differences in the *process* of mathematical reasoning. Second, other variables in which sex differences favoring males have also been found appear to correlate with mathematical ability, leading to the conclusion that they might, therefore, explain the sex difference in SAT-M scores. However, because correlations can only tell us that two or more things may be associated, we cannot know the direction of influence between them. Moreover, so far as I can tell, several of these variables (left-handedness, for example) have not been independently tested or controlled for in relation to sex differences on the SAT-M with the children studied by Benbow and her colleagues. For those that have (spatial ability, for example), the lack of a significant relation is explained away by other factors. Third, the discussion of "physiological correlates of mathematical reasoning ability" which concludes Benbow's article seems to confuse if not confound verbal reasoning with mathematical ability. Research is reviewed that has identified neurological correlates of verbal reasoning (but not mathematical reasoning), and no sex differences are reported for these studies. The same neurological attributions (i.e., *both* right hemisphere and bilateral representation of cognitive function) are then made for both verbal and mathematical reasoning. The problem here is that no sex differences have been found for verbal reasoning in this population of children. The biological factors that are implicated to explain the differences in SAT-M scores, therefore, are moot.

In sum, Benbow has presented an interesting set of research findings and raised some important questions in the effort to understand them. But we are left at the end with the suggestion that "remediation [for the sex differences in the performance of academically gifted children on math tests] . . . can and ought to be tried." However, we do not yet have an adequate understanding or even a clear enough definition of the problem to begin to know whether we should do something about it, let alone what to do about it.

Sex differences in mathematics: Is there any news here?

Lila Ghent Braine

Psychology Department, Barnard College, Columbia University, New York, N.Y. 10027

Benbow claims that mathematically gifted girls, equated with boys for mathematical achievement, score lower than the boys on mathematical reasoning; this claim is suspect on several grounds. First, consider the selection of high achievers in the SMPY groups. The top 3% on mathematical achievement tests consistently yielded more boys than girls (57% to 43%, on average). To yield this difference in frequency, the girls' distribution must be shifted downward relative to the boys so that the top 3% for the students as a whole cuts the girls' distribution at a higher point on their curve (yielding fewer girls, and scores that do not extend as high as the boys' scores). It is not news that boys (after elementary school) score higher than girls on math tests (Chipman & Thomas 1985).

With respect to Benbow's data, the point is that the achievement scores of the mathematically gifted girls and boys were not matched. The girls' scores would be lower, and show less variability, than those of the comparably selected boys, as indeed they are on the SAT-M (target article, Table 1). Benbow argues that the lower scores on the SAT-M reflect poorer mathematical reasoning by the girls, whereas I would argue that the selection process inevitably yielded groups that were not comparable. (This conclusion is not affected by the Center for the Advancement of Academically Talented Youth (CTY) studies with equal numbers of boys and girls, because these samples included youth with either high verbal or high mathematical scores, and it is presumably the verbal scores that brought in the additional girls.)

Second, there are other ways of cutting the top 3% on math achievement. What would happen if, for this group, the SAT-M scores of blacks and whites were compared, or those of urban and rural youth? I suspect that black and rural youth would be underrepresented, and that their means and standard deviations would be lower, because these groups as a whole tend to score lower on achievement tests. In addition, I would suspect that these subgroups would show only small differences (as did boys and girls) on socialization and attitudinal measures (target article, Sections 9 & 10). Such control comparisons are necessary to support Benbow's claim that her findings are specifically related to sex, and are not found in other comparisons (especially those in which the subgroups differ in relative representation in the gifted group).

A third concern is with the claim that the SAT-M measures math ability rather than achievement. Without belaboring the point, that view is not shared by others (Chipman & Thomas 1985, p. 9), nor even by Educational Testing Service (Jackson 1980, p. 383).

Benbow says that socialization cannot account for the sex difference found. I suggest that the crux of the problem was not discussed by Benbow. The question that needs to be asked is, "Why are women virtually absent from fields for which they clearly have the requisite mathematical skills?" There is enormous overlap in the math scores of boys and girls, despite differences in the means. Many women have the math skills required for success in the natural sciences, engineering, economics, accountancy, and so on, yet few women appear in these "male" occupations. Social factors must operate to keep women out of certain occupations (e.g., Rossiter 1982). To understand cultural effects on math performance, we need to investigate those social factors keeping women (and other groups) in some occupations and out of others, rather than focussing on individual differences in socialization practices of parents and teachers.

Asking why there are no women among the highest scorers on the SAT-M is like asking why there are no great women artists, and feminist writers have pointed to social factors that differentially nurture male talent (Nochlin 1972; Woolf 1929). Again, let us move from the sex contrast to consider national differences in science; e.g., why are there so few great scientists from South America (in contrast to great writers)? I doubt that South Americans are lacking in aptitude. Rather, historic and economic conditions have not favored the development of the scientific enterprise, and talent has been attracted to more rewarding fields, such as literature. I suggest that it is less attractive for many mathematically talented adolescent females to develop their skills intensively in mathematics than in fields where women are visibly rewarded, and which are more compatible with other cultural values of female adolescents. Ask yourself how you would respond to a "math freak" adolescent son and a "math freak" adolescent daughter.

Benbow claims that boys have higher mathematics scores "even in countries where the cultures are radically different." However, in none of these cultures is math achievement valued more highly for girls than for boys, nor are there more women than men in mathematically related fields. The "different" cultures are not different in ways that provide evidence about math achievement.

The claim for physiological correlates of mathematical reasoning stems from epidemiological findings of an association between left-handedness, allergies, and myopia. However, we have no understanding of the basis for these associations at a physiological level or at a psychological level. Given our ignorance, Benbow's leaps from left-handedness to localization of reasoning in the right hemisphere to sex differences in math ability are pure speculation, at best. I suggest that extensive cognitive analysis of math ability is required before brain–behavior relations can be sensibly examined.

Does Benbow report any scientific news? It remains to be demonstrated that the sex difference in SAT-M scores at the upper ranges tells us anything we did not know, that is, that the girls' distribution on math achievement tests is displaced downward relative to that of the boys. We do need some news, for example, analyses involving contrasts other than sex, and some better understanding of how cultural factors operate to attract and nurture talent in particular groups. I certainly think that individual differences in math ability are a function of both biology and experience (as are all other skills), but Benbow provides no new evidence that sex differences in math scores have a biological component.

Cerebral organization and mathematical ability

M. P. Bryden

Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Having satisfied herself that there really is a sex difference in the incidence of high-level mathematical ability, and that this difference cannot be wholly determined by experiential differences between the sexes, Benbow has concluded that there must be a biological basis for sex differences in mathematical ability. She offers us a variety of speculations as to the nature of these biological differences based largely on the recent claim of associations between handedness, immune disorders, and cerebral specialization (Geschwind & Galaburda 1986). Her arguments loosely connect a long chain of associations to reach a conclusion that appears unjustified on the basis of our present knowledge.

Let us examine Benbow's claims more carefully. First, she assumes that high mathematical ability is associated with high spatial ability and therefore with right hemisphere function. While Benbow cites some evidence to support this claim, there are other large-scale studies suggesting that spatial ability has relatively little to do with mathematical ability (e.g., Pattison & Grieve 1984). Furthermore, neuropsychological studies indicate that acalculias (disorders of calculation) are most likely to involve lesions of the posterior left hemisphere (Boller & Grafman 1985). Given that visuospatial functions are more likely to be dependent on the right hemisphere, it is hard to justify the conclusion that mathematical and spatial abilities involve the same cerebral substrate. Indeed, Gardner (1983) has argued that visuospatial ability and mathematico-logical ability are separate "intelligences," dependent on different modules of cerebral functioning. Benbow solves this problem by recourse to a concept of "mathematical reasoning," as distinct from computational ability, but she leaves the link to spatial ability undetermined.

Second, Benbow assumes a relation between handedness and cerebral function. Investigators have been so impressed with the observation that left-handers are more likely to suffer aphasic disturbances following right-hemispheric damage that they often forget the details of this relationship. The majority of left-handers, like right-handers, have language lateralized to the left cerebral hemisphere (Segalowitz & Bryden 1983). Furthermore, handedness predicts the lateralization of other functions even more poorly. De Renzi (1982) has concluded that the lateralization of visuospatial functions is unrelated to handedness; and Bryden et al. (1983) have shown that language and visuospatial functions are independently lateralized. Thus, even if higher mathematical abilities were related to visuospatial abilities and thereby to right hemisphere functioning, one need not expect to find any difference between left- and right-handers.

It is true that left-handers are somewhat more likely than right-handers to deviate from the modal pattern of cerebral organization. The higher incidence of left-handedness in mathematically precocious males reported by Benbow (1986b), however, does not require a link between cerebral organization and mathematical ability. Only if we know that the mathematically precocious left-handers are in the subgroup of left-handers with deviant cerebral organization, can we speculate about links between mathematical ability and brain organization. Furthermore, attempts to relate cerebral organization to cognitive ability in other domains have only been modestly successful (e.g., Bryden, 1986): There is little reason at the moment to believe that correlations with mathematical ability would be any more compelling.

Finally, Benbow also links mathematical ability to immune disorders. Geschwind and Galaburda (1986) have argued that handedness is related to immune disorders and to various cognitive deficits, most notably dyslexia. Their studies have generally examined simple associations, but not higher-order ones. Thus, we do not know whether there is, for example, a higher incidence of immune disorders in left-handed dyslexics than in left-handed nondyslexics. The same is true of Benbow's data (e.g., Benbow & Benbow 1984): we are told that the overall incidence of left-handedness, allergies, and myopia is elevated in students with high mathematical ability, but we do not know how these three symptoms are associated with one another, or how they are related to cerebral organization.

Ultimately, Benbow seems to prefer the argument that a bilateral representation of cognitive abilities can lead to exceptional verbal or mathematical ability. Although language skills are bilaterally represented more frequently in left-handers than in right-handers, the overall incidence of bilateral representation is very low, on the order of 2% of the population (Segalowitz & Bryden 1983). Other studies suggest that bilateral representation of visuospatial function is somewhat more common in left-handers (Bryden et al. 1983). However, simply showing that a particular subgroup has more left-handers in it does not demonstrate that the subgroup has more people with bilateral cerebral representation or that bilateral representation has any predictable consequence. What we need to know is whether the few people with bilateral representation of function differ cognitively from those with unilateral representation.

Before Benbow's arguments for a biological basis for mathematical reasoning ability can be accepted, it is necessary to describe the cerebral organization of the individual, rather than depending on a collection of weak associations with variables such as handedness or the presence of allergies. Without direct evidence that a specific pattern of cerebral organization is related to superior mathematical ability there is a danger that people will remember Benbow's speculations and come to think of them as facts.

ACKNOWLEDGMENT
Preparation of this manuscript was aided by a grant from the Natural Sciences and Engineering Research Council of Canada. The author thanks Patricia M. Rowe for her helpful comments.

Spatial visualization and mathematical reasoning abilities

Sarah A. Burnett
Department of Psychology, Rice University, Houston, Tex. 77251

Benbow's analysis of SAT-M test data from several hundred thousand 12- and 13-year olds provides convincing evidence of significant sex-related differences in mathematical reasoning ability among intellectually talented students. Benbow's argument is bolstered by the finding of similar sex-related differences in other countries and ethnic groups. Benbow's target article is not only informative, but presents a significant challenge: Neither our understanding of the nature of sex differences nor our explanations for what appear to be reliable between-group differences in mathematical reasoning abilities is at all satisfactory. Several issues cry out for research, two of which I have chosen as the subject of my commentary.

1. Are important sex-related differences in mathematical reasoning ability limited to the exceptionally talented? Although Benbow's analysis is restricted to the top 2 to 5% of scorers on mathematics screening tests, mean sex differences that are just as large (about 30 points) or larger (about 50 points) are also observed in the college-bound student population as a whole. About half of high school graduates enroll in college. Arbeiter (1985) reports that for college-bound seniors in 1984, the mean SAT-M for men was 495 (SD = 122) and for women 449 (SD = 112). These data were based on *n*'s of 464,881 and 499,804, respectively. Around 1% of the women compared to around 4% of the men scored above 700 on the SAT-M (6,707 and 18,760, respectively); above 750, the corresponding numbers were 1,156 and 5,846. The years of high school math taken does not seem to be an adequate explanation for the sex-related difference in SAT-M scores for this population either. The median SAT-M of women who have had four years of mathematics in high school is 475 (*n* = 231,231); five years of math in high school, 538 (*n* = 49,482). The corresponding medians for men are 518 (*n* = 239,408) and 489 (*n* = 65,912). There were 16% more women than men who scored below 500 on the SAT-M. One practical implication of these differences is that any university for which the SAT-M score figures prominently in its selection process – either singly or combined with the SAT-V score – will admit fewer women than men. It is no accident that in recent years some of the most selective universities (e.g., Harvard, Princeton, Stanford, Rice) have enrolled approximately 20% more men than women despite the fact that there are now more women than men attending college. The difference decreases as colleges become less selective; however, one does not have to limit one's attention to the extremely talented to observe practically significant test score differences (see also Burnett 1986).

2. Can sex-related differences in spatial ability account for the sex difference in mathematical aptitude? Benbow cites mixed support for this hypothesis. A critical analysis of the relevant studies, however, reveals certain trends. Studies reporting negative findings have tended to use (1) some index of mathematics *achievement* (i.e., grades) or problem-solving accuracy rather than speed or aptitude as the dependent measure, (2) high school students of average ability as subjects, and/or (3) inadequate measures of spatial ability. This suggests that negative findings may be attributable to the specific measures taken and subjects tested. Form board spatial tests, for example, often fail to reveal sex differences reliably found with other spatial visualization tests such as the DAT and the Guilford-Zimmerman. On the other hand, when the spatial ability test is speeded, complex, and requires a mental rotation, sex differences are more reliably demonstrated. Furthermore, reliable correlations between the mathematical and spatial constructs are more often found with higher ability high school and college students (Burnett et al. 1979). Our data indicate that students gifted in spatial ability successfully use spatial strategies to solve mathematical problems; however, other strategies are also appropriate, and some students not high in spatial visualization use these effectively to solve mathematical problems. There are obviously a number of ways to solve spatial problems and mathematics problems (see Just & Carpenter 1985). Since there seems to be no sex difference in the ability to use nonspatial strategies, however, the spatial difference appears to be the most likely candidate for the locus of the sex difference in mathematical aptitude.

Benbow points out that mathematical reasoning ability and spatial ability may be directly related and/or these two mental abilities may rely on similar cognitive processes. In recent years the methods and theory of cognitive psychology have been applied to spatial tests with the hope of identifying particular component cognitive processes presumed to be involved in these tests. A number of cognitive processes (e.g., rate of mental rotation of stimuli and quality of image representation) have been identified as making important contributions to individual differences in spatial test performance and sex-related differences have been found on some of these. Lansman et al. (1982) found that mean rotation latency was highly correlated ($r = .78$) with the common variance in six spatial tests. Just and Carpenter (1985) found that low spatial subjects "mentally rotated" cubes at half the rate of high spatial subjects. Pellegrino and Kail (1982), Kail et al. (1979), Petrusic et al. 1979; and Tapley and Bryden (1977) have all focussed on the rate of mental rotation as the major locus of sex differences in spatial test performance. In addition, differences in the precision or quality of image representation have been found to be an important factor in spatial test performance (Poltrock & Brown 1984) and significant sex differences in both image quality and mental rotation have been found in our laboratory (Burnett & Peters 1985). Therefore, the cognitive processes that mathematical reasoning and spatial ability tests are likely to have in common (because these are the processes on which one finds sex differences in spatial tests) seem to have to do with the quality of the visual image representation and the speed with which one manipulates a coded image.

Whereas Benbow's important findings emphasize the need for educational programs aimed at reducing male-female differences in mathematical abilities, one promising research orientation would seem to be exploring the biological underpinnings of the cognitive processes that have been linked to mathematical and spatial reasoning abilities regardless of sex. There are fortunately attractive alternatives to focusing on the sex differences per se.

Sex differences in parallax view?

Susan F. Chipman

Cognitive Science Program, U.S. Office of Naval Research, Arlington, Va., 22217

Benbow and Stanley (1980; 1981; 1982a; 1982b; 1983a; 1983b; 1984) have amply demonstrated that there are more young boys than young girls of junior high school age who will obtain high scores on the SAT mathematics section, but what does that mean? Few members of the general public who have been exposed to headlines like "Are Girls Born with Less Ability?" (Kolata 1980), "Do Males Have a Math Gene?" (Williams 1980), or "Is Biology Really Destiny?" (Goleman 1987) would suspect that representative survey samples of the mathematics performance of students have results like these: In the National Assessment of Educational Progress Second Mathematics Assessment (1977–78), 13-year-old females performed better on computation items (53.4% versus 49.9% correct) and males performed slightly better on problem-solving items (44.1% versus 42.5% correct). A performance difference of comparable magnitude favored girls in algebra (53% versus 51%) but was not statistically significant.

The phenomenon under discussion is really the phenomenon of the exceptional few. Problematic aspects of Benbow's implicit interpretation of this phenomenon lie in two areas. One is the uncritical treatment of the concept of mathematical reasoning ability as it is measured by the SAT-M. The other is some inaccuracy and lack of care in the treatment of research results that might contribute to the interpretation of this phenomenon.

The term "mathematical reasoning ability" seems to say a great deal more than is justified by the reality of the testing behind it. It refers to performance on word problems. The possible effects of sex bias in the content of word problems (Donlon 1973; Donlon et al. 1976; Graf & Ruddell 1972) should be considered. Good problem-solvers work with the content of the problem as much as with the mathematical form (Paige & Simon 1966), using informal understanding of the content to guide choices of mathematical operations. Benbow suggests that mathematical talent is best defined as "the ability to handle long chains of reasoning," but apparently the long chains of reasoning involved in algebra or in generating geometric proofs do not quite count.

Apart from some possibility of sex bias in content, what else might account for the sex differences in mathematics performance that are observed? I was startled to find that Benbow had somewhat misleadingly characterized a book I edited: "In a collection of studies appearing in one book (i.e., Chipman et al. 1985), the most consistent finding emerging throughout was the important role of parents in influencing their daughter's participation in mathematics." Among other things, that book contains a chapter (Chipman & Wilson 1985) that attempts a synthesis and summary of research results somewhat comparable in scope and structure to Benbow's target article; in its conclusions, the research was *not* considered to provide evidence for the role of parental influence in affecting participation in mathematics.

One issue is whether the observed differences in test performance are accounted for by differential course-taking or other indicators of learning time and effort. Far from being the "classical" explanations, obvious environmental variables like schooling were long neglected (Fennema 1974). It is odd that Benbow cites Wise (1985) as reporting small sex differences in high school participation in mathematics when a major point of this secondary analysis of Project TALENT data was the demonstration that most of the sex difference in mathematical performance at the end of high school could be accounted for by the sex differences in course participation. Enrollments in "extra" courses such as calculus, statistics, and computer science, as well as in physical science courses that provide additional practice in solving mathematical problems and maintaining skill, show sex differences (for SMPY students too) that are quite large.

Despite a general lack of evidence that measures of spatial ability can improve upon the prediction of performance in mathematics provided by measures of general ability, Benbow joins many others in citing the alleged "well-documented sex difference in spatial ability favoring males"; but she fails to cite a recent, thorough meta-analysis of this literature (Linn & Peterson 1985) that considerably limits the scope of such conclusions.

As an alternative to the physiological "nerd" factor (pallid, asthmatic, left-handed, and myopic), large sex differences in early developing fundamental interest patterns – that is, "interest in things versus interest in people" – deserve serious consideration (Dunteman et al. 1979). Although mathematics itself seems to be about equally interesting to males and females who are both about equally unlikely to select it as a major, few females select the physical sciences and engineering majors that give mathematical preparation high personal utility. Narrowly focused, intense interest patterns also seem to be associated with high levels of professional achievement and are probably more characteristic of males (Terman 1954; Terman & Oden 1959). Of course, it could be that relative interest in people has a biological basis; rather little is known about early interest development.

Sex, brain, and learning differences in rats

Victor H. Denenberg, Albert S. Berrebi, and Roslyn H. Fitch

Biobehavioral Sciences Graduate Degree Program and Department of Psychology, University of Connecticut, Storrs, Conn. 06268

Benbow proposes that sex differences in mathematical reasoning scores among intellectually precocious students are based upon a combination of biological and environmental factors.

Surely no one can argue with such a general thesis. The interesting questions arise when one tries to specify what biological factors are involved, and the nature of their involvement. To answer such questions one often turns to studies with animals since biological determinants of behavioral processes – even processes as complex as mathematical talent – are unlikely to be restricted solely to humans. We have recent data showing sex differences in a major anatomical structure of the rat's brain; in an independent study we also found sex differences in learning, thus suggesting a brain–behavior association similar in principle to what Benbow hypothesizes.

The corpus callosum of the adult male rat is significantly larger than that of the female, even after adjusting for differences in brain size (Berrebi et al. 1988). The sex difference suggested a hormonal involvement, and thus in another study we injected newborn female rats with testosterone propionate on day 4 and castrated male pups when approximately 24 hours old. Controls were given sham treatments. Measurements in adulthood found that testosterone had masculinized the corpus callosum of the female rats, while castration had only a minor effect, tending to reduce callosal size (Fitch et al. 1987). These findings indicate that testosterone is acting to masculinize the corpus callosum during the perinatal period.

The corpus callosum contains fiber tracts passing from one neocortex to the other. No neurones originate in this structure. Size differences in the callosum therefore, indicate that there are sex differences in cortical regions in one or both hemispheres that are affected by perinatal gonadal hormones. Although we have not yet identified discrete sexually dimorphic cortical areas, our callosal data suggest that they exist. This assumption is supported by direct evidence of sex differences in the brain, both at the cortical (Diamond et al. 1981) and subcortical (Diamond et al. 1982; Gorski et al. 1978; Nordeen & Yahr 1982) levels.

In other studies we have found that male rats learn a swimming version of the Lashley III maze more effectively than do female rats (Freter et al. 1987). There are two major methods by which rats can learn in this experiment. One is by using extramaze information to guide them to the goal box (much as we would use a tall building as a guide post when driving to a particular location). The second is by using intramaze information, such as by memorizing the pathway (turn right, then go left, left again, right, and so forth) or learning their place in the maze. Again there are obvious parallels to the way that a human moves through space. Of course, the most efficient approach is to use a combination of the two procedures.

An analysis of the errors made by the rats in the Lashley maze revealed that females were relying primarily upon extramaze information to navigate through the maze. In contrast, male rats utilized both sources of information and thus learned the maze pathway more effectively.

Extramaze information is spatial information, thus suggesting that the females were using a spatial strategy almost entirely in their maze learning. The quantitative difference between the two sexes in learning may be due to a qualitative difference in strategy selection.

These studies suggest the hypothesis that sex differences in learning are related to brain processes that have been affected by gonadal hormones in infancy. In a very general sense, we believe this to be the case with the human as well. It is noteworthy that Benbow emphasized the importance of spatial ability in mathematically talented students. Spatial skills certainly are adaptive and thus have an evolutionary history. The spatial skills in rats are obviously expressed very differently from the way they are expressed in humans. However, if it is found that the learning differences between the sexes in the rat can be changed by gonadal hormones in the perinatal period, and if these can be related to brain processes, that would lend strong support to Benbow's thesis. In the most general sense, we are arguing for strong linkages between hormones, brain, and behavior in the rat and we propose that similar strong linkages exist in the human as well.

O Tempora, O Mores!

H. J. Eysenck

Institute of Psychiatry, University of London, London SE5 8AF, England

The Benbow target article gives an excellent account of the widely observed superiority of males in mathematical reasoning ability. The account, however, runs into serious difficulties in the section on "Socialization and Sex Differences." Here the author considers "several possible socialization or environmental hypotheses for these differences," and apparently "these classical hypotheses were developed and have received some support for explaining sex differences among children of average ability." These alleged "environmental" causes include the possibilities that (1) females have a lower liking or more negative attitude toward mathematics; (2) females feel mathematics is less useful or important to future career goals; (3) females have less confidence in their mathematical ability; (4) mathematics is stereotyped as a masculine discipline; (5) "significant others" have different expectations for males and females as far as mathematical achievement is concerned; (6) males and females take different mathematics courses; and (7) males and females have different career and achievement motivations. It should be obvious that whether or not differences along these lines are observed, they cannot throw any light on the question of genetic or environmental factors. Let us assume, for the purpose of the argument, that mathematical ability is determined 100% by genetic factors. If that were so, all the alleged environmental factors would appear as a simple consequence of the genetic differences between males and females.

Quite naturally, females would have negative attitudes toward mathematics; they would feel that mathematics was less useful for them; they would regard mathematics as a masculine discipline, and so on. In other words, the consequences of a 100% genetic determination would be identical with those of a 100% environmental determination as regards these various differences. Benbow shows some appreciation of this point, but does not make it clear that these studies are quite incapable of throwing any light whatsoever on the important question of the relative importance of genetic and environmental factors in determining sex differences in mathematical ability. This can only be done by a proper application of the methods of behavioural genetics, which would apportion the phenotypic variance to a variety of genetic factors, like additive genetic variance, epistasis, dominance, and assortative mating, and the environmental variance to within-family and between-family variance. There is simply no alternative to a proper genetic analysis of the variables involved if we want to find an answer to the question Benbow posed.

It is curious that Benbow not only fails to mention this possibility, but that in all the studies carried out there has not been a single twin study that would throw ample light on this problem. Among the childrens' studies there must be well over a thousand twin pairs, certainly a sufficient number to give quite accurate information on all these problems. Furthermore, such a study would be relatively cheap, even if it included siblings as well as twins, and possibly parents as well. Such a study clearly would be of the utmost importance, and we now have the methodology and the statistical sophistication to test different models and to apportion the variance accordingly.

It is also curious that Benbow does not mention studies that have been carried out in the field of the genetics of mathematical ability (e.g., Schaub 1971; Weiss 1982). Possibly she does not consider these studies to be free of criticism, but in that case it would have been advisable to cite them and mention such criticisms as might be appropriate. Altogether, the treatment of the genetic component in this paper is superficial, the data cited are irrelevant, and such rudimentary knowledge as we have accumulated in this field is omitted from the discussion. It is not unusual to find such errors in areas that are supposed to be "delicate" or "sensitive" for ideological and political reasons, but this should not inhibit a scientist from giving a straightforward discussion of the evidence.

Benbow gives the game away when she says that "practically speaking, one must be an environmentalist." Scientists should not, even "practically speaking," adopt a position they know to be erroneous. They should work out the precise quantitative position in a given field, and possibly on the basis of such quantitative knowledge give advice on practical matters. What Benbow says suggests she may not be aware of the true meaning of "heritability" as a population statistic, not an absolute measure. Thus Heath et al. (1985) have demonstrated that very marked changes in the degree of heritability can be produced in educational achievement by means of greater equalization of the educational process, and such studies could and should be conducted in a similar manner in the United States, with special reference to mathematical ability.

On the practical side, the determination of degree of heritability is only a first step; the second step is a determination, by means of experimental studies, of what precisely it is that is being inherited. Thus in phenylketonuria, what is inherited is not mental defect directly, but an inability to convert phenylalanine into tyrosine, with some of the incomplete breakdown products proving toxic to the nervous system. Changes in diet can alter the situation profoundly and prevent mental defect. Thus the most practical thing we can do is to look at the genetic side of the equation and determine just precisely what it is that is being inherited. It seems likely that we will be much more successful if we work along these lines than if we disregard the genetic determination of mathematical ability and carry on the very unsuccessful type of "environmental" manipulation that has characterized this field.

Predicting who our future scientists and mathematicians will be

Helen S. Farmer

Department of Educational Psychology, University of Illinois at Urbana-Champaign, Champaign, Ill. 61820

There is an underlying assumption in Benbow's target article that the SMPY (Study of Mathematically Precocious Youth) sample will produce most of our future eminent scientists and mathematicians. Yet, by her own report, follow-up data on various SMPY samples indicate that of those continuing on to graduate school (i.e., 47%), 42% of the men chose science or mathematical fields, whereas 22% of the women did. These data indicate that less than 25% of the SMPY subjects (i.e., those who score in the top 3% of the seventh grade on the SAT-M) obtain graduate-level training in science or mathematical fields. Where do the other SMPY subjects end up? Perhaps more important is the question, who are the other students in graduate training in science or mathematics and what percentage of the total in these fields do they represent? Without an answer to this question it is impossible to judge the significance of the sex-difference findings associated with the SMPY subjects.

Some support for Benbow's view that the SMPY sample contains many of our future scientists and mathematicians is provided by findings from the longitudinal study Project Talent (Wise et al. 1979). The ninth graders in this study were followed up at age 29. Those who were in occupations related to mathematics or science had scored in the 90th percentile on math achievement tests in the twelth grade. The Wise et al. findings suggest that those who end up in mathematics or science careers are in fact those who earn better scores on math achievement tests, at least in high school. Benbow uses a higher cutoff in identifying the SMPY sample (i.e., top 3%) than the Wise et al. findings would suggest (i.e., top 10%). It is important that future researchers provide an empirical basis for what this cutoff score should be in order to predict most effectively who our future scientists will be.

I turn now to Benbow's review of sex differences in mathematical achievement. In the first section she provides an overview of the SMPY sample and findings related to sex differences on the SAT-M over a 15-year period. Most researchers accept these findings (see, for example, Dix 1987). Benbow also reports another finding regarding SMPY subjects that is consistent with the literature, namely, that girls and women compared to boys and men earn higher grade-point-averages (GPAs) for mathematics and science courses in high school and at a university. Dix noted that among those applying to the fields of physics, engineering, mathematical science, biology, and behavioral and social sciences, men score significantly higher than women on the quantitative part of the Graduate Record Exam (GRE-Q). However, these same women, earn significantly higher grades than men on their course work in these fields. The GRE-Q score consistently underpredicts women's performance in graduate school. It is not known whether this discrepancy is due in part to sex bias inherent in standardized tests such as the SAT-M and the GRE-Q, or due to some sex difference in aptitude and socialization affecting mathematical reasoning and achievement. Benbow does not discuss this discrepancy or its possible causes and I found this omission a weakness in her review.

Benbow also reviews some of the socialization factors thought to affect sex differences in mathematical reasoning and achievement. She is to be commended for placing this review (sect. 9) in the larger context of the literature on sex differences found for boys/men and girls/women in general. Her review, however, is disappointing and far too cursory to permit the reader to have confidence in her conclusions. For example, section 9.2 presents some research on the relationship of math utility and value to career goals. Benbow ends this section by citing a study by Brush (1980), who found no relation between perception of the future value of mathematics and intention to take future mathematics courses. The reader may conclude that math utility and value have no relation to future mathematics course enrollment, and that these variables have no relation to future choice of college major or occupational field. However, other researchers such as Chipman and Thomas (in press) and Eccles et al. (1984) noted that although thinking that mathematics is valuable doesn't have much power in predicting enrollment in mathematics courses during high school, thinking that mathematics is needed for certain occupations *is* predictive of math course enrollment in high school. It appears that how ***math utility*** is operationalized is critical to the outcome obtained.

Another socialization factor reviewed is career and achievement motivation (sect. 9.7). Three studies are considered. Two of these report findings indicating that females are less intrinsically motivated to achieve (Dweck et al. 1978; Nicholls 1980). The conclusions of those authors have been challenged by Frieze et al. (1982), who reviewed 21 studies produced during the years 1968 to 1979 using meta-analysis and found that there was little support for sex differences in success/failure attribution patterns. The other motivational study reported by Benbow was my own (Farmer 1987). In her review of this study Benbow noted the findings concerning the effect of parent and teacher support on motivation. Such support was found to be more predictive of high motivation for females compared to males. Implications of this finding include the view that women need more support than men from important others to be motivated to achieve in challenging careers. The multivariate model used by Farmer to explain career and achievement motivation is based on the view that multiple factors contribute to such motivation, and these in turn to achievement in our society. The search for the *best* predictor is contrary to a multivariate viewpoint.

In conclusion, I do not deny that Benbow and her associates are making a valuable contribution to the understanding of a highly talented sample. However, this may not tell us much about who our future scientists and mathematicians will be.

The new math: Is XY ≥ XX?

Patricia S. Goldman-Rakic and Ann S. Clark

Section of Neuroanatomy, Yale University School of Medicine, New Haven, Conn. 06510

Based on a review of the environmental and physiological variables that may contribute to sex differences in mathematical reasoning abilities, Benbow has concluded that "sex differences in extremely high mathematical ability may be, in part, physiologically determined." The failure to identify causal environmental factors in the studies reviewed is indicative of the complex nature of sex differences in human behaviors and the difficulty of examining them under natural life circumstances. However, Benbow's conclusion that sex differences in mathematical reasoning ability may reflect sex differences in biology and in brain structure and/or function is highly plausible and it is that point we would like to address. Our comments focus on the complexities associated with examining the biological bases of sex differences in behavior and raise the issue of rate of maturation as a variable that might reflect the interaction of biological and environmental factors. Finally, we would like to underscore the advantages of studying sexual dimorphisms under controlled conditions, specifically using nonhuman primates as experimental models.

One strategy for elucidating the biological basis of sex differences may be to study a cognitive process that, unlike mathematical reasoning, can be shown to be sexually dimorphic in nonhuman mammals as well as humans. Spatial thinking (perception and memory) is such a process (Caplan et al. 1985). The spatial abilities of nonhuman primates are among the interests of our laboratory; we have used a combination of anatomical, physiological and pharmacological methods to understand the circuit basis of spatial memory capacity in the rhesus monkey. For example, we have identified a neural network that involves about 17 cortical regions and extensive communication between and among cortical and subcortical structures (see Goldman-Rakic 1987 for review). It is likely that similar networks are present in man. For example, Roland and Friberg (1985) demonstrated different patterns of cerebral blood flow across multiple cortical fields during a variety of assigned thinking tasks, including spatial thinking. Because of the complex nature and demands of cognitive tasks, including the mathematical reasoning abilities evaluated by Benbow and colleagues, the physiological underpinnings of these tasks are probably not so simple to identify or analyze. If we are to understand sexual dimorphism of complex behaviors, we will have to acknowledge the complexity of the neural networks underlying them. Unfortunately, a common error for investigators of sex differences has been excessive eagerness to find simple physiological bases for complex behaviors.

Another consideration with respect to sex differences is the idea that males and females may mature at different rates rather that have fixed differences in function (Goldman et al. 1974). One aspect of our studies on spatial memory that may be particularly relevant to Benbow's finding of sex differences in mathematical reasoning abilities is the role of gonadal hormones in the ontogeny of performance of spatial memory tasks. A number of years ago, we demonstrated a sex difference in performance of the delayed response task by monkeys at 18 months of age (Goldman et al. 1974). After lesions of the orbital prefrontal cortex, male rhesus monkeys showed deficits in their performance of delayed response, whereas female monkeys that received similar lesions were not impaired on the task. These results suggested to us that male and female brains, and in particular the orbital prefrontal cortex, matured according to different timetables. Follow-up studies revealed that females given androgen replacement perform similarly to males, suggesting that gonadal hormones may act directly to modify the cortex during development (Goldman & Brown 1975). These findings emphasize the transient nature of biologically based sex differences that could interact with environmental forces in development. Such differences may easily be overlooked and should receive attention.

Despite the difficulties of studying sex differences in cognitive function, there have been considerable insights into possible mechanisms as a result of studies on nonhuman primates. Biochemical studies of the distribution and ontogeny of gonadal hormonal systems in the developing rhesus monkey have been particularly promising. In recent studies, from our laboratory and others, it has become clear that the cerebral cortex is responsive to androgens during gestation as well as during postnatal life (Pomerantz et al. 1985; Roselli & Resko 1986; MacLusky et al. 1986; Clark et al. 1986). Androgen- and estrogen-binding sites were identified in many regions of the cerebral cortex in fetal, juvenile, and adult rhesus monkeys. The presence of androgen binding and metabolizing systems in the neocortex of the developing primate lends credence to the possibility that the cerebral cortex might undergo sexual differentiation. The biochemical and behavioral studies considered together are consistent with the idea that physiological factors may indeed contribute to observed sex differences in cortical function and the expression of cognitive skills.

However, there remains the problem of pinpointing the specific molecular and/or cellular events or endpoints that are being modified by gonadal hormones – be they synaptogenesis, axon sprouting or degeneration, protein synthesis, or the like – and to determine how these events are translated into behavioral changes. At this time, there is very little information available about the influence of gonadal hormones on these measures in the primate brain. Another perplexing issue is whether hormones act globally across all cortical areas or have limited focused actions on or within discrete cytoarchitectonic regions. For example, Benbow writes that although mathematical ability is superior in males, verbal ability in females is not correspondingly elevated. This lack of reciprocity raises the issue of whether additional cognitive functions that were not evaluated in this study might also be sexually dimorphic. It would be valuable to continue to study these subjects on additional dimensions to explore this possibility. The biological implications of such sexually dimorphic neural networks are great, and these findings in humans point to the need for further studies of nonhuman primates in which experimental models can be developed in circumstances where social and environmental variables can be experimentally analyzed.

Sex differences in mathematical reasoning ability: Let me count the ways

Diane F. Halpern

Department of Psychology, California State University, San Bernardino, Calif. 92407

Benbow has concluded that "sex differences in extremely high mathematical reasoning ability may be, in part, physiologically determined." This is a surprisingly cautious conclusion given that most of the target article documents repeated failure at identifying psychosocial variables that can explain the large and consistent sex differences found with mathematically precocious youth. I concur with Benbow's conclusion as stated. There is a large body of data that suggests there are biological mediators for sex differences in mathematical, verbal, and visual-spatial abilities (for reviews, see Halpern 1986a; 1986b). I am left wondering, however, about the nature of the other "part." To what extent are psychosocial variables involved in the development of sex differences among mathematically precocious youth? In answering this question, I will discuss: (1) possible biases in sampling; (2) differences among ability, aptitude, and achievement; (3) conclusions based on nonsignificant results; and (4) the relative strength of psychosocial and biological hypotheses.

1. Possible biases in sampling. The test results are clearcut: Many more 12- and 13-year-old males score in the highest ranges on the SAT-M than females, but research of this sort is fraught with the obvious problem of subject self-selection. We can only conclude that these results are generalizable to all high-ability females and males in this age range if Benbow sampled equally well from both high-ability populations.

Participation in the Study of Mathematically Precocious Youth (SMPY) is voluntary, with recommendations for student participation made by either teachers or counselors. The inescapable reality is that we can never fully identify girls with high mathematical ability in a sex-typed society because they either never develop their ability to its fullest extent or they learn to devalue and hide it. This bias cannot be corrected by matching scores with same-sex 12th graders or by screening with standardized achievement tests because these other indices are subject to the same biases. Cross-cultural comparisons can only be meaningful if the societies investigated have different sex roles and expectations vis-à-vis mathematics than our own. Thus, the data presented from West Germany and China, though interesting in their own right, cannot be used as evidence of sex-equitable sampling.

Consider the following example: Suppose all (or virtually all) high-ability boys participate in the talent search and from this elite group we focus on the top 3% to make generalizations about all precocious boys. Correspondingly, only 50% of the high-ability girls participate in the search. (The true percentage can never be known.) From this group, the top 3% are identified as mathematically precocious. Not surprisingly, fewer girls than boys are identified at each SAT-M cut-off value.

2. Differences among ability, aptitude, and achievement. Mathematical ability was indexed with the Scholastic Aptitude Test, a test designed to predict how much an individual will benefit from instruction. An aptitude test – unlike an achievement test, which assesses knowledge of previously learned material – should be a measure of potential. SMPY-identified girls benefit from instruction at least as much as SMPY-identified boys. As noted in Benbow's target article "SMPY females received better grades in their high school mathematics classes than did SMPY males" and "SMPY students completed college . . . with female grades being somewhat higher." Benbow also noted that "no sex differences are found in the ability to apply already learned knowledge." It seems that high-ability males have greater mathematical achievement (i.e., they score higher on the SAT-M and are overrepresented in math-related professions), but this does not mean that they have greater mathematical aptitude (capacity to benefit from instruction) or superior skill at applying learned concepts.

3. Conclusions based on nonsignificant results. Benbow examined the "classic" psychosocial explanations of mathematical sex differences that are found in the general population with samples of mathematically precocious youth. She found that although in the general population girls have more negative attitudes toward math, feel it is less useful, and have less motivation to succeed in math, SMPY youth did not replicate these general findings. The only "usual" result was that SMPY girls, like their more ordinary counterparts, are less confident in their mathematical ability than the SMPY boys. It is not possible to draw any meaningful conclusions from these nonsignificant results.

The overwhelming majority of the studies that have examined psychosocial explanations of cognitive sex differences show that these variables are important. The failure to replicate classic psychosocial findings with mathematically precocious youth supports my earlier position that the highly able girls who maintain the "usual" math attitudes did not participate in the search and consequently were not identified as mathematically precocious.

4. The relative strength of psychosocial and biological hypotheses. The involvement of several biological systems in the development of mathematical and spatial ability has been consistently documented. For example, findings that handedness or amount of testosterone available at puberty influences mathematical ability cannot be explained with psychosocial variables. A colleague and I recently found that sex and familial sinistrality (number of immediate family members who are left-handed) had a significant effect on university mathematics and English placement test scores (Martinez & Halpern 1987). Biological theories (e.g., sex-differentiated cerebral organization) are needed to explain results of this sort.

The real question is not whether psychosocial or biological variables are the determinants of sex differences in mathematics, but what is the relative contribution of each. I would amend Benbow's assertion that biological variables must be *in part* responsible for some of the observed differences by adding specific reference to psychosocial influences as well. The relative importance of psychosocial and biological contributors to mathematics differences in highly able youth is impossible to determine. It may be that for this elite group, biological variables, which are also responsible for the increased incidence of left-handedness, allergies, and myopia, play a greater role than for the rest of the population. Future comparisons across ability group levels will allow us to examine this possibility.

Causes of mathematical giftedness: Beware of left-handed compliments

Curtis Hardyck

Departments of Education and Psychology, University of California, Berkeley, Calif. 94720

I claim no expertise at all about the area of mathematical reasoning ability in intellectually talented preadolescents. I will accordingly, limit limit my commentary to Benbow's section on "Physiological Correlates." However, I did read the entire target article, and in so doing, noted an interesting shift in literary style beginning with section 12, "Physiological Correlates of Mathematical Reasoning Ability." Such phrases as "may be;" "would seem reasonable;" "is possibly more," and similar phrases designed to convey incertitude occur more often in section 12 ff. than in the preceding 11 sections of the paper. When tallied in my review copy, the first 22.5 pages of test contain five such hedges. The next six pages of text beginning with section 12, contain 25 modifiers in various forms.

This may reflect an overzealous concern for proper scientific caution. Alternatively, it might represent uncertainty on the part of the author as to the durability of the interpretations offered – a more plausible view in my opinion.

After considerable experience in pointing out the flaws in research oriented toward proving the left-handed stupid (Hardyck et al. 1976), it is refreshing to see an article arguing that they have superior ability. However, it is not reassuring to see the same errors of interpretation, the same distortions of data to fit preconceptions, and the same selective reviews and interpretations of existing literature. The data presented here are no more convincing in showing that the left-handed are especially gifted than the cumulative record is in trying to show that left-handedness is proof of cognitive deficit.

Benbow operates from a rather antiquated view of cerebral lateralization. Early views of lateralization held that specific functions were almost, if not absolutely, hemisphere-specific. Benbow refers to tasks mediated by the right hemisphere, such as those involving musical or spatial ability, and to higher frequencies of left-handed mathematics teachers and students, architects, artists, and so forth. She even refers to dyslexics, "who are often left-handed [and], have superior talents in certain areas of nonverbal skill." The relation of this group to the mathematically precocious of any gender is not evident. (And, just for the record, there is *no* statistical evidence that left-handedness is more common in dyslexics.)

Benbow argues that "mathematical reasoning ability, especially in contrast to computational ability, may be more strongly under the influence of the right hemisphere." She then predicts that the left-handed may have an advantage. The reasoning is baffling. We know from a rather extensive set of studies (Dimond & Beaumont 1974; Hardyck 1977), that the left-handed show greater *bilaterality* of function than do the right-handed; especially in verbal tasks. How a right-hemisphere advantage can be present for mathematical reasoning ability is not at all clear, unless (a) one assumes that the right hemisphere controls all spatial ability, (b) one equates spatial ability with mathematical ability, (c) one assumes that the left-handed are somehow right-hemisphere superior and therefore (d) that they can perform mathematical reasoning tasks more effectively. Such a hypothesis, while no more convoluted than others presented in Benbow's article, has no support in what is known about lateralization in the left-handed. It is unfortunately necessary to add that the author seems unaware of the familial–nonfamilial distinction commonly observed in current studies of handedness (McKeever et al. 1973) or of recent developments in theories of laterality (Friedman & Polson 1981; Sergent 1982), which suggest that tasks are not hemisphere-specific.

The interpretation is confounded when Benbow later hypothesizes that "it may be that bilateral representation of cognitive functions (rather than greater specialization of the hemispheres) is associated with extremely high mathematical and verbal reasoning abilities." It should not be necessary to point out that this hypothesis is contradictory to the earlier one postulating right-hemisphere superiority. What is puzzling is the evidence offered in support – a pilot study on "extremely precocious students" in which a right visual field advantage for letter-matching tasks is found. Benbow's specific statement is: "The response times for the letter-matching task . . . were somewhat higher when the problem was flashed to the left rather than the right visual field." If "higher" response times mean longer, then the result is identical to the sort of result that is usually found for this task, whether one's subjects are extremely precocious, just average, or not very bright. (For a review, see Bradshaw & Nettleton 1981). If "higher" means shorter response times, then this result could be interpreted as support for the hypothesis of right-hemisphere superiority advanced earlier. I see no way to interpret this result as supportive of bilaterality of function. Benbow concludes this section by stating: "Thus, bilateral or diffuse representation of cognitive functions and/or a strong right hemisphere may be another possible biological correlate of extremely high mathematical reasoning ability." Given such a statement, it becomes increasingly difficult to understand what the author is advocating. The discussion that follows, referring to sex differences in brain asymmetry and differences in the thickness of the corpus callosum, is equally clear and supportive, especially when Benbow concludes: "Thus it is not clear whether differences in the size of the corpus callosum could have any references to sex differences in ability."

The discussion of allergies, hormones, and myopia suffers from similar problems in interpretation. It does lead to some interesting selection possibilities, however. Given the concerns and suspicions currently extant about ability testing, I would turn the data presented here to advantage. If I wanted to find a group of mathematically precocious youth, I would ingore the SAT and lurk around the offices of allergists and optometrists. If these factors are as important as claimed, they might work just as well as the SAT.

In summary, I think this would have been a much better paper if sections 12.0 ff had been thrown away. The author should be commended for her efforts to understand the bases of an ability that it is clearly important to identify and understand. The fact that I think her physiological evidence nonexistent or irrelevant does not detract from the rest of the material presented and I will leave commentary on that topic to those qualified to evaluate it. Benbow concludes: "Of course, some of the above discussion on physiological correlates is speculative." I agree, and I think there is too much premature speculation, adding nothing to our current level of knowledge about mathematical ability and nothing to our knowledge of cerebral lateralization.

A variety of brains?

Richard A. Harshman
Department of Psychology, University of Western Ontario, London, Canada N6A 5C2

My colleagues and I have found results, in less gifted individuals, that are highly consistent with Benbow's two main conclusions. Our data indicate that (a) individual differences in reasoning ability are related to differences in brain organization and (b) sex differences in cognitive abilities are at least partly neurological in origin. We have also found clear sex differences in the cognitive correlates of left-handedness, and these may help explain what would otherwise be troublesome inconsistencies in Benbow's handedness data. However, Benbow's initial explanation of her results may be too simple.

Data supporting Benbow's main claims. Many of our findings have been presented elsewhere (Harshman et al. 1983; Harshman & Hampson 1987), so only a few key points will be noted. (a) Sex differences in cognitive abilities are not the same for right- and left-handed individuals, suggesting that brain variation may be involved. (b) Many of these sex-by-handedness interactions will change form, depending on the subjects' level of reasoning ability. For example, in our above-median subjects, the spatial ability of left-handed women was better than that of their right-handed counterparts, but the spatial abilty of left-handed men was worse than that of their right-handed counterparts. The reverse pattern was usually found in the below-median subjects. (c) Significant three-way interactions of handedness, sex, and reasoning level were also found for other cognitive abilities, including verbal fluency and perceptual speed, but the *form* of the three-way interaction was different for each ability (d) Many apparent inconsistencies in the literature on handedness and spatial ability can be reconciled if one considers the level of reasoning ability of the samples involved.

Our interpretation is consistent with Benbow's approach to her data. We find that a purely environmental explanation for *interactions* of handedness, sex, and reasoning level is very hard to construct, particularly one that explains the different form that the interactions take for different cognitive abilities. In addition, our dichotic listening data show similar three-way interactions, consistent with the idea that variations in brain organization may underlie the cognitive differences (Harshman & Hampson 1987). We therefore interpret these cognitive differences as arising in part from handedness, sex, and reasoning-related differences in brain organization. In fact, the distinctive cognitive profiles of each subgroup suggest that each different *combination* of these three factors may pick out subjects with somewhat different brains – there may be a variety of normal brain organizations. These variations presumably arise during embryological development, when biological factors such as sex hormones interact with others related to the development of handedness, specific cognitive abilities, and so on. [See Gualtiori & Hicks: "An Immunoreactive Theory of Selective Male Affliction" *BBS* 8(3) 1985.]

Neurological evidence for such variation is provided by Mateer et al. (1982). They found indications that sex and verbal IQ are both related to differences in the localization of verbal functions in the left hemisphere, as determined by direct cortical stimulation.

In light of our findings, and those of Mateer et al., it is not surprising that exceptionally high reasoning ability would correspond to unusual patterns of handedness and brain organization. The data also support Benbow's physiological interpretation of sex differences.

Complexities in the handedness data. Certain complexities of Benbow's handedness data seem inconsistent with the single-factor simplicity of her biological hypothesis. Here are two examples: (a) Although gifted students of both sexes show an increased incidence of allergies and myopia, only the males show a substantially elevated incidence of left-handedness (gifted females do not have significantly more left-handedness than controls; see Benbow 1986b, Table 1). If an increase in left-handedness is the sign linking unusual brain organization with giftedness, why is this sign so weak in gifted women? (b) Both the mathematically gifted *and the verbally gifted* males show increased left-handedness relative to females, yet only the mathematically gifted males outperform their female counterparts. Why?

These (and other) apparent inconsistencies are not an insurmountable problem. It may be necessary, however, to modify the idea that gifted males and females differ along a single continuum (defined by degree of hormonally induced right hemisphere or bilateral brain development). The brain organizations that result in giftedness may be somewhat different in females and males, with the male forms more strongly associated with left-handedness. This would be consistent with previous findings of sex differences in the cognitive correlates of left-handedness and with our three-way interactions, which suggest that (moderately) high-reasoning males and females may differ in brain organization.

The second inconsistency can be resolved in a similar way. Students selected for exceptional SAT-V scores probably do not have the same brain organizations as those with exceptional SAT-M scores. In the embryological development of verbally gifted brains, somewhat different genetic factors may be involved, ones which do not react in the same way to sex hormones. As a result, verbally gifted male brains develop characteristics that lead to a particularly elevated rate of left-handedness, but not to any verbal advantage (or disadvantage) relative to verbally gifted females. As in Harshman and Hampson (1987), sex differences in the neurological correlates of left-handedness may vary across groups that have different cognitive specializations. Once again, the complexities in Benbow's data make sense when one allows for "a variety of normal brain organizations."

As one last note of comparison, it is interesting that the highest incidence of left-handedness was found in Benbow's *verbally* gifted males. This is consistent with our finding (in the high-reasoning group) that the verbal ability of males but not females benefited from left-handedness (Harshman et al. 1983). If, as Benbow suggests, the SAT-M is not highly spatial, then this may explain why it also identifies a gifted sample in which left-handedness is elevated among males.

There will doubtless be debate about the practical implications of Benbow's work, but I believe its real significnce is as basic research. Benbow and her colleagues have contributed rare, invaluable data on individual and sex differences in cognition. Their search for physiological correlates bears directly on basic questions concerning the development of the brain, and the relationship between brain structure and function. I hope they continue this important work.

Hormonal influences on human cognition: What they might tell us about encouraging mathematical ability and precocity in boys and girls

Melissa Hines

Departments of Anatomy and Psychiatry and Biobehavioral Sciences, UCLA School of Medicine and Neuropsychiatric Institute, Center for the Health Sciences, University of California, Los Angeles, Calif. 90024

Benbow offers two main arguments to support the suggestion that biological factors contribute to sex differences in mathematical precocity: Sociocultural factors alone cannot explain the sex differences and mathematical precocity is correlated with physiological characteristics, such as left-handedness and allergies. Neither argument is logically sound. The first depends on proving the null hypothesis – an impossibility. The second assumes that correlation implies causation – it does not. These logical principles have been discussed extensively elsewhere. Hence, I will not address them further here. Instead, I will address three questions related closely to my own area of expertise, as well as to possible biological influences on mathematical ability: (1) Is there support for the hypothesis that gonadal hormones influence mathematical ability? (2) Why has this hypothesis not been tested directly? (3) Why is it important to test the hypothesis directly, instead of interpreting indirect evidence to support it?

Animal research indicates that gonadal hormones have powerful influences on the development of neurobehavioral characteristics that show sex differences (cf., Arnold & Gorski 1984; Beatty 1979; Goy & McEwen 1981; Hines & Gorski 1985). There is also some evidence that hormones exert similar, though subtler, influences on human development, including the development of characteristics, such as visuospatial ability, language lateralization, and handedness, that might relate to mathematical ability (cf. Gordon & Galatzer 1980; Hier & Crowley 1982; Hines & Shipley 1984; Money 1964; Nass et al. 1987; Resnick et al. 1986; Rovet & Netley 1982).

No reliable information is available relating hormones directly to mathematical ability, however. Although Benbow states that "progesterone exposure enhances numerical ability," the study cited (Zussman, Zussman & Dalton 1975, cited in Reinisch, et al. 1979) used methodological procedures that preclude this conclusion. Studies of hormone-exposed people are not truly experimental. Hence, results are most conclusive when outcomes agree with predictions from animal research, when same-sexed relatives serve as controls, and when changes are seen only on measures that show sex differences (cf. Hines 1982; Reinisch & Gandelman 1978). The cited study satisfies none of these criteria. First, progesterone typically reduces masculine characteristics in other species. Second, progesterone-exposed offspring were not compared to relatives. Third, progesterone-exposed offspring showed increases in verbal abilities, as well as numerical ability. Early reports suggested that androgen exposure increased intelligence (cf. Money & Lewis 1966), but subsequent studies found that both androgen-exposed individuals and their relatives had elevated intelligence test scores. Apparently, nonhormonal factors, such as socioeconomic background, caused the increase (cf. Hines 1982; Resnick & Gottesman 1978). Similar factors probably explain the apparent cognitive enhancement in progesterone-exposed offspring. The hypothesis that gonadal hormones influence mathematical ability has not been tested directly because tests of mathematical ability do not show large enough sex differences. This may seem surprising, given the ratio of mathematically precocious boys to girls. Sex differences in mathematical ability seem large when the focus is on the upper tail of the distribution (the mathematically precocious). However, they are not as large in the general population as are sex differences in other characteristics, such as certain visuospatial abilities, for which effect sizes are 0.80 or larger (cf. Sanders et al. 1982). The root of the problem may be that we do not know exactly which aspects of mathematical ability show sex differences. If more specific measures that show larger sex differences are developed, it should be possible to investigate hormonal contributions to mathematical ability, as has been done for visuospatial ability, by examining mathematical ability in individuals who have developed in unusual hormonal environments. However, practical considerations such as the limited number of hormone-exposed individuals and difficulties in measuring hormones prenatally will probably preclude examining hormonal contributions to mathematical precocity.

It is interesting to contrast our knowledge of sociocultural and hormonal influences on mathematical ability. Sociocultural factors appear to contribute to sex differences in the general population, but Benbow's studies suggest that they are less important among the precocious. In contrast, we do not know whether hormones contribute to mathematical ability in any population, and we probably cannot test their impact among the precocious. Thus, sociocultural factors have more support than hormonal factors. Hormonal factors may seem promising merely because they have not been examined. In addition, the hypothesis that they contribute to mathematical precocity could remain viable simply because it cannot be tested. Common sense suggests that both sociocultural and biological factors contribute to mathematical precocity. On the most basic level, adequate neural development, as well as adequate educational experience, is needed to develop mathematical ability, both within and beyond the normal range. Direct assessment of specific biological and social contributions is important, however, to achieve the full understanding required to modify their effects.

For instance, assessments of hormonal influences on play behavior and visuospatial ability have suggested avenues for increasing visuospatial ability. Females exposed to high levels of male hormone during early development show more male-typical play patterns (Ehrhardt & Baker 1974) and more male-typical (i.e., enhanced) visuospatial ability (Resnick et al. 1986) than do unexposed female relatives. Other evidence suggests that normal girls with male-typical play patterns also show greater visuospatial ability than normal girls with female-typical play patterns (cf. Sprafkin et al. 1983), and one study has found that encouraging girls to play in male-typical ways increases their visuospatial ability (Sprafkin et al. 1983).

As our world becomes increasingly technological, mathematical ability will become increasingly important for the well-being of our society, as well as for the success and personal satisfaction of individuals in it. Currently, schoolchildren in the United States in general, regardless of gender, perform worse on tests of mathematical ability than do children in many other countries. It seems that one of our responsibilities is, therefore, to determine how all children can be encouraged to develop mathematical ability. The available data suggest that both biological and environmental factors contribute to visuospatial ability. The same is likely to be true for mathematical ability. Yet the data on visuospatial ability also suggest that childhood play experiences may provide a final common pathway, a point where intervention can enhance visuospatial ability, regardless of biological factors. Research aimed at identifying specific aspects of mathematical ability that show sex differences and the specific biological and social factors influencing them may reveal ways to increase mathematical skills in girls and boys as well.

Sex differences in variability may be more important than sex differences in means

Lloyd G. Humphreys

Department of Psychology, University of Illinois, Champaign, Ill. 61820

The phenomenon of extreme ratios of boys to girls with respect to high attainment in mathematics is solidly established by Benbow's data. Her target article also helps to clarify the issues involved and to suggest, in terms of correlates, possible causal factors. It is not possible, however, to determine from the article the size of most of the correlates reported. Thus their importance in suggesting possible causes is uncertain.

For the first time in what I have read of this literature, attention is paid to a sex difference in variability. This is a welcome contribution. It is obvious that a large mean difference with equal variability about the means can produce the larger proportion of boys at gifted levels. It is also obvious that a large difference in variability about equal means can produce the same phenomenon, but this possibility has been overlooked by most investigators. Then, of course, somewhat smaller differences in both means and variances can produce ratios just as extreme. Causes for the two kinds of differences could be independent of each other.

In Project Talent (Flanagan et al. 1962), more than 50 different tests were administered to students in the ninth through twelfth grades in a probability sample of more than 900 high schools. Sex differences in means were found for approximately equal numbers of the tests, but a substantial proportion of those tests showed greater male variability. Talent tests were short and many distributions were skewed, making differences in variability difficult to interpret, but the data are at least suggestive.

Tests in the growth study of the Educational Testing Service (Hilton et al. 1971) provide better data concerning variability. Tests were administered in a wide sampling of schools to students in the fifth, seventh, ninth, and eleventh grades during the sixties. Eight of the cognitive tests had adequate length and were scaled longitudinally from form to form in order to measure growth. The four grades and eight tests provide 32 possible differences in means and variances. In somewhat more than one-half of the test-grade combinations, mean differences favored males, but there were only five such combinations for which males were not more variable. Furthermore, for all five the differences were small, and three of the five occurred in one of the eight tests. This was Listening, a measure of aural comprehension, of the Scholastic Tests of Educational Progress (STEP). The Writing test of STEP on which females had the largest advantage showed differences in variability favoring males. The mean sex difference on the Quantitative test of the School and College Aptitude Test (SCAT) favored females in grade 5, but a difference in variability produced an excess of males at gifted levels.

An example of an interaction between differences in means and in variances can be found in the data of Hilton et al. (1971) and is shown in Table 1. As the difference in means on the Quantitative test increased from grade 5 to grade 11, the distance from the female mean in female standard score units to the point in the male distribution three standard deviations above the male mean in male units also changed in minor ways. However, the changes did not increase with the increase in the mean differences. Perhaps it is only coincidental that the most

Table 1 (Humphreys). *Estimated ratios of gifted boys to girls on SCAT-Quantitative at four grade levels as a function of differences in means and variances (data from Hilton et al. 1971)*

	5		7		9		11	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
N	1334	1718	1334	1718	1334	1718	1334	1718
X	257.0	257.3	275.5	274.4	290.5	289.2	296.6	292.4
Sx	9.2	8.1	13.9	12.3	17.1	16.0	18.1	17.5
Ratio	13/4		13/2		13/5		13/4	

extreme ratio of males to females occurred in grade 7, which is the point in the educational system in which the Hopkins searches for mathematical talent have been largely concentrated.

The growth study data were longitudinal. Even though the broad sampling used was not strictly random, the longitudinal feature of the data drastically reduced the effects of sampling errors in the comparisons. Sample sizes in the report cited varied slightly from test to test, but the smallest Ns provided sampling stability. These considerations provide confidence in the phenomenon portrayed in Table 1, but do not explain the even larger ratios of males to females in the talent searches. The SAT-M used in the talent searches is a much more difficult test for seventh graders than the form of SCAT-Quantitative used in the growth study and has a much higher ceiling. The present data do show a complex relationship between central tendencies and variabilities for the two sexes from grade school to high school. The research on differential course taking in high school is irrelevant to sex differences in ratios of gifted children in late high school on SCAT-Quantitative.

The information about the physical correlates of mathematical attainment is interesting, but more and better data are needed. The correlates described are probably only a small sample of the number that can be found. In a recent study of the self-reports of tenth-grade students in Project Talent (Humphreys et al. 1986), we found small significant correlations with the Talent intelligence composite of responses to 43 different questions about health and physical disabilities. All but five were in the direction of illness and disability being associated with lower intelligence. Included in the five exceptions were three questions about visual problems. One of these asked about seeing things at a distance. Being informed by a doctor of having an allergy was also an exception to the general trend, but separate reports about hay fever and asthma were negatively correlated with intelligence. The fifth exception was a report of having had mumps.

Samples of males and females, each with Ns of almost 10,000, were analyzed separately in obtaining the preceding data, but sex differences in the correlations with intelligence were minimal. A composite formed with multiple regression weights derived from one sample was correlated with the intelligence criterion in an independent sample. Correlations for both sexes were slightly smaller than .40, and there was a small amount of common variance with a measure of socioeconomic status. Male means were smaller than female means by about .15 of a standard deviation, indicating that males reported a somewhat higher incidence of illness and disability, and male variances were larger. Succumbing to environmental accidents is also a more variable phenomenon in males.

Better data are also needed than the p-values associated with the simplest form of hypothesis testing. That a difference in means is greater than zero, or that a correlation is greater than zero at a given level of alpha is not satisfying. This test informs the reader about the sign of the relation between two variables but provides no information about the size. We need to know how much variance the two variables have in common and with what precision the amount of common variance is determined.

The prevalence of opportunistic sampling accompanied by little solid information about the characteristics of the population from which the sample was drawn represents a third area in which better data are needed. Every sample cannot be a probability sample from the national population, but the populations from which samples are drawn should be carefully described. The description should also be as quantitative as possible.

All in all, Benbow's target article describes an important psychological problem with broad social implications and makes a good start in outlining its dimensions. Psychological problems cannot be solved on the basis of either conservative or liberal ideology. Legislation can only influence the rules of the game, allowing equality to appear if it exists.

Sex differences in mathematical talents remain unexplained

Earl Hunt

Psychology Department, University of Washington, Seattle, Wash. 98195

Preadolescent boys are much more likely to do well on the SAT-M than preadolescent girls (Benbow & Stanley 1980; 1983b). Is this due to biologically produced differences in cognition? The evidence is weaker than Benbow suggests.

A crucial step in Benbow's argument is her assumption that the SAT-M tests mathematical reasoning in preadolescents. The reference cited is a talk given at a scientific meeting (Minor & Benbow 1986) and speculation about how SAT-M problems might be attacked. This is weak and virtually unverifiable evidence. Nevertheless, the SAT-M is a good test, so Benbow has established a case.

Do preadolescent differences on the SAT-M predict subsequent sex differences in mathematical activities? Yes and no. Girls in the Study of Mathematically Precocious Youth (SMPY) do less well than boys on subsequent achievement examinations but have slightly *higher* grades in related courses (Benbow & Minor 1986). The discrepancy is not surprising. High levels of cognitive performance are usually more test-specific than are low levels. Benbow et al. (1983) found this when they contrasted the factor structures of mental test scores in SMPY participants and their parents. Given specificity, what counts as evidence of talent? When faced with test-grade discrepancies, Benbow and Minor (1986), as well as Stanley and Benbow (1986), stated that grades are not really due to "ability." More correctly, timed mental tests and course accomplishments tap somewhat different psychological variables. All may be related to later achievement.

If one is interested in predicting concrete achievements in mathematics instead of predicting test scores, it may be appropriate to consider "noncognitive" variables, such as persistence and willingness to study, as well as measuring pure mental power. [See also Macphail: "The Comparative Psychology of Intelligence" *BBS* 10(4) 1987.]

The conclusion that environmental effects do not influence mathematical ability can only be drawn if the appropriate variables have been measured using accurate techniques. The issue here is one of accuracy of measurement, not statistical power. Are the SMPY team's variables accurate anough to permit drawing a conclusion based on a failure to find relationships? Probably not. The SMPY relied upon a retroactive report of parent-child interactions, after the child's "precocious" status was known. Participants' retroactive reports do not always agree with observation. The cited reference states that the SMPY explicitly investigated parental tutoring in mathematics, such as helping with homework. Differential involvement of boys and girls in situations requiring quantitative, logical, or spatial reasoning (e.g., carpentry) was not studied. The SMPY measures appear to have concentrated on adult–child interactions, as is typical in this field. Peer influences may be important. The environmental question is still open.

The direct evidence for biological links to mathematical talent is weak. Are well-established sex differences in cerebral organization linked to mathematical ability? Benbow (1986b) reported that the incidence of left-handedness was higher in exceptionally gifted SMPY children than in their siblings. But different criteria of handedness were used for the groups! For the SMPY group, sinstrality was counted if there was *any* reported left-hand preference in a variety of activities. Sibling classification was based on parental classification. Furthermore, in calculating the incidence of sinstrality, Benbow (1986b) excluded Asians, although Asians were markedly overrepresented in the SMPY study, and their incidence of left-handedness is about half of that in the general population. Benbow (1986b) claimed that perhaps Asians might not report sinstrality accurately. Perhaps. However, excluding data opposed to one's hypothesis calls for careful explanation.

Dyslexia occurs more frequently in boys than girls. Do dyslexics "have superior talents in certain areas of nonverbal skill"? Benbow cites a paper by Geschwind (1982) as evidence, but Geschwind's paper contains only opinion and anecdote. Einstein's speech developed slowly – which is not dyslexia – but he read Kant's *Critique of Pure Reason* at thirteen (Clark 1971). More generally, the correct question is whether or not the rate of gifted mathematicians is higher in dyslexics than in the general population. Vellutino (1979) does not mention this in his extensive review of dyslexia.

If sinstrality, mathematical ability, and gender covary, Benbow (1986b) should have found sex by handedness effects in her data. She did not.

Benbow (1986b) found elevated incidences of self-reported asthma and myopia in the exceptionally high scorers, compared to their siblings and other control groups. However, there were no indications of sex differences in incidence, so the relevance to sex differences in mathematical ability is not clear. This finding alone does not compel us to accept a biological hypothesis. Myopia and asthma may restrict a child's activities to a life-style more suited to the development of mathematical talent than would that of the prototype junior high schooler.

Benbow makes the link between sex differences, biology, and mathematics sound firmer than the data warrant. Her target article could mislead the trusting reader.

ACKNOWLEDGMENT

Preparation of this comment was supported in part by Contract N00014-86-C-0065, Computer-Based Assessment of Cognitive Abilities, between the Office of Naval Research and the University of Washington. (The opinoins expressed here are the author's and do not represent the views of the Office of Naval Research.)

To understand sex differences we must understand reasoning processes (and vice versa)

Nancy Ewald Jackson

Department of Psychology, University of Colorado, Boulder, Colo. 80309

A substantial and reliable sex difference has been found in the SAT mathematics test performance of intellectually talented 12 and 13 year olds. Benbow patches this evidence together with the results of studies describing experiential, psychological, and physiological correlates of performance on various tests of mathematical reasoning and concludes that the SAT performance difference is the result of biological as well as environmental causes. I question the appropriateness of seeking, in weak data, causes for performance differences that have yet to be defined at a process level. When the nature of mathematical reasoning is better understood, a search for developmental causes of sex differences in components of that performance is likely to be more fruitful.

As Benbow has noted, "we do not have a theory to explain or define mathematical talent or mathematical reasoning ability, nor do we currently have data to suggest one." Although there is much relevant work that Benbow has not cited (e.g., Nesher 1986), she is right, in that current theories tend to focus either on elementary mathematics or on particular kinds of advanced problems; they have not been designed to account for performance on the diverse set of reasoning problems represented on the SAT. Thus, the construct validity of the SAT and other tests of mathematical reasoning ability is weak. It is hard to know what theoretical significance the differences reported by Benbow might have. Nonetheless, the finding of reliable sex differences on the SAT and some other mathematics tasks suggests that the ability to account for sex differences should be one criterion against which emerging theories of mathematical reasoning are evaluated.

In order to account for data such as Benbow's, a theory of mathematical reasoning must specify the processes that account for average differences in problem difficulty, those that account for individual differences in performance, and those that account for sex differences. It is possible that the sources of each kind of variance will be different, and the overall pattern will be important for developing hypotheses about developmental origins of performance differences.

Benbow's suggestion that individual differences in mathematical reasoning ability are accounted for by differences in the ability to handle "long chains of reasoning" is not, at least when taken at face value, a promising start on the needed theorizing. It implies that problem-solving is a mechanical process of working sequentially through a series of steps that are fixed by the nature of the problem. Such a conceptualization is not consistent with evidence that the steps, or rules, change as a problem-solver learns about a class of problems (Nesher 1986). Neither does it suggest how individual differences in preferred strategies might influence the accuracy and efficiency of the solution process (Siegler, in press).

Differences in strategy choice may be an important component of sex differences in performance on mathematical reasoning tests. Dweck (1986) and her colleagues have found that intellectually talented boys and girls differ in their responses to failure; girls are demoralized by getting a problem wrong and prefer easy problems that they can solve without failure. Over time, children who try to avoid failure by habitual reliance on careful but unadventurous problem-solving strategies may develop increasingly large gaps in their success rates for familiar and unfamiliar types of problems, relative to the gaps shown by less cautious problem solvers. When faced with the repeated and varied challenges of a test such as the SAT, they may avoid approaches such as working with unfamiliar modes of representation that seem risky but that offer the best chance of success. The influence of this sort of developmental process on mathematical reasoning needs to be understood before sex differences in performance are attributed by default to biological factors.

When sex differences in mathematics performance are understood in terms of both problem-solving and developmental processes, it may be possible to provide the "remedial" instruction for girls that Benbow suggests. However, a process model of mathematical reasoning may provide evidence that girls' reasoning processes are different from, but not necessarily inferior to, those of boys. The current higher mathematics has been constructed almost entirely by men who were doing the best thinking of which their male minds were capable (Keller 1985). This mathematics is beautiful and useful, but it is not the only mathematics that could exist. Perhaps, as more women go far enough within the male system to do creative work of their own, they will produce new conceptualizations and operations that take advantage of characteristically female modes of reasoning that have yet to be identified.

Sex differences in arithmetic computation and reasoning in prepubertal boys and girls

Arthur R. Jensen
School of Education, University of California, Berkeley, Calif. 94720

In her review of the evidence for sex differences in mathematical reasoning ability among intellectually gifted students, Benbow states: "We do not know how these findings may relate to students of average ability. Differences may be smaller at that level." She also states: "It is in junior high school that the sex difference in mathematics first becomes apparent. Girls excel in computation, boys on tasks requiring mathematical reasoning." She notes that, although the sex difference in mathematical reasoning is apparent by age 12, "It is rather difficult to obtain data below that age since there are no tests of mathematical reasoning ability for younger students, probably because the elementary curricula tend to cover mainly computation and basic arithmetic facts."

I have obtained data, not previously published, that throw some light on these and other issues posed in Benbow's target article. They consist of scores on three subtests (Arithmetic Computation, Arithmetic Concepts, and Arithmetic Applications) of the Stanford Achievement Tests battery obtained on *all* of the 3,112 pupils attending regular classes in the fourth, fifth, and sixth grades of all the elementary schools in one California school district. Virtually all of the boys and the vast majority of the girls are prepubescent, averaging about 9, 10, and 11 years of age in the three grades. The Computation test requires a knowledge of numeric facts and the various arithmetic operations known as mechanical arithmetic. The Concepts and Applications tests involve knowledge of quantitative concepts and the use of such concepts in reasoning about quantitative "thought problems" in which the required arithmetic operations are not explicit but must be inferred.

Table 1 shows the mean standardized sex differences (male minus female) on these tests. The reported difference (D) is what Benbow refers to as "effect size," that is, the raw-score mean difference divided by the average of the within-groups standard deviations. Also shown is the variance ratio (*F*) for male/female. In all three ethnic groups, girls perform very significantly better than boys on Computation (with an average effect size of −.20). On the Concepts and Applications tests, however, white boys significantly outperform their female counterparts (average effect size of +.185). The sex difference in the Asian (Chinese and Japanese) group favors girls in grades 4 and 5, then reverses in grade 6. The number of Asians is small, however, so we cannot make too much of this finding. The black pupils (total $N = 1{,}282$) show hardly any sex difference (average effect size of +.01) on Concepts and Applications, which agrees with Benbow's statement that the sex difference in mathematical reasoning ability is smallest among blacks. Also, only in the white group is the male/female variance ratio (*F*) consistently greater than 1. It is never significantly greater than 1 (overall mean $F = .95$) in the black group, even with its large *N*. The marked ethnic difference in the magnitude of the sex difference raises the question of whether this effect is attributable to cultural or biological factors. Benbow seems to favor the hypothesis that the sex difference in mathematical reasoning ability in her white sample is attributable to a biological difference between the sexes. But how would she explain the

Table 1 (Jensen). *Standardized mean male/female difference (D) and male/female variance ratio (F) on Arithmetic Computation, Concepts, and Applications of the Stanford Achievement Test taken by 3,112 elementary school pupils*

		White			Black			Asian		
		Grade			Grade			Grade		
Test		4	5	6	4	5	6	4	5	6
Computation	D	−.11*	−.08	−.27**	−.13**	−.29**	−.23**	−.41**	−.42**	.11
	F	.95	1.17	1.15	.93	.61	1.23	1.66	.69	1.00
Concepts	D	.17**	.21**	.21**	.09	.00	.02	−.33*	−.15	.16
	F	1.09	1.29*	1.28*	1.08	.63	1.12	1.73*	.63	.74
Applications	D	.18**	.18**	.16**	.00	−.10*	.10*	.32**	−.15	.28*
	F	1.12	1.43**	1.19	1.13	.69	1.03	1.94*	.85	1.00
N (Males)		269	274	280	218	219	226	36	44	40
N (Females)		264	223	278	212	216	191	43	43	36

Note: D = the difference between male and female means (M − F) divided by the average within-group standard deviation; F = the variance of males divided by the variance of females.
*$p < .05$, 2-tailed test.
**$p < .01$, 2-tailed test.

absence of a sex difference in mathematical reasoning in the present black sample? If the absence of a sex difference in the black group or the interaction of sex difference with ethnic group is explained strictly in terms of cultural factors, perhaps cultural factors also account for the sex difference in the white population. The observed interaction of the sex difference with ethnic group would seem quite problematic for Benbow's biological theorizing.

The male superiority in Arithmetic Concepts and Applications cannot be attributed to a sex difference in general intelligence. The Lorge–Thorndike Intelligence Test was given to all of these pupils, and in every ethnic group, at every grade level, the mean IQ of girls is about 3 IQ points *higher* than the mean IQ of boys – on both Verbal IQ *and* Nonverbal IQ. The overall male–female variance ratios are $F = 1.27$ for Verbal IQ and $F = 1.18$ for Nonverbal IQ (both F's significant at $p < .001$).

Most of Benbow's statistics are based on students who are highly selected for mathematical talent. What do we find when we select the top talent from the white group in the present study? If we select from above a cutoff 2 standrad deviations (*SDs*) above the overall mean of the total white distribution of combined scores on the Arithmetic Concepts and Applications tests, the ratio of the proportion of boys to the proportion of girls falling above the cutoff is approximately 4 to 1 (4% of boys and 1% of girls). If we select all pupils who score above a cutoff that is 2 *SDs above the mean for the boys' and the girls' distributions separately, so that approximately the same proportions of boys and girls are selected, the boys' mean turns out to be about two-thirds of a SD* higher than the girls' mean, which fully accords with Benbow's report of the effect size for the sex differences on tests of mathematical reasoning ability in groups of adolescents who are highly selected for mathematical talent.

Benbow's apparent hope for environmental remediation of the sex difference in mathematical aptitude seems to me a farfetched fantasy, if we believe that the hypothesis of a biological basis for the sex difference is correct. Remediation of a biological nature would raise the ethical question of the desirability of manipulating hormonal or other biological factors to achieve gender equality in high-level mathematical talent. Only an extremely small fraction of the male population displays the very high level of mathematical talent at which sex differences are marked. Therefore I would question the desirability, either to women or to society in general, of attempting to remedy a gender gap of such small consequence if the achievement of gender equality involved applying extreme or strenuous educational or biological interventions to any large segment of the population.

Biology: Si! Hard-wired ability: Maybe no

Douglas T. Kenrick

Department of Psychology, Arizona State University, Tempe, Ariz. 85287

It is interesting that Benbow opens her review by commenting on the problems of an ideologically based psychology of sex differences. Although times are changing, there are still those who believe that feminist ideology (or a particular sort of feminist ideology) should take precedence over data on gender differences. Advocates of this ideological approach have often used a two-front denial strategy. The first line is to deny that there are any substantial gender differences in behavior. This failing, the second line is to assume, a priori, that any differences that are demonstrated do not have a biological basis (e.g., Hubbard et al. 1982). Benbow does a fairly thorough job of vanquishing the denial of sex differences in math at the first front – it is hard to defend the position that there is no real gender difference here. She also makes a good attack against the second line of denial – the biological data are, if not overpowering, at least stronger than the environmental data.

Whatever one's feelings about a feminist-based approach to psychology, however, there seems cause for concern at the implications of this evidence – that mathematical problem-solving ability is biologically based. Would it not be better to suppress these data than to let the good ol' boys down at the legislature get the impression that women are innately inferior to males? I think not. There are two problems with the denial strategy, however instrumental it may seem at first. For one, it requires denial of other parts of the biologically based literature on sex differences, a literature that, in total, is hardly flattering to the good ol' boys. Along with any effects that male physiology may have on math and spatial ability, it seems to carry with it a wealth of handicaps. These include the learning disabilities and immune deficiencies discussed by Benbow, as well as a testosterone-fed proneness to violence (Rose et al. 1971) that, among other problems, leads males to be several times more likely to murder or to be murdered (U.S. Department of Justice, 1979; Wilson & Daly 1985). The other problem is that if we deny gender differences, we will not make any progress toward understanding, and thus toward remedying, those differences that displease us.

Benbow implies that the difference in math performance is due to a testosterone-induced modification in brain architecture. However, a look at the total picture emerging from biosocial research into sex differences suggests that the difference in math performance may not be due to a cognitive advantage for males at all. It may be a byproduct of other, much simpler gender differences in motivation and social behavior. Consider the clearest gender differences found to date – males are more aggressive, more oriented toward social-dominance games, and more physically active (see Eaton & Enns 1986; Kenrick 1987; for recent reviews relevant to these issues). Anyone who has seen Irven DeVore's fascinating film *The Baboon Troop* has watched the same pattern in our baboon cousins (see also Hall & DeVore 1965), and other primate research demonstrates the importance of testosterone in this masculine syndrome (e.g., Rose et al. 1971). Perhaps, then, the superior male performance on the SAT-M, and later in math-related professions, is simply another manifestation of the primate male's tendency toward hyperactive competitiveness.

Such an explanation fits with the finding that females do better than males under the more relaxed schedule of the formal math class, but worse under the time pressure of the SAT-M. In my own recollection, the quantitative sections of tests like the SAT were not so much more difficult than the verbal sections in content as they were more time pressured – and thus more likely to elicit a frantic race to finish before the bell. Even in grammar school, the boys would race to be the first to solve math problems, whereas the girls were more likely to quietly look on (even when they knew the answer). Performance on timed math exams may be like speed chess (a game that I have never seen a woman play, but which provides a peculiar obsession for several male colleagues). Thus, it may be that the SAT-M differences are simply a function of boys' higher motivation to perform on a timed competitive test.

Those higher levels of motivation could also account for the fact that the same boys go on to high levels of achievement in math and sciences. It seems unnecessary to presume that males' professional achievements are due to inherent abilities in those areas, since males' hyperactive dominance drives show up even in areas where there is no male ability advantage. In fact, males enter more competitively into the dominance hierarchies not only in physics and math, but also in art and literature, areas where female aptitude is at least on a par with that of males. In the math area, the dominance hierarchies are simply made relevant earlier because the tests draw out the male's affinity for "races." Consistent with my speculation here, Eccles (1985) reports reliably more persistent and single-minded pursuit of high levels of occupational achievement among males, despite the existence of only small differences in abilities.

Benbow may be able to rule out this hypothesis quite easily. If females taking math aptitude tests complete as many of the problems as males do, but get more incorrect answers, then a simple speed/competition hypothesis is wrong. The picture may be somewhat cloudier, though, because what begins as motivational differences may lead to different learning experiences for the two sexes (Anastasi 1985; Kenrick 1987), and eventually to differences in ability. However, it is important to keep in mind that even if the difference in math ability is rooted in "biology," that in no way limits the explanation to a sex difference in cognitive ability and/or brain architecture.

Biological influences on cognitive function

Doreen Kimura

Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5C2

The target article is a commendable review of the facts on mathematically precocious youth, but its ventures into the possible biological mechanisms underlying this special talent are often problematic, sometimes uninformed, and occasionally wanting in that scrupulous respect for evidence that has graced Benbow's psychometric research in the past. For example, her collection of data on the incidence of allergies in precocious students was inspired by Geschwind and Behan's (1982) claim of a higher incidence of immune disorders in left-handers. Unfortunately, it suffers from the same methodological weakness as the original – no information is provided for nonallergy (or nonimmune) disorders. We are unable to conclude from either study that the pathologies which apparently occur more often in their special populations are in fact limited to those disorders reported.

The probability that left-handers as a group have a less asymmetric brain organization is invoked to infer that "diffuse representation of cognitive functions . . . may be another possible biological correlate of extremely high mathematical reasoning ability." Yet there is absolutely no evidence that left-handers as a whole have superior mathematical ability, merely that among Benbow's highly select population (whether verbally *or* mathematically gifted), the incidence of left-handedness is greater than in a less select group. A similar fallacy is found in the claim that "dyslexics . . . have superior talents in certain areas of nonverbal skill." I know of no one (including Geschwind) who maintains that dyslexics as a group have better nonverbal skills than nondyslexics. Concerning left-handedness, we have recently collected evidence (among graduate and senior undergraduate students) that left-handers with high spatial ability (as well as excellent verbal ability) are *more* likely to show a right-ear superiority (implying left-hemisphere speech representation) than comparable left-handers with lower spatial ability (D'Amico & Kimura 1987). These data would not be readily compatible with an explanation that greater ability results from more diffuse brain organization. Benbow's own tachistoscopic data do not unequivocally support such a claim, as she herself acknowledges. The idea may be a reasonable one, but the evidence for it is weak or nonexistent.

The target article also makes reference to the possibility that the male superiority in mathematical reasoning may be related to the greater functional asymmetry of the male brain. Quite apart from the fact that this statement contradicts the claim that the mechanisms for superior ability might be more diffuse or bilateral organization, it has become clear that the concept of a more globally asymmetric brain organization in the male is an oversimplification. For example, the motor programming functions critical for speech and many manual activities have been found to be just as lateralized, and probably more focally organized, in females (see Kimura 1987). The superior manual dexterity and speech fluency of females may also be related to this more focal organization, but at present there is no convincing evidence directly linking the two.

Finally, I would like to comment on the possible role of sex hormones in brain organization and in cognitive makeup. To my knowledge, the first suggestion that fetal sex hormones might alter the nature of functional brain asymmetry, and thereby affect the intellectual pattern of the adult was made by Levy & Reid (1978). Nyborg (1983), without reference to brain asymmetry, later proposed that there may be an optimal level of sex hormones (specifically, estrogen) for superior spatial ability in humans. Intriguing as these and related hypotheses are, they have largely generated further speculation, and very little in the way of empirical verification. Nyborg's original study, showing improvement in spatial ability in Turner's Syndrome girls with short-term estrogen therapy, has not yet been replicated.

We have recently demonstrated in my laboratory that levels of sex hormones can influence certain abilities and, moreover, that these hormonal influences are *selective*, raising the level of some skills and lowering others. Elizabeth Hampson has shown, in women with normal menstrual cycles, that performance levels on a perceptuo-spatial task were reduced in the midluteal phase of the cycle (when levels of estrogen and progesterone are high), compared to the menstrual phase, when hormone levels are low. In contrast, performance on speeded manual and articulatory tasks was enhanced during the high estrogen/ progesterone phase, relative to the low phase (Hampson & Kimura, in press). A similar enhancement of motor skill has been shown during the estrogen phase of hormone-replacement therapy in a small group of postmenopausal women. Again, not all abilities were enhanced, and in fact there was a trend for a disembedding task to be impaired during the estrogen phase, compared to the no-hormone phase of therapy (Rosenthal & Kimura 1987).

These preliminary findings suggest that the influence of hormones on cognitive function will not be generally unidirectional, but may effect a reciprocal or trade-off relationship between certain abilities, at least within an individual. Thus the hormonal environment may modify the cognitive pattern within the limits of the genetic predisposition. Whether such reciprocal relationships reflect differential hemispheric processing or some other alteration of cerebral organization is still to be determined. In my opinion, what we most need now is intelligent *research* on these questions, to substantiate or refute some of the fertile speculation. An important area of research that has been largely ignored is the potential contribution of hormonal factors to ethnic differences in ability patterns (see Rushton, in press).

Creative mathematics: Do SAT-M sex effects matter?

Diana Eugenie Kornbrot

Psychology Division, The Hatfield Polytechnic, Hatfield, Hertfordshire AL 10 9AB, United Kingdom

Benbow's core finding is a sex difference in the performance of gifted adolescents on the mathematics section of the SAT-M. The results are impressive because they span a very large sample and are highly consistent over a period of 15 years, large in the life of a research worker, if not in the history of mathematics or feminism. Building on these core findings, Benbow draws several far-reaching conclusions. There are three in particular (all in the abstract) with which I would like to take issue: (a) that "differences are most pronounced at the highest levels of mathematical reasoning"; (b) that sex differences in test performance on the SAT-M result partly from biological factors; and (c) that sex differences in the performance of adolescents taking the SAT-M predict sex differences in later achievement in mathematics and science.

Differences increase with level of performance? Examining only the top 3% of 12 year olds in a talent search, the ratios of boys to girls for exceeding the SAT-M criteria of 500, 600, and 700 are 2.1, 4.1, and 12.9, respectively. Sensational stuff! But here is an alternative way of presenting the same results. The effect sizes at the 500, 600, and 700 SAT-M criteria are 0.25, 0.40, and 0.40, respectively. Because an effect size around 0.40 is about what is found in the general population, there is *no substantial evidence* that sex differences are larger at higher levels of performance. The presentation of group differences in terms of the ratio of group A:group B in attaining some criterion is inherently misleading since this ratio *always* increases with the severity of the criterion when the effect size is constant. A mathematical property of normal distributions is presented as if it were an empirical finding about sex differences.

Are sex differences in the SAT-M in part biological? The arguments in Section 7 suggest that sex differences in the SAT-M must be biological because the results of all studies of sex differences in superior mathematical performance favour males. Fraser and Cormack present a substantial counterexample to this generalisation (in press; also reported in Times Higher Educational Supplement, 16.1, 1987). They found no significant sex difference in the proportion of students obtaining first class honours in pure mathematics courses at Edinburgh University. Their results are particularly interesting because (a) the proportion of women students was 50%, so both sexes are equally select and (b) the performance level is high, with only 16.5% of an already highly select group achieving the criterion. Furthermore, an analysis of the results published by the University of London Examination Board, 1985, with more than 20,000 candidates in mathematics, shows no overall significant sex difference in the proportion achieving the top grade. Significant sex effects that do occur are on some of the more advanced elements and actually *favour women* (effect sizes of 0.16 for further mathematics syllabus B, and 0.79 for distinction in "special" paper of the "modern mathematics" SMP syllabus). These results may be contrasted with earlier studies showing a male advantage (Cockroft 1980).

The arguments Benbow presents in Sections 9 and 10 suggest that environmental mechanisms fail to account for the observed sex differences in cognitive performance. The evidence against *social* environmental factors such as differences in motivation and attitudes of the students and/or the surrounding society are highly convincing. The arguments against *cognitive* environmental factors, such as differential course-taking, are another matter. Evidence presented in sections 9.6 and 10.6 covers only mathematics, although the really big sex differences in course-taking are in the physical sciences. Several studies have shown that students who take science courses generally do better, even in degrees where the science content is not directly relevant (e.g., Fraser & Cormack, in press). Benbow herself admits that "some of the discussion on physiological correlates is speculative." Indeed, although the discussion concerns nice, firm, physiological data, the logic of many of the arguments is dubious. For example, all the physiological correlates turn out to be just as highly associated with high verbal, as with high mathematical ability, although the absence of adolescent gender effects in the SAT-V is used to argue for a sex-linked biological factor peculiar to mathematical ability. Furthermore, left-handedness is more common in males, allergies are equally common in both sexes, and myopia is more common in females! Most of the physiological correlates, like mathematical ability itself, unfold with development, so why any of them should be causes rather than consequences of superior mentation is unclear.

Adult achievement in mathematics and science. Clearly, sex is only one of many predictors of mature achievement. If there is an interaction between sex and other, perhaps larger, predictors that are not part of the analyses, then effects due to these variables may be erroneously attributed to sex. Without knowledge of these other factors and their interaction with sex, no valid conclusion can be drawn. For example, the association of parental education and SES (socioeconomic standing) with performance in the SMPY data is clearly very large (Benbow 1986a, Table 1), yet the effect of social variables has never been explored.

Sex differences favouring males appear in all prestigous occupations (Deaux 1985) and the social mechanisms responsible presumably also operate in scientific fields. Thus if adolescent SAT-M is responsible for some part of the sex differences in scientific achievement one would predict larger sex differences in the sciences than in verbal areas. Perhaps surprisingly, although the proportion of women in the hard sciences is small, they frequently perform better relative to men in physics and mathematics than in the humanities. Women tend to perform even better in biological sciences (e.g., Kornbrot, in press, for UK degree performance; Ferry 1982 for UK university academics).

Summary. The elimination of many plausible hypotheses about social environmental mechanisms is surely valuable for any understanding of scientific thought. Hypotheses that adult differences are either "part biological" or have any substantial component predicted by sex differences in adolescent SAT-M seem at best premature. Great effort has been devoted to elucidating small sex differences that could be biological. Meanwhile, large effects of social class have been ignored. Why?

Sex differences in mathematical reasoning ability: Causes, consequences, and variability

Brian Mackenzie

Department of Psychology, University of Tasmania, Hobart, Tasmania, Australia 7001

Benbow provides a valuable review of recent evidence showing a clear and substantial sex difference in mathematical reasoning ability in a very large sample of talented children, boys scoring consistently and substantially higher than girls. Her discussion and dismissal of various socialization hypotheses to account for the differences are persuasive if not definitive. Nonetheless, there are three major areas in which her account seems defective.

First, Benbow several times raises, but never deals with, the vital question of whether the sex differences in bright children are a result of sex differences in means or sex differences in variability. If the former, then average and dull boys would tend to do better than average and dull girls. If the latter, then dull boys, at least, would tend to do worse than dull girls. She notes that more boys than girls seem to have learning disabilities in mathematics, but later dismisses such distributional factors, writing that "even if one concludes that our findings result primarily from greater male variability, one must still explain why males are more variable." Of course one must: The point of considering the question of means versus variability is to decide what it is that must be explained. The most ingenious and elaborate hypothese to account for mean differences between boys and girls might be not only false but irrelevant, if it turns out that there are no mean differences but only (or mainly) a greater spread of scores for boys. Benbow explicitly acknowledges this fact. After spending exactly one-third of her paper discussing various socialization hypotheses for why girls do less well at mathematics, she concludes that "A possible reason why the various classical socialization hypotheses do not help . . . is that those hypotheses were formulated to explain differences between means. [Our] sex differences data may be the result of greater male variability." Having thus raised the question of whether her long discussion was in effect a waste of space, she drops it.

Second, Benbow assures the reader that the sex differences in mathematical ability she describes have clear consequences beyond high school, where they were measured. Her attempts are unconvincing. She finds significant sex differences in choice of major subjects at college among mathematically talented youth, with 63% of boys versus 35% of girls majoring in mathematics or science, and similar differences in graduate school. But to get these differences she had to combine the figures for mathematics with those for physics, engineering, computer science, and so forth. There were no substantial sex differences in enrollments in mathematics departments. There may be many reasons why girls are less likely to study physics or engineering than boys, but lesser interest or ability in mathematics is not likely to be a major one if girls are no less likely than boys to study mathematics itself. To make her case about the long-term consequences of the ability differences measured in grades 7 and 12, Benbow needs data on differences in performance in university mathematics courses or in the quality of Ph.D. dissertations and the like. In the absence of such data, she does not have a basis for her claim that the test score differences found in high school have important consequences later.

Third, despairing of socialization explanations, Benbow offers some physiological speculations about the possible source of these sex differences. It is certainly to her credit that she tries to identify possible mechanisms rather than merely concluding that "it must be in the genes." Unfortunately, her speculations are inconsistent and self-contradictory. In discussing laterality, she tentatively offers the following points: (a) left-handed people seem to have an advantage at right-brain tasks; (b) mathematical reasoning (as distinct from computation) seems to be largely a right-brain activity; (c) in her elite sample of 340 exceptionally bright children there was twice the incidence of left-handedness as in the population at large, and 50% more left-handedness than in the sample's siblings and parents; and (d) in the same sample, boys were significantly more likely to be left-handed than girls. The conclusion follows neatly: Left-handedness (perhaps indicative of "bilateral or diffuse representation of cognitive functions"), mathematical reasoning ability, and being born male all go together. It is made even neater by adding a speculation by Geschwind and Behan (1982) that fetal testosterone, produced in the male fetus's testes, enhances right-brain development.

The only jarring note is Benbow's finding that, in her elite sample, the greatest incidence of all of left-handedness (three times the population incidence) was found in the boys with high verbal rather than high mathematical achievement. She accommodates this finding by proposing that verbal reasoning, as distinct from verbal fluency or vocabulary size, is itself largely a right-brain ability. This proposition makes the whole argument rather shaky, however, since her repeated finding is that there are no sex differences (in bright children) in verbal ability. If verbal reasoning ability, too, is concentrated in the right-brain, and boys are more likely to have highly developed right-brains (incidentally producing more left-handedness) due to fetal testosterone syndrome, boys should surely surpass girls in verbal ability as well as mathematical. But they don't.

Furthermore, even the greater tendency to left-handedness in males in Benbow's elite sample is questionable. Either the figures she gives, or the calculations she makes, are in error. She reports that in the sample of 340 children, 11.4% of the girls and 16.4% of the boys were left-handed. However, with a total sample of 340, the greatest possible value for the z for the difference between proportions is 1.29 (assuming equal girl and boy Ns; unequal Ns would make the z value still smaller). This value falls far short of the .05 significance level she claims (a z of 1.96 is required, two tailed).

In conclusion, these three criticisms seem to cover the major components of the case Benbow tries to make. They are concerned with (a) what the sex differences consist in, (b) whether or not they are important beyond high school, and (c) what might be causing them. I hope Benbow can answer them, either now or subsequently. The task of understanding sex differences in cognitive functioning is sufficiently difficult, and sufficiently important, that we must value highly every approach that leads to progress in the task. Benbow and her colleagues at the Study of Mathematically Precocious Youth have developed a productive and versatile approach to investigating the area, and future output from the study can only be eagerly awaited.

What we really need is a theory of mathematical ability

Richard E. Mayer

Department of Psychology, University of California, Santa Barbara, Calif. 93106

Benbow's description of the Study of Mathematically Precocious Youth (SMPY) is a classic example of research data in search of a theory. The major to-be-explained finding is that 12- and 13-year-old boys selected on the basis of extreme mathematical talent averaged higher scores on the mathematics section of the Scholastic Aptitude Test (SAT-M) than corresponding girls. What is needed, of course, is a theory of mathematical ability (Ginsburg 1983; Mayer 1985; Resnick & Ford 1981; Schoenfeld 1985). Before we can adequately explain the SMPY finding of sex differences in mathematical ability, we need to know what mathematical ability is and how to measure it.

The need for a theory of mathematical ability is exemplified by an intriguing contradiction in the SMPY data: SMPY boys outperformed the SMPY girls on the SAT-M but the girls outperformed the boys on school mathematics grades. Benbow's target article seems based on the premise that the SAT-M scale is the best measure of mathematical ability. This key assumption can be challenged on the grounds that other measures (such as mathematics grades) may be more representative and less biased indicators. First, the SAT-M scale is based on responses to 60 problems produced over the course of 60 minutes; high school mathematics grades are likely to be based on thousands of responses produced over hundreds of hours. Second, the SAT-M scale, when given to adolescents, covers mathematical problems requiring knowledge that must be acquired outside of school; because we cannot insure that boys and girls in the SMPY sample had equal access to nonschool mathematics-based experiences, the SAT-M could be biased in favor of subgroups that received more nonschool mathematics experience. In contrast, mathematics grades represent students' ability to learn mathematics in a situation where boys and girls receive equivalent exposure to the material. If we accept school grades as a measure of mathematical ability – namely, the ability to learn mathematics – we seem to turn Benbow's argument on its head; now we must explain why girls outperform boys! A sound theory of mathematical ability is needed in order to help select a meaningful measure of mathematical ability.

The remainder of this commentary explores two issues that I have discussed in more detail elsewhere (Mayer 1983): Are there sex differences in mathematical ability? If so, why are there differences?

Are there sex differences? Benbow asserts: "Studies of mathematical ability and achievement have consistently found sex differences favoring males." However, the references cited and Benbow's own data appear to contradict this assertion in at least three ways.

Sex differences depend on age. First, Maccoby and Jacklin's (1974) classic book is cited as support for the claim that there are consistent sex differences favoring males. However, in their review of about two dozen studies, Maccoby and Jacklin found that for young children (ages 3 to 8) there were either no differences or girls averaged better than boys, for children in the 9- to 12-year-old range there were either no differences or boys averaged better than girls, and for children in the 13- to 21-year-old range boys tended to average higher scores than girls in most studies. A similar pattern has been obtained in state-wide assessments of mathematical performance (Mayer 1983).

Sex differences are not always strong. Second, of the four studies listed by Maccoby and Jacklin (1974) involving very large samples of children above age 12, only one yielded effect sizes that Benbow would classify as more than "small" – data from an early sample of SMPY students. Similarly, state assessments of mathematical performance reveal sex differences of one or two points (Mayer 1983).

Sex differences are not present for all math tests. A final contradiction to the claim for consistent sex differences favoring boys is Benbow's observation that "SMPY females received better grades in their high school mathematics courses than did SMPY males." To the extent that school mathematics grades measure mathematics achievement and ability, there *fails* to be consistent evidence.

In summary, the data do not support the assertion that there are consistent sex differences in mathematical ability. However, there is support for a more limited assertion: For adolescents in the SMPY data base, average performance on the SAT-M was better for boys than for girls. The data in Table 1 provide overwhelming support for this limited assertion, and there is certainly no need for any further replication.

Why are there differences? If we accept the above limited assertion, how can we explain it? Of the three kinds of explanations suggested by Benbow – socialization, environmental (or experiential), and physiological (or biological) – the latter appears to the most innovative.

Biological explanation. Benbow offers the conclusion that "biological factors seem to be involved in determining the sex difference in mathematical reasoning ability." As support, she examines four physiological correlates of mathematical ability: handedness, allergies, hormonal levels, and myopia. A convincing argument involving any of these correlates would need to be based on the following evidence: (a) the physiological characteristic is over-represented in the SMPY sample as compared to the general population, (b) within the SMPY sample the physiological characteristic is over-represented in males as compared to females, and (c) the characteristic is not related to non-mathematical ability such as verbal ability.

Unfortunately, the SMPY data fail to meet the three criteria necessary to support a biological explanation of sex differences in mathematical ability. For handedness, Benbow presents evidence for (a) and (b) but not (c) – that is, left-handedness is over-represented in high mathematics and in boys but is also over-represented in high verbal students. For allergies and myopia, Benbow presents evidence for (a) but not (b) or (c) – that is, incidence of allergies and myopia is overrepresented both in high mathematics and high verbal students and no sex differences are reported. For hormonal level, no SMPY data appear to be reported. Although males and females (or high and normal ability students) may differ in hormonal levels or brain structure or some other physiological characteristic, it does not follow that these differences cause sex differences in behavior specific to mathematical ability.

Conclusion. We are left as we began this commentary – with some interesting data in search of a theory. Although Benbow's target article fails to provide convincing empirical evidence for the biological explanation, and although such a theory might be distasteful to some people, the proper test of a theory lies in careful study and research data rather than one's own biases. Additional research is needed to assess Benbow's conclusion that sex differences in mathematical ability "result from both environmental and biological factors." Benbow's paper will be useful and fruitful to the extent that it stimulates a deeper understanding of the nature of mathematical ability.

Socialization versus biology: Time to move on

Diane McGuinness

Department of Psychology, University of South Florida, Ft. Myers, Fla. 33907

Despite the overwhelming evidence that Benbow and Stanley have been providing for the past 16 years on the sex differences in mathematical ability, most people, including psychologists working in the field, refuse to believe it. Or if it is "true," as some now must grudgingly admit, it must be due to a set of socialization practices by which females are discouraged and males encouraged to excel in mathematics. Benbow devotes a large portion of her target article to a review of the various hypotheses by which such socialization practices could influence actual mathematical talent.

I have several reactions to this. First, it is unfortunate that Benbow is obliged by the current Zeitgeist to expend so much energy in this futile endeavor, as she concludes that socialization hypotheses not only fail to explain the sex differences in the normal population, but also in her own population of talented students. She is also almost too kind and too patient, because she does not really address the logical fallacy of a socialization perspective. That is, how could it be possible for society to encourage the aptitude of females in arithmetic, where they often show superior skills to males, but to discourage female performance in mathematical reasoning or higher mathematics? Furthermore, why would society want girls to do badly in math, but not in other school subjects like foreign languages, art, and biology?

Second, Benbow does not point out the central problem with all of the research on the socialization model, and that is, the design of the research is entirely correlational and can therefore tell us nothing about the direction of causality. Do attitudes to math cause poor performance, or does poor performance cause the attitudes? Furthermore, as Benbow does indicate in her target article, the research on attitudes, confidence, parental involvement, and so forth often fails even to carry out correlations at all, assuming that attitudes cause mathematical performance while neglecting entirely to measure that performance.

The third problem that arises from being forced to address the Zeitgeist, not only in the tedious amount of reading required as well as the experiments Benbow herself has had to carry out, is that it takes up so much valuable time from the real issues, which Benbow outlines in section 3, paragraph 6. These are the issues concerning the nature of mathematical ability per se. It is perhaps our area of greatest ignorance in cognitive psychology, and one looks forward to the future article by Benbow in which she begins to review the results of her work on mathematical reasoning.

Last, the notion that socialization can explain everything not only produces a diversion from real issues, such as "what is it that boys are doing that creates their advantage?", but can also lead to equally zealous and time-wasting attempts to demonstrate a biological basis for cognitive skills, when the socialization arguments are found to be in error. One does not need to locate the sex difference on a particular gene, or within a hemisphere, to make a valid argument for the biological basis of these sex differences. This argument is made quite cogently by Benbow's review of the cross-cultural data, showing that even in China (where everyone is supposed to be good at math!) the boys are better than the girls.

In fact, we are wasting an inordinate amount of time on why-questions, because people find sex differences in cognitive functioning to be emotionally unsettling. But how can we begin to answer why-questions until we have the answers to what-questions: Such as, "What is mathematics, anyway?" Until we understand mathematical reasoning and are able to pin down which brain systems function in which mathematical processes, we are premature in hypothesizing about where mathematics is "located" in male and female brains.

The central fact is that the sex differences are just not going to go away. As Benbow points out, they have, if anything, increased during the years when feminists were pushing for greater equality of opportunity. To get bogged down in a nature/nurture dispute does not really seem productive in view of the fact that millions of our children keep going through the school system, and no one has the faintest idea how to teach mathematics to boys, much less to girls. Even if we discovered the "truth" about these sex differences, whether it was socialization or biology wouldn't make the slightest bit of difference. Not only is it extremely difficult for society to rearrange its stereotypes, but locating mathematics on a gene in the Y chromosome would hardly help us either. Furthermore, biology is certainly not destiny. No biological function, apart from a few primitive reflexes, is immutable.

Thus I feel that spending time on Geschwind's ingenious, but perhaps misguided, theory on the connection between handedness, allergies, testosterone, and sex differences in cognitive function, is again counterproductive. First, the kind of lateral thinking that goes on in the clinic, where the good clinician begins to detect patterns of connections, is a great starting place for some major insight in clinical diagnosis, but it may not serve too well in solving the mystery of mathematical talent. Second, too many "links" in the chain of coincidence are ultimately messy. Would all allergic, left-handed males who are myopic be the world's greatest mathematicians? What of the other 85% of right-handed students in the Talent Search population who are gifted in mathematics? How does Geschwind's theory explain them? And if all of these factors or coincidences lead back to testosterone, we already knew that in the first place, because of the male advantage in mathematics. We still do not know what mathematics is about, however, nor what it is that males are doing to make them good at it.

As a final comment, one paragraph from the paper is disturbing, suggesting that Benbow might be beginning to overreact to the attacks from the socializationists by seeming prone to the belief that "biology is destiny." In section 6, she cites an interesting study by Senk and Usiskin (1985) in which the sex differences in geometry were eliminated by extensive training in how to formulate proofs. Benbow dismisses this finding with the following words: "It should be noted, however, that this measure was testing the ability to develop geometric proofs after prolonged instruction. Thus, the ability measured by Senk and Usiskin is very different from SMPY's mathematical reasoning ability." Benbow writes as if she had a vested interest in rejecting this result. Why isn't formulating proofs a reasoning process, given that mathematical reasoning is so hard to define, as Benbow points out? I am also concerned about the word "however," used as if a training program that eliminated sex differences were somehow a fluke, or due to a semantic error in the use of the word "reasoning" rather than as an exciting result to be followed up. After all, isn't the goal to be able to train everyone to enjoy and excel in mathematics?

Rival hypotheses about sex differences in mathematics: Problems and possibilities

Carol J. Mills

Center for the Advancement of Academically Talented Youth, The Johns Hopkins University, Baltimore, Md. 21218

The Benbow target article provides an important discussion of the stable finding of sex differences in SAT-M scores in a population of high-ability adolescents (12 year olds). The summary of the possible correlates or causes of the sex difference in exceptional mathematical reasoning ability, however, fails to discuss adequately the methodological problems and alternate interpretations endemic to many of the key studies. Nor does this paper help the reader to form the necessary connections that would help us understand possible causal pathways.

Sampling bias. Although Benbow states that "systematic sampling bias in the Talent Search data seemed unlikely," the fact is that numerous factors are quite likely to influence who takes the SAT at 12 years of age and who does not. For example:

1. School personnel decide which students (if any) in their school are given information and an application for the Talent Search.

2. Parents' decision to have their child participate in the Talent Search is probably influenced by such factors as: cost, value and importance of a child taking such a test at that age, and desire to have their child take accelerated courses.

3. A student's decision to participate in the Talent Search is probably influenced by self-perceptions of abilities, academic goals and interests, as well as a host of other conflicting interests. It is possible that fewer girls with high mathematical ability than comparable boys are electing to take the SAT because of less commitment to advancing their mathematics education outside of school, a fear of being seen by their peers in a negative light, or simply less confidence in their own ability (Reis 1987).

Since the taking of the SAT in the seventh grade (compared to the more typical undertaking in high school) is a much more selective venture with less obvious value and more risk involved, there is no way to be sure at this point whether the Talent Search population is a representative sample of mathematical reasoning ability in the two sexes. The statement that the boys and girls are "matched" is misleading. They simply enter through a common portal loosely defined by age and grade, and a variety of acceptable test scores. Until a myriad of uncontrolled variables are examined more carefully, sampling bias is still a valid hypothesis for at least partially explaining the smaller number of girls achieving higher scores on the SAT-M.

Environmental correlates. The reader should be cautious in accepting Benbow's summary statement that no sex differences were found for the majority of the environmental/social hypotheses in the SMPY population. Many of the studies reported suffer from one or more of the following problems:

1. *The group being tested was inappropriate.* Is there any reason to expect that girls and boys with equally high SAT-M scores (over 700) would differ in their attitudes about mathematics? And yet, this is the very group most often studied when Benbow refers to the "SMPY population." If one wants to study environmental factors that may be suppressing the number of girls with exceptional mathematical reasoning ability, a more appropriate comparison would be between high- and low-scoring SAT-M girls.

2. *The test measures used were questionable.* In most cases, when studying self-concept, interests, attitudes, or personality, the staff of SMPY used a few select items (sometimes as few as one item to measure a construct, see Raymond & Benbow 1986) either taken from other instruments or constructed by SMPY staff. In a study using several established personality measures with the Talent Search participants from 1976, I found support for a positive relationship between the possession of cross-sex personality characteristics and measures of intellectual ability, particularly for girls (Mills 1981).

3. *Test bias was introduced by the identification process and subsequent testing situation.* SMPY students have been singled out (and honored) for their mathematical reasoning ability by the very organization now asking them about their attitudes, behaviors, and interest in mathematics. Response biases such as social desirability are quite likely to be operative in such a situation (Anastasi 1982).

4. *Retrospective studies asking parents or students about*

past behaviors or preferences (see Benbow 1986a) are subject to response bias, distortion, and selective remembering. The fact that the Fox et al. (1982; 1985) studies looking at SMPY populations did find sex differences in some environmental variables (e.g., confidence in mathematical ability) is quite remarkable, and should be investigated further.

Possible connections and other factors. It has been suggested that spatial ability could account for the "edge" that boys have at the higher levels of the SAT-M (Stanley 1982). Indeed, Benbow uses the spatial ability hypothesis as a pivotal point and a transition from her discussion of environmental influences to that of biological correlates. Although she emphasizes the biological evidence for sex differences in spatial ability, she fails to point out the evidence for an environmental influence, which in fact is as strong as the biological. It is even possible that the environmental influence on mathematical reasoning ability is *through* its influence on spatial ability.

Another possible correlate of intellectual ability is cognitive style, defined as "preferred or habitual patterns of mental functioning; a disposition to seek out learning environments compatible with one's attitudes and interests" (Lawrence 1984). Using the *Myers-Briggs Type Indicator* with 1983 and 1984 Talent Search participants, I found sex differences on two dimensions of "cognitive style" (Mills 1985). These same dimensions differentiated the verbally gifted from the mathematically gifted students, and significant differences in SAT-M scores were found for the girls with differing cognitive styles (Mills 1984).

As important as it is to investigate the determinants of the sex differences in exceptionally high mathematical reasoning ability (both environmental *and* biological), it is equally important to begin looking at what happens to the girls who *do* score high on the SAT-M. What distinguishes high-scoring girls who have high aspirations and achievement levels from high-scoring girls who do not? In addition, understanding the social/environmental differences in high- and low-scoring groups of girls may help us to understand the ability differences that occur between boys and girls.

Mathematics as male pathology

John Money

Professor of Medical Psychology and Pediatrics Emeritus, Departments of Psychiatry and Behavioral Sciences and of Pediatrics, The Johns Hopkins Hospital, Baltimore, Md. 21205

Psychology and psychiatry are divided between the proponents of mindlessness and the proponents of brainlessness (Eisenberg 1986). Under the impact of psychoanalysis, brainlessness held sway until, in very recent years, the rapid growth of neuroscience gave new glamor to brain research. It is now fashionable for psychologists to make conjectural leaps from cognitive and intellectual function to brain function. One such leap is from male-female differences in cognitive and intellectual performance to right and left cerebral hemispheric functioning. Benbow's target article addresses itself to this leap with respect to superiority in mathematical performance.

The initial findings of the SMPY (Study of Mathematically Precocious Youth) confirmed an existing phenomenon, namely, that in high school and subsequently, mathematical giftedness is more prevalent among boys than girls, whereas in the earlier years both sexes are either equal or girls are superior. Good phenomenological recording is a prerequisite of good science, but causal explanation is the prize and incentive that sustains scientific investigation. Without a new causal hypothesis, an investigator falls back on an old one, in this instance one that is not only old, but outmoded, namely, the causal dichotomy of nature versus nurture, of biology versus learning.

Benbow reviews research designed to test an unsystematic and more or less haphazard set of social-environmental doctrines that have been popularly used to explain the male–female difference in mathematical test scores. Insidiously, the difference is not stated as a greater prevalence of high scores among males, but as male superiority in mathematical reasoning or mathematical reasoning ability. One reads the entire text without obtaining any idea of what constitutes mathematical reasoning ability other than that it is the mathematics score on a test, for example, the Scholastic Aptitude Test. One might condone such terminological imprecision under certain circumstances, but not in the context of attempting to relate mathematical reasoning to differential lateral dominance of the cerebral hemispheres in boys and girls.

Even when the evidence for or against a particular doctrine of social-environmental determinism is inconclusive, Benbow finds the doctrine wanting and favors its rejection. This paves the way for the eventual endorsement of biological determinism as the cause of male superiority in mathematics. More precisely, it paves the way for a very skewed and outmoded idea of what constitutes biological determinism in psychology. There is a biology of learning and remembering. It exists in the brain. The fact that what is learned and remembered is programmed through the senses into the brain neonatally and postnatally does not necessarily make it any less powerful and permanent than that which is programmed into the brain phylogenetically and prenatally. One must always pay attention to the lesson of the psychology of native language: It is not present in the brain on the day of birth, but once it gets programmed into the brain it cannot be deprogrammed except by surgical extirpation or by the trauma of a brain injury. Individual variations as well as male–female differences exist in native language phenomena. There is no a priori reason to reject the possibility of a mathematical parallel.

Native language demonstrates the invalidity of juxtaposing nature versus nurture, and the necessity of following geneticists and embryologists by substituting the two-term paradigm with a new, three-term one, namely, nature/critical period/nurture. To illustrate: Nature, as represented by the genetic code, cannot unfold itself except in a hospitable environment, either before birth or later, and then it must do so at a preordained critical period, after which the resultant product becomes permanent. Psychology learns from ethology that the same principle applies in psychological development, and that the resultant product becomes imprinted in the brain. If not directly, then circuitously, this principle may apply to mathematical development, regardless of the stage of development, early in infancy or later, that constitutes the critical period.

Having given up on doctrines of social-environmental determinism to explain the differential sex ratio in mathematical test scores, otherwise known as mathematical reasoning ability, Benbow turns to three possible biological correlates, namely a differential sex ratio in left-handedness, allergy, and myopia. The rationale for selecting these three correlates is theoretically convoluted, and derives from extreme speculations, first published, with associates, by the neurologist Norman Geschwind in 1982, shortly before he died.

Geschwind had a professional lifetime's interest in learning disability and dyslexia, which have long been known to be more prevalent in boys than in girls, and to be associated with left-handedness more than right-handedness. Finding also an additional association of learning disability with migraine headache, immune disease (chiefly allergy and asthma), hair color, and vision, Geschwind formulated a grand theoretical synthesis that tied together handedness, brain laterality, prenatal hormonal effects of melatonin on hair, skin, and the nervous system, prenatal effects of testosterone on the differentiation of sexual dimorphism in the brain, and prenatal effects of testosterone on the immune system by way of the fetal thymus gland.

It is an ingenious theory, but it raises more questions than it answers about learning disability, and even more questions when applied to mathematical superiority. Perhaps because it endangers their theorizing, neither Geschwind nor Benbow have referred to a considerable body of clinical psychoendocrine and achievement data gathered developmentally from childhood to adulthood from people with a known prenatal history of hormonal deficiency or excess – much of it gathered over the last 35 years at Benbow's own former university, Johns Hopkins.

Theorizing aside, what Benbow is left with is a finding that children (both sexes, apparently) selected for high mathematical test scores are more likely than other children to be left-handed; to be allergic, chiefly with hayfever; and to be myopic. The biology of the relationship of these three phenomena to mathematical reasoning remains a complete mystery, as does also the sex ratio in favor of boys as high achievers in mathematics. Likewise, if the relationship between all four factors should eventually prove to be socioenvironmentally mediated, that also remains a mystery, but it cannot be totally dismissed in the present era of the new sciences of psychoneuroendocrinology and psychoneuroimmunology. [See Engel: "An essay on circulation as behavior." *BBS* (9)2.] In both sciences, cause and effect work both ways. Ironically, after having struggled to prove otherwise, Benbow herself more or less capitulates and accepts this two-way wisdom when, in the last sentence of her abstract, she writes, "It is therefore proposed that the sex difference in SAT-M scores among intellectually talented students, which may be related to greater male variability, results from both environmental and biological factors."

Nature/nurture in male/female mathematical giftedness

Nora Newcombe and MaryAnn Baenninger
Department of Psychology, Temple University, Philadelphia, Pa. 19122

We would like to suggest that Benbow's target article is flawed in at least two ways. First, it fails to justify the suggestion that the mathematically gifted require study as a special group, and that conclusions drawn from studying normal children do not apply to the gifted. Small differences between the sexes in mathematical ability, on the average, with much larger differences at the positive end of the distribution, suggest continuous, normal or near-normal, distributions of mathematical ability for both males and females. Benbow seems to endorse such a model, but fails to notice that it conflicts with her repeated assertion that the mathematically gifted need to be studied as a separate group, a strategy that has the clear disadvantage of restricted range and ceiling effects on many variables, leading to null results or weak effects.

If mathematical giftedness *is* produced by the operation of special variables, the observed distribution of mathematical ability is really a mixture of two different normal distributions. This hypothesis should be evaluated statistically. Even if the mixture-of-normals hypothesis is supported, Benbow should evaluate problems with ceilings and restriction of range in her gifted sample quantitatively, and correct for them if necessary, rather than merely assert that the problems do not exist.

A second problem in Benbow's article is that she is considerably harder on the environmental than the biological hypotheses. She emphasizes the inconsistencies in the data on environmental hypotheses, ignoring the fact that there are few if any reversals of the direction of effects, but merely some null findings. Within a meta-analytic perspective, this is clearly compatible with overall support for the hypotheses (e.g., Rosenthal & Rosnow 1985). Signorella and Jamison (1986) used meta-analysis to conclude, for instance, that masculinity and femininity *are* correlated with mathematical ability. Furthermore, Benbow often fails even to mention null findings relevant to her biological hypotheses; for instance, Bornstein (1984) and Herring and Reitan (1986) both failed to find the sex differences in the effect of unilateral cerebral lesions emphasized by McGlone (1980).

Also, in evaluating environmental hypotheses, Benbow consistently emphasizes that variables should be correlated with ability *within* ability groups. By contrast, biological variables are evaluated solely in terms of mean differences between the talented groups and controlled. For instance, sex differences in lateralization are discussed without attention to the issue of whether indices of lateralization are correlated with mathematical ability, within sex or ability level. For spatial ability, correlations with lateralization within sex are rare (Newcombe 1982), and may vary with sex and ability level (Ray et al. 1981). As another example, Benbow points out that left-handedness and allergies are twice as likely in her gifted groups as in comparison groups, without presenting correlations of ability with handedness or allergies within ability groups. By contrast, for environmental variables, Raymond and Benbow (1986) emphasize within-group analyses, downplaying the fact that, for instance, paternal encouragement of quantitative interests was almost twice as common among mathematically gifted children as in comparison groups who were themselves far above the average in quantitative skill. In addition, much of the SMPY work evaluating environmental hypotheses has involved looking for sex differences on environmental variables within ability groups, that is, after matching for ability. It is hard to see why sex differences would be expected in this case.

Recent work on sex differences in spatial ability. As Benbow notes, there is a vast literature on sex differences in spatial ability. We would like to update *BBS* readers on two aspects of this literature. First, Waber's (1977) hypothesis that late maturing is associated with higher spatial ability has become widely known, but efforts at replication have had variable results. A meta-analysis by Newcombe and Dubas (1987) showed that the association was, at best, small, and that its significance was dubious given that only a handful of "filedrawer" studies with null results would be needed to render it nonsignificant. Second, efforts to pursue environmental hypotheses for the sex difference in spatial ability have recently been summarized by Signorella and Jamison (1986) and Moore et al. (1987).

In conclusion, we would like to say that we are not radical environmentalists. Theorists who favor environmental hypotheses often ignore or denigrate biological data, and this is clearly an error. But the symmetric point is that biologically inclined investigators often dismiss environmental data and demand less of biological data. We believe Benbow's article shows this bias.

Mathematics, sex hormones, and brain function

Helmuth Nyborg
International Research Center for Psychoneuroendocrinology, Psychology Institute, Aarhus University, DK-8240 Risskov, Denmark

Benbow's proposal that prenatal testosterone exposure may in part determine extremely high mathematical reasoning ability is suggestive, but estradiol exposure seems a better bet (Nyborg 1979; 1983), because it has more direct organizational and activational effects on sensitive brain tissues (see Toran-Allerand 1986).

Estradiol may actually explain many of the relations among sex hormones, brain growth, spatial ability, and mathematics in humans. Thus, women with Turner's syndrome (Turner 1938) are deficient in sex hormones and have severe problems in spatial ability tasks and mathematics (e.g., Nielsen et al. 1977). However, Turner women who have received about 1 year of cyclic estrogen/gestagen treatment perform at the same level as

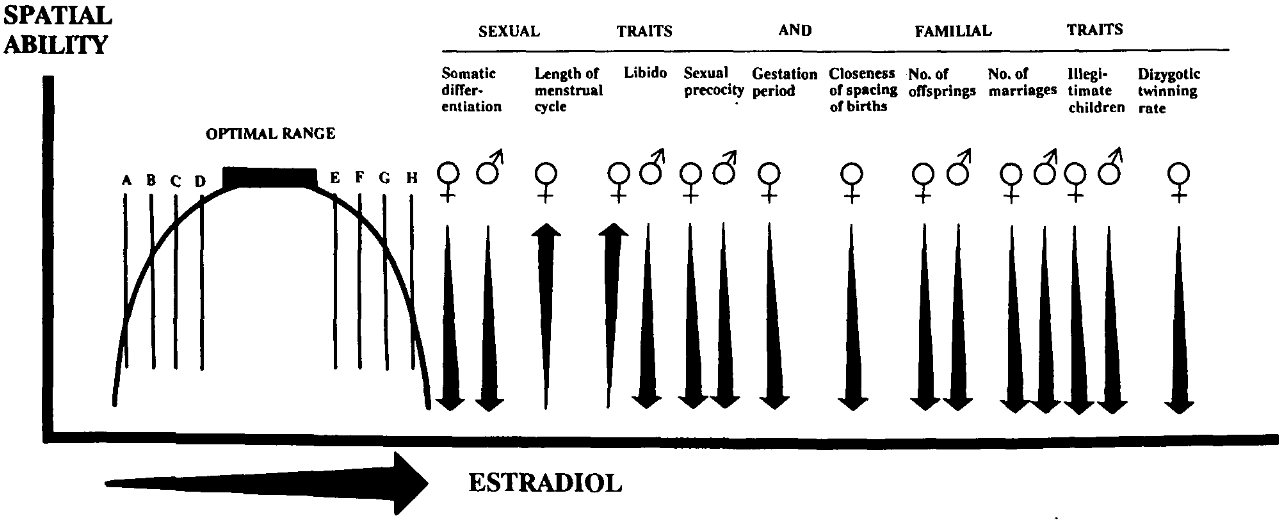

Figure 1 (Nyborg). The General Trait Covariance-Androgen/Estrogen (GTC-A/E) balance model for the effects of variation in gonadal hormones on body, brain, and behavioral development in masculinized (A-D) and feminized (E-H) individuals (from Nyborg 1987; for explanation, see text).

their age-matched sisters on various tests for spatial ability as well as in mathematics (Nyborg & Nielsen 1981). These and other findings have led me to propose (Nyborg 1979; 1983; 1984; 1986) a curvilinear model for the effects of estradiol concentrations on body, brain, and spatial ability development as illustrated in Figure 1. According to this model, normal females (E-H) typically overshoot the range of estradiol values for the full expression of spatial ability at puberty. This explains the common female teenage regression in spatial ability and mathematical achievement (Nyborg 1983). A tilt in the sex hormone balance toward testosterone will slightly masculinize the female body and keep the brain concentration of estradiol within the range for the optimal expression of spatial ability. In this way, the model explains why late maturing, androgynous females tend to show higher spatial ability than do their more feminized counterparts. Normal males (A-D) undershoot the optimal range at puberty. A tilt toward testosterone will further masculinize the body and depress the expression of spatial ability (for details, see Nyborg 1983). Moderate estrogenization will feminize the body and lead to the full expression of spatial ability. This could explain why spatial ability tends to be higher in androgynous men than in very masculine men (e.g., Maccoby & Jacklin 1974). It has further been found that women show high spatial ability during low estrogen phases of the menstrual cycle, and low spatial ability during high estrogen phases (see Nyborg 1983). My model has since been elaborated to incorporate covariant intellectual and personality development and is now referred to as the *general trait covariance androgen/estrogen* (GTC-A/E) model (Nyborg 1984; 1987; submitted).

The GTC-A/E model allows for testable predictions about the sexual development of mathematically eminent people. These can, for example, be expected to show a moderate surge in sex-related gonadal steroid plasma concentrations at puberty, and a low degree of secondary sexual differentiation. The male mathematician will be either tall and slender or pyknic, will have a low muscle content and accordingly show decreased muscular strength, will be long-lived, and will have been a sissy as a child. The GTC-A/E model further predicts that the female mathematician will have a low body-fat ratio, will be tall, slender, and strong, and will reflect a childhood history of tomboyism. Mathematically eminent people go into puberty late and show a prolonged period of brain development. They are typically first-born, come from a family with few children, and have, themselves, few offspring (in particular, few dizygotic twins). They show reduced physical aggressiveness, high behavioral restraint, introversion, and prefer abstractions and objects to people. They prefer controlled political development and appreciate a formal to a loose social organization of society (Nyborg 1987). Let me, therefore, suggest that Benbow initiate a person-specific search for the mathematically eminent, unrestrained he-man and for the opulent, very fertile, extroverted female mathematician in her large populations. The finding of more than a few such "Black Swans" would falsify my GTC-A/E model.

Evaluating explanations of sex differences in mathematical reasoning scores

Robert Rosenthal
Department of Psychology, Harvard University, Cambridge, Mass. 02138

Benbow's scholarly target article is thoroughly compelling in its documentation of sex differences in SAT-M scores. In addition, it seems impossible that anyone could quarrel with her modest conclusion that both social and biological factors may contribute to the difference, each to an as yet unknown degree.

The major goal of this commentary is to enter a plea. That plea is that future reviews of social and biological correlates of both sex and mathematical reasoning scores go beyond the narrative listing of studies finding or not finding that there is a nonzero correlation. The developmental stage of the behavioral sciences is now such that reviews of the literature should give readers an estimate of the typical effect size found (Cooper 1984; Glass et al. 1981; Hedges & Olkin 1985; Hunter et al. 1982; Light & Pillemer 1984; Mullen & Rosenthal 1985; Rosenthal 1980; 1984; Wolf 1986). [See also Prioleau et al. "An Analysis of Psychotherapy Versus Placebo Studies." *BBS*:(6)2 1983.]

Thus, while we know that there is a typical sex difference of about a half standard deviation favoring males among the intellectually talented youngsters studied by Benbow and her colleagues, we have no idea of the typical magnitude of the effect of the social or biological variables narratively described in the target article. Are these effects, even if significant statistically, so small that they cannot begin to account for the sex difference in mathematical reasoning scores? Are they large enough to swamp the one-half standard deviation effect (which is equivalent to a Pearson r of .24)?

If we assume a chromosomal definition of sex for the moment, then sex is an antecedent variable to the dependent variable of mathematical reasoning score. The various social and biological variables described by Benbow are potential mediating (explanatory) variables. If they are to serve as effective explanatory variables they must show a relationship both to sex and to mathematical reasoning scores. If a mediator variable shows a strong relationship to mathematical ability scores but no relationship to sex, it cannot serve as an explanation of the sex-mathematics relationship. This logical state of affairs implies that it will not be enough simply to record the relationship between possible explanatory variables and either sex or mathematics scores; *both* relationships must be recorded. We can gain some useful insights even if the sex-mediator correlations and the mediator-mathematics score correlations come from different studies (Harris & Rosenthal 1985). Ideally, however, there will be a subset of studies that provides both correlations for the same persons.

Once we have the three requisite correlations of sex and mathematics score, sex and mediator, and mediator and mathematics score, we can examine the effect of sex on mathematics score after partialing out the effect of the mediator. If the partial r approaches zero the mediator may be a plausible explanatory variable. If partialing out the effect of the mediator yields a partial r that is not greatly diminished, the mediator, although a possibly useful predictor of mathematics scores, is not a very plausible explanatory variable. Partialing procedures can easily be generalized to deal with several mediators simultaneously, although interpretations then become more complex than is often realized even by experienced users of multiple-regression procedures.

One practical suggestion is to begin by reducing a larger number of social and biological variables to a smaller number of composite variables that can then serve as mediators in the partialing process (Rosenthal 1987).

In addition to my basic plea for a more quantitative handling of full or even partial reviews of the literature, I have a comment on Benbow's analysis of changes over time of sex differences in mathematical reasoning scores. I agree that for the seventh graders of the years 1983 to 1986 (sect. 5) there is no change in the magnitude of the sex differences. However, if we add her data for seventh graders for the years 1972 to 1979 (Table 1 of the target article) a trend does appear as shown in my Table 1 [this commentary]. For the 10 data points of 1972 to 1986 the correlation between date of study and the sex difference favoring males is $-.50$, a rather dramatic correlation showing a decreasing superiority of males over females over the 15-year span. If we drop the year 1973, which has a somewhat extreme difference score of 55, the correlation drops only modestly to $-.39$. There is precedent for a decrease over time in male superiority of quantitative performance. For the far less selected samples of college and high school students studied by Hyde (1981), Rosenthal and Rubin (1982) reported an r of $-.21$. This tendency for a decrease in sex difference over time for Hyde's samples was also reported in an independent analysis by Becker and Hedges (1984).

Table 1 (Rosenthal). *Sex differences in SAT-M scores as a function of year*

Year of research	SAT-M difference (Male–Female)
1972	37
1973	55
1974	33
1976	34
1978	35
1979	32
1983	31
1984	36
1985	34
1986	32

In conclusion, we know very little about the reasonableness of explanations for sex differences in mathematical reasoning scores. The evidence is quite suggestive, however, that these sex differences may be decreasing over time and at a rate rather faster than the gene can travel (Rosenthal & Rubin 1982, p. 711).

ACKNOWLEDGMENT
Preparation of this commentary was supported in part by the National Science Foundation.

Mathematical ability, spatial ability, and remedial training

Barbara Sanders
Department of Psychology, University of Connecticut, Storrs, Conn. 06268

The general framework within which Benbow presents the case for a biological explanation of the male advantage in SAT-M performance in mathematically gifted students was erected originally to explain the male advantage in spatial ability. The central idea is that the sex difference in proficiency is attributable to a sex difference in hemispheric specialization (Levy 1976; Witelson 1976; Harris 1978). Although Geschwind's version of the theory (Geschwind & Behan 1982), to which Benbow makes specific appeal, also extends the model to mathematical ability, the case which can be made for the appropriateness of the model in this realm is far weaker than for spatial ability.

The major problem with applying Geschwind's theory to the sex difference in mathematics is the assumption that mathematical reasoning is a lateralized skill. Data supporting this assertion are meager by comparison with the numerous findings linking visuospatial processing and the right hemisphere. Moreover, there is little direct evidence that the male right hemisphere differs from the female right hemisphere in executing mathematical problems, whereas indications of such a sex difference are fairly common in the more abundant literature on spatial ability. On the face of it, mathematical reasoning would seem to involve both verbal and nonverbal holistic skills. As Annett and Kilshaw put it (1982): "Mathematics is a discipline for representing those aspects of the world which would otherwise be represented as complex spatial images. It would be fruitless to argue whether mathematics is a verbal *or* spatial activity when it is the discipline which coordinates and unifies these two aspects of human intellectual activity" (p. 564). Benbow's finding that a verbal factor *and* a spatial factor accounted for the high performance of a sample of mathematically precocious students (cited in Section 11) is consistent with the view that proficiency in mathematics requires both left- and right-hemisphere skills. To the extent that these two types of skills *can* be disentangled when mathematically gifted individuals solve mathematical reasoning problems, one might speculate, as Benbow does, that the sex difference is produced solely by a sex difference in spatial ability. Benbow's alternative, that is, that males and females differ in some *independent* right-hemisphere function related specifically to mathematical reasoning ability, while not inconceivable, rests on very little empirical evidence.

A potential difference between the sex difference in spatial ability and that in mathematical reasoning is that the latter may be primarily a matter of greater male variability. Benbow is careful to restrict her conclusions to mathematically gifted students, implying that there may be a difference in the mechanisms responsible for the sex difference in high- and low-ability groups. It is clear from the data she presents that the sex difference is magnified among the more talented students. There is at present no evidence that this is true of the sex difference in spatial ability. A substantial male advantage (a full standard deviation) has been found on some spatial tests in unselected populations of high school and college students (Sanders et al. 1982). The question of whether this sex difference might be even larger in a sample of spatially gifted students, however, has not been systematically addressed. The sex × handedness × general reasoning ability interaction that has emerged across several large data sets (Harshman et al. 1983) suggests that there may indeed be a difference in mechanism between high- and low-ability groups. In view of the results reported by Benbow in the target article, this question deserves further exploration.

Benbow has noted several physiological correlates of SAT-M performance. Her findings, although suggestive of a physiological basis for mathematical reasoning ability, certainly do not necessitate a physiological explanation of the sex *difference* in this domain. The link to the latter is through Geschwind's proposal that these physiological differences reflect differences in prenatal exposure to testosterone and that testosterone also affects the relative development of the two cerebral hemispheres. In evaluating Benbow's argument, it is well to remember that Geschwind's ideas, though extremely provocative, are far from demonstrated. Even in the realm of spatial ability, support for his model requires a fairly selective reading of a complex and frequently contradictory set of results.

On a practical note, I object strongly to Benbow's suggestion in this and other presentations of her work that remediation programs in mathematics be established for girls. Although Benbow has shown that gender is correlated with mathematical aptitude, the correlation is far from perfect. Selection for remedial programs in mathematics or in any other cognitive domain should be based on measured aptitude (or achievement) rather than on gender.

Neuropsychological factors and mathematical reasoning ability

Alan Searleman
Department of Psychology, St. Lawrence University, Canton, N.Y. 13617

After describing the evidence that a sex difference exists in extreme mathematical reasoning ability in favor of male preadolescents, Benbow discusses a number of environmental hypotheses that have been invoked to account for the observed sex difference. She maintains that these explanations, while perhaps accounting for achievement differences among average-ability males and females, probably don't apply to the mathematically precocious. Benbow then lists three "physiological correlates" that she claims are related to high levels of mathematical reasoning ability – left-handedness, allergies, and myopia.

First, I would caution against considering left-handedness as being predominantly a physiological trait. Most enlightened researchers believe that hand preference has both biological and environmental components. For instance, it is well known that left-handers have often been coerced into becoming right-handers by environmental pressures arising from familial and/or cultural biases, prohibitions, and superstitions against left-sidedness (for reviews, see Dawson 1977; Harris 1980; Porac & Coren 1981). Recently, it was discovered that overt pressures to influence hand preference formation were still occurring as late as the 1960s. Porac et al. (1986) found that 68.42% of left-handed or ambidextrous children raised in North America had experienced direct environmental pressure to switch hand preference to the right. Although approximately equal percentages of males and females endured such pressures, females were significantly more likely than males (61.5% versus 26.3%, respectively) to report that the pressure was at least moderately successful in altering their hand preferences.

Second, given the results of the Porac et al. (1986) study, an alternative explanation can be provided to account for the excess of left-handers among mathematically gifted and/or verbally gifted children. Directly opposite to Benbow's general position, the higher proportion of left-handers may be related to *environmental,* rather than biological, differences between average-ability and precocious children. This would occur if differential amounts of environmental pressure are applied to each group to change handedness preferences. It seems reasonable to assume that parents and teachers may not try as vigorously, if at all, to alter the hand preferences of left-handed children who are obviously very precocious (not wanting to disturb a good thing), whereas, in contrast, left-handed children of lesser ability would be subjected to the typical pressures. If this is an accurate scenario, the overall effect would be to find an increased percentage of left-handers among those with exceptional reasoning skills. Somewhat ironically, this alternative explanation is in accordance with Benbow's own theme concerning the effects of environmental influences (i.e., environmental factors that influence average-ability children do not affect extremely precocious children to the same extent).

Benbow also advances the controversial hypothesis that people (especially males) with left-handed relatives are more likely to have a bilateral or diffuse representation for cognitive functions. She further suggests that such a bilateral representation may be associated with extremely high mathematical and verbal reasoning abilities.

In the more typical parlance of neuropsychology, I believe she is suggesting that there should be a greater incidence of familial sinistrality (FS) among the precocious subjects, particularly for males. A positive history of FS is usually defined as having at least one biological parent or full sibling who is left-handed or ambidextrous. Benbow (1986b) collected data on family handedness patterns for precocious and nonprecocious children that would have allowed her to determine whether or not FS was associated with superior reasoning abilities. Unfortunately, the data were not reported in a way that would allow the hypothesis to be examined.

There are some data, however, that do address this topic. Searleman et al. (1984) examined the mathematical and verbal SAT scores of 86 left-handed college students. When combined SAT scores were examined it was discovered that strongly left-handed subjects with FS had much lower scores than did other groups of left-handers. In addition, there was a significant triple interaction between FS, Sex, and Type of Aptitude Test. It was observed that a positive history of FS was associated with *poorer* performance for males on both mathematical and verbal tests. For females, FS was related to poorer performance on tests of verbal ability but better performance on mathematical ability. Post hoc tests revealed that the males with FS had significantly lower mathematical ability test scores than did the males without a history of FS. The latter result, in particular, would not be in keeping with Benbow's implied hypothesis that there should be a positive correlation between FS in males and superior ability.

Finally, with regard to the general supposition that bilateral representation of cognitive function is related to extraordinary reasoning abilities, I find the evidence less than compelling. To support the supposition, Benbow cites the results of a pilot study that lacks a proper control group (Benbow & Benbow 1986) and a personal communication from Levy. For more convincing evidence, I would suggest performing a series of standard dichotic listening and visual field tests on samples of precocious and nonprecocious subjects. Is there evidence that precocious children are less lateralized than their nonprecocious counterparts? Within each sample, are the least lateralized children the most skilled in reasoning? Is this only true for the precocious subjects? Furthermore, as Benbow herself acknowledges, it is widely held that males are more likely than females to have greater, not less, functional hemispheric specialization. Unless there is empirical evidence that this pattern is somehow reversed among the precocious subjects, wouldn't this suggest (at least according to Benbow's general supposition) that more females than males should be precocious, the exact opposite of her overall premise?

Causes of things and nature of things: Advice from Hughlings Jackson

Daniel W. Smothergill
Department of Psychology, Syracuse University, Syracuse, N.Y. 13244

It is trivially true that heredity is responsible to some degree for differences in virtually all human abilities from automobile driving to xylophone playing; mathematical reasoning included. How these abilities group themselves and what the criteria should be for determining group membership are subjects of considerable interest (Gardner 1983). About the only dissent to this is to be found among the descendants of John B. Watson, many of whom still exist at heart despite the prevailing opinion that nothing as simple as theoretical behaviorism (Kessen 1965) could possibly be true. Humans really are not infinitely malleable, able to be shaped into doctor, lawyer, beggar man, thief, depending solely on rearing conditions. In fact, it was pointed out just recently in these pages that children raised under conditions traditionally regarded as extremely similar (i.e., in the same family) routinely turn out to be very different from each other (Plomin & Daniels 1987).

Nor is it true that sex differences are likely to be understood simply in terms of societal pressures operating on an infinitely malleable blob that happens to possess either male or female genitals. I would suspect, instead, that nature in her wisdom has endowed males and females with certain specialties that psychological science is now able to discern only dimly. The highly reliable evidence reviewed in Benbow's target article concerning sex differences in mathematical reasoning is important and deserves the attention a BBS treatment draws. At the same time, it seems important to point out that the research strategy underlying this work is very similar to what made a muddle of a more famous problem. Hughlings Jackson, the great nineteenth-century neurologist, is given credit for a research maxim that has been largely ignored in whole areas of psychological research: "The study of the causes of things must be preceded by the study of things caused" (cited in Garvey 1977, p. 4).

Perhaps the clearest example of what can ensue when the search for causes takes precedence over the study of things can be found in IQ research. [See Jensen: "The Nature of the Black–White Difference on Various Psychometric Tests: Spearmen's Hypothesis." *BBS*:(8)2 1985.] As Block and Dworkin (1974) pointed out in great and careful detail, the history of IQ research has been a true scientific anomoly in proceeding under the assumption that questions about the causes of individual differences (heredity/environment) can be meaningfully pursued without prior (or, at least, contemporaneous) progress in understanding the nature of the thing itself. As they point out, the result has been that we are now no closer to a conceptual understanding of IQ, in the sense that *temperature* is understood in physics, than we were when the merry chase to understand the causes of differences in IQ began more than three-quarters of a century ago.

The parallel to sex differences in mathematical reasoning is too clear. Again, differences are found to exist on some poorly understood construct that can be measured reliably by paper and pencil tests and has been shown to have some real world correlates. Question immediately asked: Are these differences due to heredity or environment? More than half the target article is devoted to this question and, not incidentally, the controversial nature of the question is a major source of interest in the research. But, what about the kind of question Hughlings Jackson claimed should be asked first? What is mathematical reasoning? A section (3) of the target article addresses this question, but the answer is all too predictable from the history of IQ research. "We do not have a theory to explain or define . . . mathematical reasoning ability." And, regarding the test on which the observed differences Benbow seeks to explain have been found: "Although the College Board reports that the SAT measures developed mathematical reasoning ability, we do not know what the SAT-M measures, especially among these young students."

It is hard to escape feeling somewhat embarrassed by that and wishing that the many pages of Benbow's target article devoted to possible *causes of differences* in mathematical reasoning had instead been spent reviewing research on the *nature* of mathematical reasoning. True, the state of the art does not permit that kind of review. The larger point, however, is that the state of the art has come about as a result of a particular style of research, one that gives priority to studying the causes of differences rather than the nature of things.

Nor would there seem to be much reason to anticipate that the nature of mathematical reasoning will be any more illuminated by future research of this type than the understanding of intelligence has profited from studies of IQ. Hughlings Jackson seemed to understand that the order in which scientific questions are asked is important, that it is much more than a matter of fashion or preference because it determines what comes to count as knowledge.

The male/female difference is there: Should we care?

Robert J. Sternberg
Department of Psychology, Yale University, New Haven, Conn. 06520

If we were to find out that the appendices of intellectually talented boys are larger than those of intellectually talented girls to an extent that was statistically significant but not substantial, we probably wouldn't care; the same could be said for the respective sizes of their feet. If, on the other hand, we were to find out that the intellectually talented boys earned more money over the course of their lifetime or exhibited less depression, we probably would care, even if the difference was small in magnitude but statistically significant.

The difference between the first two items and the latter two shows that what matters is not the fact of a difference between two populations, but the theoretical or practical importance of the difference, and the interpretation that can plausibly be assigned to it. In the case of Benbow's target article, four questions need to be addressed.

1. *Is there a statistically significant difference between intellectually talented, preadolescent boys and girls in their mathematical reasoning talent?* If we accept Benbow's definition of mathematical talent as "the ability to handle long chains of reasoning," the answer is unclear. The tests that Benbow and others have used measure only a very limited aspect of the ability to handle long chains of reasoning. For example, various kinds of logic problems (such as syllogisms) require long chains of reasoning, but we do not have in this article a demonstration of significant differences on such problems. Formulating important problems often requires long chains of reasoning, but we have no idea whether boys, in their thinking, formulate more important problems than do girls. Actually, I believe Benbow's

definition of mathematical talent is too broad, as not all long chains of reasoning involve mathematics. Verbal problems can involve long chains of reasoning, even if they have no quantitative aspects. But even if we restrict our definition of mathematical talent to quantitative problems involving long chains of reasoning, we still do not have a particularly complete test of whether boys are better than girls. I do think, however, that the article is persuasive in arguing that the intellectually talented boys do significantly better than the intellectually talented girls on fairly narrow operationalizations of mathematical reasoning talents, especially the mathematical section of the Scholastic Aptitude Test.

2. *How large is the difference?* Benbow's data appear to argue for a difference of about a half of a standard deviation. With large numbers of cases, which Benbow has assessed, this difference is certainly large enough to be statistically significant. Whether it is practically significant will depend on the context of the difference. This brings us to the next question.

3. *Should we care about the difference?* Whether or not we should care about the difference depends on whether the difference is large enough to achieve consequential differential outcomes between boys and girls in well-specified situations. In other words, the question is, sizeable enough for what? Benbow's target article is largely silent on this point. Even if there were enough of a difference to predict average grade-point differences between boys and girls – and such a difference is not persuasively documented in Benbow's article – I am not particularly taken by the importance of grades in any subject. They just do not predict later performance in any career very well, including academic careers. If what we care about is the contribution that intellectually talented persons ultimately make in their lifetimes – and this is one thing we might well care about – we are not going to get a handle on that contribution by looking at grades or SAT scores. So there is no point in getting excited about the difference unless we are convinced that it matters for some important outcome, preferably in the long term, but even in the short term. It is not clear for what outcome the difference matters. Why should we care about this difference any more than we care about differential in size of appendix or feet?

4. *What is the source of the difference, if it exists?* I am convinced that there is a difference between boys and girls on the narrow operationalizations of mathematical talent, but I am not convinced that we know much about why it is there. Benbow has quite carefully surveyed possible origins of the difference that do not seem to be able to account for it, but she tells us much less about sources that might succeed in accounting for the difference. The fact that intellectually talented children have more allergies, left-handedness, and myopia than intellectually average students does not seem to bear directly on the source of a difference in mathematical ability between intellectually talented boys and girls, interesting as the other differences may be. Moreover, it is important to take note of Benbow's own admonition that the weakness of the environmental factors is only with respect to the intellectually talented children, not with respect to all children. I am not convinved that the operationalizations of the environmental factors Benbow has used always fully test the environmental factors. But she has certainly done a creditable job of considering alternative explanations, even if she has not hit upon one or more of the correct ones. And she is to be admired for her courage in tackling a problem that most people don't want to hear about.

In sum, the target article does tell us that there is a relatively small but statistically significant difference between intellectually talented boys and girls on certain narrow operationalizations of mathematical talent. It does not make clear why we should care about these differences, or what their likely origin is. Given that many people will assign more importance to SAT scores than to foot sizes, if only because the former are used for college admission purposes and the latter are not, one would hope that researchers in this area will soon shed some light on the origins and importance of the differences Benbow has documented.

Hormones and sexual differentiation

Heidi H. Swanson

Netherlands Institute for Brain Research, Meibergdreef 33, 1105 AZ Amsterdam, Netherlands

The study of sex differences in brain organization is important because it contributes to the broader question of individual differences in brain organization and may lead us to understand how particular biological characteristics may enhance or constrain certain functions.

Sex hormones secreted during fetal development in both males and females affect many morphological, physiological, and biochemical characteristics of the brain, which may be reflected in differences in behaviour, attitudes, and abilities (De Vries et al. 1984a). In the most intensively studied animal model, the rat, endogenous or exogenous (experimental) androgenization of the brain during fetal development has wide-ranging behavioural consequences, some directly related to copulatory behaviour, others affecting activity, exploration, aggressiveness and fearfulness (Beatty 1979; De Vries et al. 1984). Sex differences in learning may reflect differential sensitivity and responsiveness to environmental stimuli. For example, sex differences in performance during active and passive avoidance procedures might be interpreted in terms of a relatively strong tendency in males to suppress activity in stressful situations, in contrast to a female predilection for escape. It is possible to increase or decrease the proficiency of one or the other sex on specific tasks, either through hormonal interventions or by altering the parameters and contingencies of the test situation (Van Haaren & van de Poll 1984; Van Hest et al. 1987).

Whereas some of the organizational actions of early hormones may be expressed in juveniles, most only become manifest in the adult, in response to postpubertal secretion of sex hormones. Although most subjects in Benbow's study were preadolescent, an increase in the secretion of sex hormones in humans begins well before physical signs become apparent (Donovan & van der Werfften Bosch 1965), so that both organizational and activational effects of hormones could have been involved in the observed sex differences. It is important to remember that hormones *allow* a behaviour to occur but do not produce the behaviour. Hormones may thus act as a trigger to set off a specific pattern of behaviour in response to environmental conditions.

The question as to whether it would be desirable to try to selectively improve human performance by manipulating the internal or external environment can be examined practically and philosophically. Practically, it *might* be possible to affect performance on specific tasks (such as the SAT test) through special training, alterations to the test situation, or endocrine interventions. Reinisch and Sanders (1984) conclude from quasi-experimental studies in humans that exposure to steroid hormones during gestation affects human behavioural development in a manner consistent with that seen experimentally in laboratory animals. Nyborg (1984) describes a model in which circulating testosterone and estrogen may act as intervening variables to coordinate the development of gender-related traits, including spatial and verbal ability, but he emphasizes that the model is "person-specific." Even if it were possible to predict the outcome, most people would probably object to the use of hormones, either prenatally or prior to testing, to improve intellectual function. Inadvertent prenatal influences may nevertheless produce unexpected alterations in hormonal balance. Thus, in rats, prenatal stress depresses fetal testosterone secretion, inhibits masculine development of the sexually dimorphic nucleus (Anderson et al. 1985), and prevents cerebral lateralization (Fleming et al. 1986).

Philosophically, the question may be phrased: "Is it desirable to eliminate sex differences?" It could be argued that it may be more profitable to reinforce *individuals* at whatever they do best, and if there happen to be more members of one sex or the other it does not matter, as long as all abilities are equally valued. I would be interested to know whether any studies similar to Benbow's have been carried out with adolescents initially selected on the basis of their verbal ability. In view of the generally accepted female advantage in verbal skills, a bias to the advantage of girls might have emerged. Because skills in verbal communication are just as important for the functioning of modern society as mathematical proficiency, special facilities could then be provided for the further development of verbally gifted individuals of either sex.

Finally, it has often been suggested that because of possible political or social misuse, research on sex differences should not be carried out (cf. Sapolsky 1987). On the contrary, I propose that such research is vital precisely because there are so many popular preconceived ideas and prejudices. Considered application of findings such as those presented by Benbow can ensure that the best facilities can be provided for, and the best use can be made of every individual's capacities.

On throwing bones to environmentalists

Donald Symons
Anthropology Department, University of California, Santa Barbara, Calif. 93106

At the risk of appearing to use one of Stephen Potter's (1950) classic gambits of reviewership – viz., find out the quality for which an author is most famous, and then blame her for not having enough of it – I must aver that Benbow's conclusion slights biology. After meticulously reviewing the evidence bearing on various socialization/environmental explanations for the widespread and persistent sex difference in mathematical reasoning ability among intellectually talented 12- to 13-year-old students, Benbow reports "that many years of research by the staff of SMPY and by others using the SMPY data base have not turned up results that provide support for the various socialization hypotheses" (sect. 10.4, para. 4). But because "there are well-documented differences in the socialization as well as in the biology of boys and girls" (Summary, para. 2), Benbow concludes that this sex difference results "from both environmental and biological factors" (abstract).

Boys and girls may indeed be socialized differently. In the absence of evidence that socialization has anything to do with a given sex difference, however, there is no more reason to assume that socialization magnifies that difference than there is to assume that socialization mitigates it. The latter possibility is by no means farfetched: A number of innate sex differences in sexuality, for example, probably are mitigated by socialization (Symons 1979; 1987). Benbow's conclusion may, of course, turn out to be correct, but it certainly does not follow from the evidence that she has so ably marshalled.

Sex differences in mathematics: Why the fuss?

Lionel Tiger
Department of Anthropology, Rutgers University, New Brunswick, N.J. 08903

Perhaps the most interesting question about the Benbow material is why there should have been any expectation that the sexes would be utterly the same rather than that they produce differences in behavior. After all, males and females are different, so why should they not act differently? The absurd consequence of commencing with the problem of explaining why differences occur, rather than why they do not, is in that form of American psychological transcendentalism represented by the Bem test for androgyny, which in effect seeks to examine the extent to which real male and female respondents to tests differ from a theorized profile of responses of a fantasy androgynoid that does not exist (Bem 1974).

The human brain, like all other brains, evolved not to think but principally to act and since sexual selection has to do precisely with sex differences it is peculiar to be puzzled that the sexes will think differently, just as they act differently. Part of the problem is psychologists' obsession with tests, which leads them to regard behavior on tests as predictive of real behavior rather than the other way around, which is empirically more parsimonious. Also, in preparing many tests, psychologists try to remove questions that distinguish between the sexes – on the ground that such distinctions must reflect culturally induced differences, which of course begs the question. Then when differences nonetheless continue to emerge, there is all this fuss about the need for even purer cultures that inflict less distinction on the different sexes. This is again part of a general pattern, "the psycho-industrial complex" (Tiger 1987), in which peoples' behavior is seen as a product to be monitored, evaluated, and rewarded; any seemingly primordial blemish on a free market, such as sex, is subject to banishment. And of course the emphasis placed in all this discussion upon exotic performances of highly circumscribed people in relatively limited strata of societies, based largely on literacy, skews any conclusions away from those that might also reflect the overwhelming number of human beings in other cultures and periods of history in which the arcane matter of doing mathematics was broadly unimportant to the conduct of life in the community.

There are interesting variations, however, in sex difference discussion, depending on the perhaps political interpretation of the finding. If many females do less well than many males in mathematics, this is clearly a bad thing about which no one would in principle disagree, something to be stamped out. But when Gilligan (1982), on the basis of rather perilous extrapolation from a study of some schoolchildren, described differences in moral order between males and females, she was celebrated for this finding, perhaps by precisely the people who castigate Benbow for the results she reports.

Benbow's emphasis on the adolescent period seems very appropriate. In our study of sex differences in the kibbutz movement, where boys and girls were by and large raised together in children's houses, in a setting of strong ideological and practical commitment to equality, Shepher and I found that girls tended to do better in school before adolescence but their relative performance declined thereafter (Tiger & Shepher 1975). This is difficult to attribute to conventional socialization theory insofar as for the first eight years of school girls were superior to boys and the overall ethic of the community stressed that there should be no differences. Adolescence made the difference – hardly a surprise in view of the fact that it is a period centrally connected with reproduction – when whatever sex differences there are should become more salient than in childhood.

Could these sex differences be due to genes?

Steven G. Vandenberg
Institute for Behavioral Genetics, University of Colorado, Boulder, Colo. 80309

Whenever one reads about a sex difference in means and variances, the question is raised whether these differences

Table 1 (Vandenberg). *Concordances of male and female twins on the Mathematics Achievement Test in the National Merit Scholarship study*

Males			
	MZ	.74	$h^2 = .66$
	DZ	.41	$c^2 = .08$
Females			
	MZ	.71	$h^2 = .50$
	DZ	.46	$c^2 = .21$

Note: h^2 is the genetic variance; c^2 is variance due to common environment.

might be due to the fact that males receive only one X chromosome, whereas females receive two. This could mean that on the quantitative trait in question males have a distribution determined by the effects of one chromosome, whereas females have a distribution that is the result of two chromosomes, perhaps an average of some kind. It is true that Mary Lyon (1961) has shown that one of the two X chromosomes in human females is turned off ("inactivated") rather early in embryonic development in all cells and that the choice of whether the paternal X or the maternal X is turned off in each cell is initially random, but thereafter continues to be the same for each cell's descendants. However, much of the sex-determining work has been done by the time the X is inactivated, not so much because of the presence of two X chromosomes as because of the presence (or absence) of a Y chromosome. Whatever effect the X chromosome might have on a trait – for example, mathematical ability – could thus result in a different distribution for one-X males than for two-X females. In the strict sense of genes on the X chromosome directly influencing the trait in question, we speak of sex linkage. As far as I know, there has been no sex-linkage study, in that strict sense, of mathematical ability. Sex linkage has been proposed for spatial ability, but careful examination of the evidence by Vandenberg and Kuse (1979) and Harris (1979) led to the conclusion that this idea was not tenable. What are we to make of this? A possible explanation is only somewhat more complicated. The fact that the newly developing embryo has one or two X's does make a difference, but only indirectly through the differential development of the two sexes, not just in their sexual characteristics but also in their brain structure and functions, as summarized by Kelly (1985) and by Kimura (1987).

Behavioral genetic studies have sometimes included measures of mathematical abilities. On closer inspection, however, these are usually more measures of numerical calculation or arithmetic. Benbow makes a clear distinction between this ability and more abstract mathematical reasoning. One of these studies (Loehlin & Nichols 1976) is somewhat more relevant: The National Merit Scholarship study of twins included a measure of mathematical achievement. Although Benbow warns against equating mathematical achievement (based in large part on applying learned rules) with mathematical reasoning (at least in part based on new solutions to unfamiliar problems), the National Merit twin data have some relevance. The correlations for mathematical achievement are shown in Table 1.

Another study of some relevance is one by Arleen Garfinkle (1982). In this study 137 MZ (monozygotic) and 72 DZ (dizygotic) same-sexed white twins aged 4 to 8 years were administered a battery of 15 tests, entitled the Piagetian Mathematical Concepts Battery (PMCG), derived by Garfinkle from Piaget's work. The 15 tasks are briefly described by Garfinkle (1982, p. 33).

Table 2 shows the results of a factor analysis of all 15 tasks. No sex differences were found. The MZ and DZ correlations for the age-corrected total scores are .73 ± .04 and .56 ± .06. Applying Falconer's formula, $h^2 = 2(r_{MZ} - r_{DZ})$, gives a value of .34 ± .18 for the genetic variance, or heritability. The total score correlated only moderately with the Coloured Progressive Matrices (PM), the Peabody Picture Vocabulary Test (PPVT), and a

Table 2 (Vandenberg). *Varimax rotated factor matrix of the tasks in the PMCB for 418 children in the Colorado Piagetian Twin Study*

	Factor[a]			
Task	I	II	III	C[b]
Conservation of number	.20	.78	.17	.68
Counting	.50	.38	.34	.50
Seriation-T	.34	.62	.39	.66
Parts and wholes	.56	.21	.15	.38
Transitivity	.20	.32	.07	.15
Addition and subtraction	.61	.10	.22	.43
Conservation of number-identity	.64	.26	.23	.53
Conservation of number-equivalence	.12	.76	.17	.62
Discrimination	.46	.07	.08	.23
Seriation-K	.40	.34	.75	.84
Insertion	.43	.46	.57	.72
Numeration	.60	.27	.29	.51
Sorting	.29	.12	.14	.12
Some and all-class inclusion	.52	.17	.08	.31
Multiple class membership	.54	.24	.20	.39
Percentage of common variance	44.2	34.5	21.3	

[a] A factor loading is a correlation coefficient of a variable and the factor. Factor loadings of .50 or more are underlined.

[b] Communality, the squared multiple correlation of each variable with all other variables.

Table 3 (Vandenberg). *Correlations among the PMCB, PM, and VM, with age partialed out, for 418 children in the Colorado Piagetian Twin Study*

	PMCB[a]	PM[b]	PPVT[c]	VM[d]
PMCB	—			
PM	.41			
PPVT	.36	.23		
VM	.22	.19	.19	—
Correlation with age	.75	.59	.70	.43

Note: For $N = 418$, the critical value ($p < .01$) of the correlation coefficient is .13.

[a] Piagetian Mathematical Concepts Battery.

[b] Progressive Matrices (Colored).

[c] Peabody Picture Vocabulary Test.

[d] Visual Memory.

measure of visual memory (VM) as shown in Table 3, leading Garfinkle to conclude that the PMCB measures an ability that is largely independent from conventional intelligence measures. Garfinkle says that sex differences in the ability to learn mathematical concepts may not show up until children are older.

Bias and sampling error in sex difference research

Douglas Wahlsten

Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Benbow argues that, because she can find no environmental differences to explain a small sex difference in Scholastic Aptitude Test-Mathematics (SAT-M) scores in a highly select sample of volunteers from the top 3% of seventh-grade children, we should therefore look to biology for an explanation. The reasoning here is reminiscent of arguments about racial differences in intelligence used by Jensen (1969), who declared unabashedly (and prematurely) that "Compensatory education has been tried and it apparently has failed" (p. 2) and then used this to support his assertion that differences must be largely genetic. [See also Jensen: "The Nature of the Black/White Difference on Various Psychometric Tests" *BBS* 8(2) 1985.] However, there is no necessary relationship between the ease of modifying a characteristic by changing the environment and the extent of influence by chromosomal or genetic variation. Any evidence for a biological source of a sex difference in SAT-M scores must come from well-controlled studies of the biology of the children.

Among the four kinds of physiological correlates discussed in Benbow's target article, not one provides consistent evidence of a causal link with a sex difference in mathematical ability. For allergies, there is no sex difference in the Study of Mathematically Precocious Youth (SMPY) sample. For myopia, SMPY girls are even more extreme than boys. Concerning prenatal hormones, the only indicator in the SMPY sample is birth order, which is an exceedingly indirect "measure" of physiology and is strongly related to social processes. Likewise, hand preference is not a physiological measure at all, and it is sensitive to cultural influences. Benbow's discussion of physiology is indeed "speculative." Her conclusion shows that she is not sufficiently critical of some very weak research.

For example, Benbow gives credence to two studies relating the size of the corpus callosum (CC) to sex and laterality differences. One study (deLacoste-Utamsing & Holloway 1982) claiming that females have a larger splenium (posterior portion) of the CC was based on only 9 male and 5 female brains at autopsy, and the other (Witleson 1985) claiming that the whole CC is smaller in right-handed people examined 42 decreased cancer patients. However, Nasrallah et al. (1986) failed to find any relation between CC size and sex or handedness using nuclear magnetic resonance imaging with 41 normal adults. Five other studies failed to find a sex difference in the human CC (Bell & Variend 1985; Byne et al. 1986; Demeter et al. 1985; Oppenheim et al. 1987; Weber & Weis 1986). Seven studies with a total sample size of over 113 male and 108 female brains have failed to replicate the results for females in the widely cited deLacoste-Utamsing and Holloway (1982) study, although results for male brains have been more consistent. The 5 female brains in the 1982 study had a mean cross-sectional area of the splenium of the CC equal to 218.3 square mm, whereas combined data from three other studies involving female brains measured in a similar way yielded a 99% confidence interval for the mean area of the female splenium from 161.3 to 178.52 square mm. The sex difference in the splenium reported in *Science* in 1982 was truly a case of sampling error.

Benbow strongly asserts that sex differences in SAT-M scores do not result from different experiences. Because human boys and girls are treated differently from the cradle to the grave, the only way Benbow can prove her point is to know in considerable detail those features of the environment that are in fact most relevant to the nurturance of mathematical ability and how they function jointly. Then she can compare the experiences of boys and girls fairly and rigorously.

However, the SMPY project does not do this. Instead, it compares boys and girls on several rather indirect measures of experience one at a time. This approach is not likely to detect relevant features of the environment. To see this problem, let us suppose there are more than 20 different aspects of early experience, each of which can augment or impair later mathematical ability by a small amount. Now measure the correlation between SAT-M score and an indirect indicator of each element considered separately. By fragmenting the totality of relevant experience and attenuating each effect by indirect measurement, correlations will generally be low and undetectable.

What if we were to use the same approach in a genetic study of SAT-M scores? Suppose there are more than 20 autosomal genes, each of which exerts a small influence on the development of mathematical ability and each of which is linked to (located near) a detectable marker gene, such as one specifying some protein in the blood. Now determine what alleles each person has at each marker locus, and then look for associations between each genotype and test score. Only rarely will any significant and replicable association be found. This is why attempts to identify genes relevant to complex behaviors or abilities using genetic linkage analysis (Ashton 1986; Sturt & McGuffin 1985) or DNA restriction fragment length polymorphisms (Ellis 1986) show so little promise. A series of separate tests of indirect measures of many small environmental influences on mathematical ability applied to an extremely narrow range of boys and girls would not be expected to yield much in the way of significant sex differences, so it comes as no surprise that most results of the SMPY study concerning environment are negative. In view of this, surely our current ignorance about the nature of experience cannot justify support for a biological view by default.

Factors influencing educational productivity

Herbert J. Walberg

College of Education, University of Illinois at Chicago, Chicago, Ill. 60680

Few policy makers would propose to exclude girls from rigorous mathematics even if they were shown to be less able by nature or nurture. Perhaps for this reason, Benbow concludes that environmental factors are of greatest practical importance. Indeed, considerable research, accumulated in the past five years, indicates what improvements are likely to strengthen the mathematics performance of both boys and girls. Recent evidence supports the view that mathematics performance in the United States leaves much to be desired and that such improvements are required to make learning in mathematics, science, and other subjects more productive.

In *A Nation at Risk* (1983), the National Commission for Excellence in Education warned about the under-performance of U.S. students on achievement tests; recent data provide an even grimmer picture with respect to the subject of mathematics (although science and other subjects could be cited; Walberg 1983). Among samples of students completing elementary school in a dozen countries and two provinces of Canada, U.S. eighth-grade students scored third from the bottom. Among the top 5% of twelfth graders completing secondary school in each of the countries surveyed, U.S. students ranked last among Western countries and Japan; they exceeded only the less economically developed countries in Africa and South America (U.S. Department of Education 1986, pp. 28, 30). The gap between Japan and the U.S. is sufficiently large that most Japanese girls exceeded most American boys.

The causes of such poor performance are best illustrated in an important paper in *Science* by Stevenson et al. (1986), who studied the learning of mathematics in Japanese, Taiwanese, and U.S. elementary school classes. Internationally calibrated IQ tests showed that all three groups were equally able at the start of schooling; but with each year, Asian students drew further ahead in achievement. A small achievement advantage at the end of the first grade grew ever larger, so that by fifth grade, the worst Asian class was superior to the best American class. (This "Matthew effect" of the educationally rich getting richer has been observed in many U.S. studies; Walberg & Tsai 1983).

The Asian students had a far more rigorous curriculum and worked at a faster pace. They studied far more at school and at home, with their parents' encouragement and help; those who fell temporarily behind spent extra time to catch up. In the U.S., success was more often attributed to ability; in Asia, to hard work.

This study shows the powerful effects of rigorous instruction and supportive family environments. It is confirmed by several large-scale U.S. and international educational surveys and in the quantitative synthesis (or "meta-analysis") of thousands of smaller-scale educational experiments (Walberg 1984). These surveys and experiments show that nine factors increase learning. Potent, consistent, and widely generalizable, these nine factors fall into the three groups shown in Table 1.

Table 1 (Walberg). *Nine educational productivity factors*

Student aptitude
- (1) Ability or prior achievement as measured by the achievement tests
- (2) Development as indexed by chronological age or stage of maturation
- (3) Motivation or self-concept as indicated by personality tests or the student's willingness to persevere intensively on learning tasks

Instruction
- (4) The amount of time students engage in learning
- (5) The quality of the instructional experience including psychological and curricular aspects

Psychological environments
- (6) The "curriculum of the home"
- (7) The morale of classroom social group
- (8) The peer group outside school
- (9) Minimum leisure-time television viewing

Each of the first five factors – prior achievement, development, motivation, and the quantity and quality of instruction – seems necessary for learning in school; without at least a small amount of each, the student can learn little. Large amounts of instruction and high degrees of ability, for example, may count for little if students are unmotivated or instruction is unsuitable.

These first five essential factors, however, are only partly alterable by educators since, for example, the curriculum in terms of lengths of time devoted to various subjects and activities is partly determined by diverse economic, political, and social forces. Ability and motivation, moreover, are influenced by parents, by prior learning, and by students themselves. Thus educators are unlikely to raise achievement substantially by their own efforts alone.

The remaining factors – the psychological climate of the classroom group; enduring affection and academic stimulation from adults at home; and an out-of-schol peer group with academic interests, goals, and activities – influence learning in two ways: Students learn from them directly, and benefit from them indirectly because they raise student ability, motivation, and responsiveness to instruction. In addition, about 10 (not the average of 28) weekly hours of leisure-time television viewing seem optimal for learning, perhaps because more television time displaces homework and other educationally and developmentally constructive activities outside school. The powerful influences of out-of-school factors, especially the home environment, must be considered since the 12 years of 180 six-hour days in elementary and secondary school add up to only about 13% of the waking, potentially educative time available and about 8% of all time during the first 18 years of life.

To improve the performance of both boys and girls, U.S. educators, parents, and students will undoubtedly have to work longer, harder, and more efficiently (Walberg 1983; 1984; U.S. Department of Education 1986). Although research indicates which specific programs and methods work most productively, they are of little use unless they are put into place.

Neuroanatomical sex differences: Of no consequence for cognition?

Sandra F. Witelson
Departments of Psychiatry, Psychology, and Neurosciences, McMaster University, Hamilton, Ontario, Canada L8N 3Z5

Benbow discusses possible causes of boys' higher mean scores and superior performance on tests of mathematics, reviews hypotheses about the social and environmental causes of these sex differences in mathematical performance and notes that although some environmental correlates of mathematics scores can be found, their "causality cannot be truly demonstrated." This situation, combined with the growing number of findings of neuropsychological and neurobiological sex differences, makes a consideration of biological correlates of mathematical ability both reasonable and promising. Benbow reviews some of the neurobiological correlates of mathematical test performance, such as individual differences in the anatomy of the human brain, particularly in the corpus callosum (the main interhemispheric fiber tract) and in the pattern of hemisphere functional asymmetry. I will comment on the latter issue first. Benbow suggests that cognitive differences in spatial ability and its possible correlate, mathematical ability, may be related to greater bilateral representation of cognitive skills in females. In addition, Benbow suggests that the observation of a greater incidence of left-hand preference among the mathematically gifted (and in other fields such as music and architecture), may

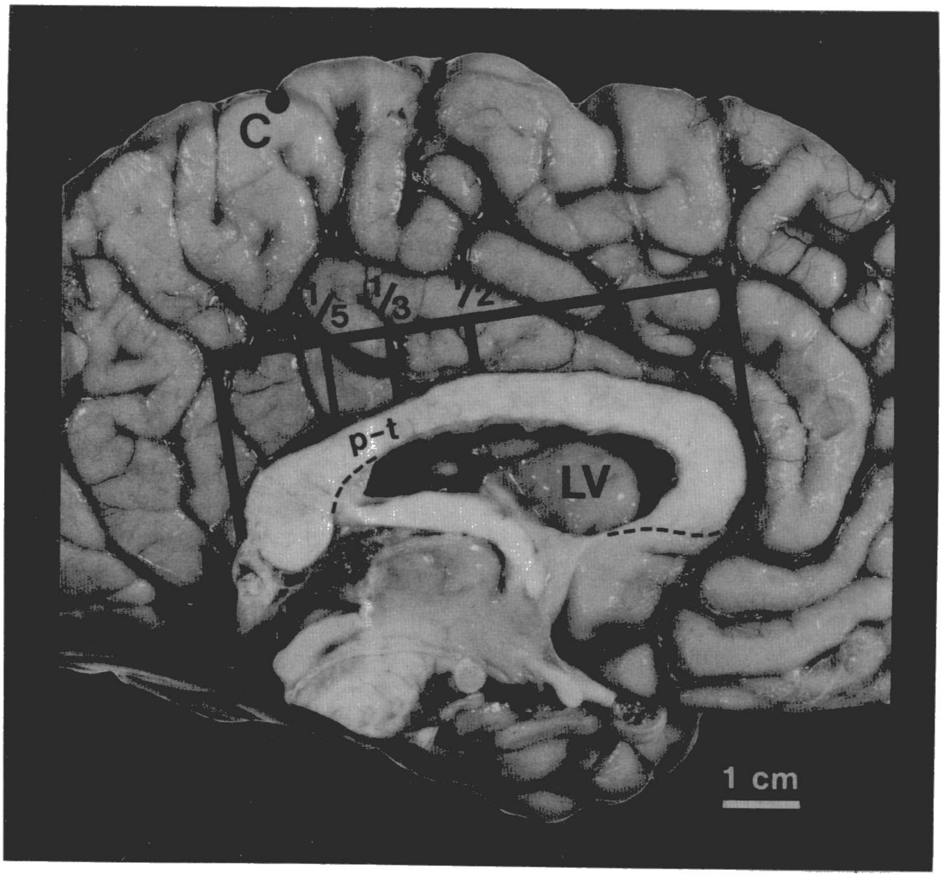

Figure 1 (Witelson). The human corpus callosum is shown in midsagittal section with its boundaries indicated by dashed lines. The line joining the most anterior and posterior points of the callosum is used as the axis to define the subdivisions as indicated. The midposterior region (the posterior third minus the posterior fifth region) is approximately the area of the callosum that connects the right and left parietotemporal regions relevant for language and spatial functions. Abbreviations: C, central sulcus at the dorsomedial aspect; LV, lateral ventricle exposed by removal of the septum pellucidum; p-t, parietotemporal region of the callosum (isthmus) (from Witelson 1986).

be related to "bilateral or diffuse representation of cognitive functions and/or a strong right hemisphere." This postulate of a strong right hemisphere is not further elaborated by Benbow but could be seen as a direct contrast to the option of bilateral representation.

It seems reasonable to infer that different patterns of cortical localization of function have cognitive consequences, especially if one notes that this aspect of brain organization not only appears to be present at birth but is associated with early sex differences (Witelson 1987a). However, the nature of the association, if any, between patterns of hemisphere functional asymmetry and cognitive skill is still unknown. Several obvious complexities make it clear that any possible relationship cannot be as simple as Benbow's discussion suggests. For example, sex differences in brain lateralization are usually in the direction of greater lateralization in men for *both* verbal and spatial tasks – yet men do better than women on spatial tasks and worse on verbal tasks. Arguments that variation in cognition correlates with brain lateralization may best be made at a general level, with explicit documentation of the complexities of any relationship. This is especially important if one aims to convince proponents of solely environmental factors.

Benbow refers to the reports of a possible sex difference in the midsagittal area of the corpus callosum as another potential neurobiological correlate of sex differences in ability. In one report (de Lacoste-Utamsing & Holloway 1982) the posterior one-fifth region of the callosum (roughly congruent with the splenium) was observed to be larger in females than in males, both in absolute area (at the .08 level of probability) and in maximal width. In a later study (Holloway & de Lacoste 1986) only the sex difference in maximal splenial width was observed. In the larger samples of the few studies done at the turn of the century, in my own work (Witelson 1985), and in the few more recent reports, sex differences were not found in total callosal size or in the posterior one-fifth region, either in adults (see review, Witelson & Kigar 1987) or in children (see review, Witelson & Kigar 1988). My interpretation of the data is that females do not have a larger (absolute or proportional) splenial region than do males.

There may be sex differences in other parts of the corpus callosum, however. Benbow very accurately reported the results of my 1985 paper, which found a nearly significant sex-by-hand interaction effect for the size of the overall posterior one-half region of the callosum.

In a subsequent study (Witelson 1986) the posterior part of the body or trunk of the callosum (the isthmus or "parieto-temporal" callosal region) showed a particularly marked difference in size between consistent-right-handers and mixed-handers, regardless of right- or left-hand writing (see figure 1). In a further study of subregions of the callosum with an expanded sample of 50 cases (15 men, 35 women), statistically significant sex differences in the size of the callosum were found in some regions (Witelson 1987b). Two sex differences are relevant here: (1) Hand preference was found to be a factor in the absolute size of the isthmus in males only. This area was 56% larger in the mixed-handers. (2) The consistently right-handed females had a larger absolute isthmus region than the consistently right-handed males; females did not have a larger area in any other callosal region.

The isthmus region connects the parieto-temporal cortical regions of the two hemispheres. These regions are involved in the representation of linguistic, spatial, and musical skills – skills that are represented asymmetrically in the cortex. The observation of a sex difference in brain anatomy that interacts with hand preference is consistent with the findings of interactions between sex and laterality in psychological studies, including Benbow's current report.

It is difficult to imagine that sex differences in the anatomy of brain regions related to cognition could be inconsequential. It remains to be demonstrated, however, whether such neuroanatomical variation is related to measures of cognition. It does seem worthwhile to consider the hypothesis that such neuroanatomical differences are among the factors leading to sex differences in behavior. They could be influential in cognitive tasks such as mathematics and also in social behavior such as moral judgment (Gilligan 1982) or the pursuit of professional and personal goals (Abramson & Franklin 1986). As Benbow remarks: "It would not and does not help females if differences are swept under the rug." I have argued elsewhere (Witelson 1985a) that it may be of no benefit to society to assume that the two sexes are basically homogeneous with respect to cognition and that the only heterogeneity is that imposed by different experience. Zero variation is not a requirement of equal opportunity.

The forgotten realm of genetic differences

Ada Zohar and Ruth Guttman
Department of Psychology, Hebrew University, Jerusalem, Israel 91905

Benbow's target article is a lucid, well-organized review that surveys and evaluates the plausibility of some biological and mostly environmental factors contributing to the considerable sex differences in mathematical reasoning ability – henceforth called MRA.

There are, however, two realms unexplored as yet:

1. The detailed definition of MRA. What constitutes the phenotype of MRA is largely unformulated. We know that it is not the application of material or rules previously learned. Yet Benbow's definition – "the ability to handle long chains of reasoning" – does not relate specifically to mathematical reasoning and could be applicable to literary criticism. Benbow does list some cognitive abilities that correlate with MRA but there is no clear definition of the function of each of these elements or its relation to the others. Mathematics includes diverse fields such as topology, algebra, geometry, calculus, set theory, logic, and so forth, each differentially related to spatial ability and with quite different modes of reasoning. Thus, is an individual with a high score in the SAT-M necessarily outstanding in *all* the above? It is unclear whether excellence in mathematical reasoning ability constitutes a simple phenotype or several.

2. Genetic factors in MRA. As Benbow shows, girls and boys have intense socializing forces operating in the direction of the observed sex differences. And yet it must be concluded that in themselves these socializing forces are insufficient to explain the enormous sex difference in outstanding mathematical reasoning ability. If, in fact, the major force operating was social, we would see the greatest change in the difference between the sexes emerge during adolescence when the sex-role pressures peak. The establishment of the difference in very early adolescence – by the seventh grade – argues that there must be other important forces to be considered.

A genetic basis for mathematical ability has been studied extensively. In two large classical twin studies in Denmark in the 1960s, Husén (1959; 1960) found the correlations for mathematical achievement to be about 0.80 for monozygotic and 0.5 for dizygotic twins. These data have since been replicated by Fischbein (1981) in Sweden and by several others in the U.S. and Great Britain (Scarr & Saltzman 1982). These studies deal with general mathematics achievement but not with outstanding ability. There is evidence of outstanding mathematical ability running in families. Barlow (1969) in his book *Mental Prodigies* describes many such cases, among them the pedigrees of the Bernouli and the Pascal families first published by Galton. In these families one finds mathematical genius reappearing in several generations and branches of the family. As Benbow's longitudinal studies show, some of the very high scores on the SAT-M are potentially creative mathematicians of exceptional talent. It is on this level that the sex difference becomes particularly evident.

Since there are so many more boys than girls with the phenotype of outstanding ability, an X-linked recessive gene may be a *major* factor in the genetic system associated with the phenotype. This gene may have something to do with MRA or may be a regulatory gene controlling a system of autosomal genes. The strong point of the X-linked gene hypothesis is that it can be easily refuted on the strength of pedigree analysis: Father–son transmission would rule out MRA's being X-linked.

Since MRA is a multivariate profile, separate elements of it should be clearly defined and their transmission subjected to genetic analysis. We are aware that a similar hypothesis was proposed for the inheritance of spatial ability (Caplan et al. 1985; Stafford 1964) and has, to date, not found support. This should not deter researchers from making a similar attempt with regard to MRA, and in particular with exceptional mathematical reasoning ability. Pedigree studies of this phenotype may show either sex-linkage or other forms of transmission. An additional contribution of such a study will be a fuller understanding of MRA and its component structure.

Benbow states that "Even though biological factors seem to be involved in determining the sex-difference in mathematical reasoning ability this does not imply that efforts at remediation cannot make a difference. They probably can and ought to be tried. Thus practically speaking, one must be an environmentalist." To design an effective intervention to improve women's MRA we need to know more than the extent of the sex difference. It is important to know what components are involved, how they develop with age, and the length and nature of critical periods. Thus we would like to amend Benbow's statement to read: Practically speaking one must be a behavior geneticist taking into account genetic processes to design an effective environmental intervention.

Author's Response

Sex-related differences in precocious mathematical reasoning ability: Not illusory, not easily explained

Camilla Persson Benbow
Study of Mathematically Precocious Youth (SMPY), Department of Psychology, Iowa State University, Ames, Iowa 50011-3180

I appreciate the many thoughtful commentaries on my target article. They should help greatly as I and others plan future research endeavors, for after my initial despair in the face of hundreds of pages of massed criticism, it became apparent that my commentators had outlined a research program for my next 40 professional years. (Now all I need is a grant for half the national debt and 124 hours per day!) My response will focus on eight themes common to the set of commentaries.

What does the SAT-M measure?

Bloom, Braine, Chipman, Hunt, Jackson, Mayer, McGuinness, Money, Smothergill, Sternberg, and **Zohar & Guttman** all address the problem that we do not know precisely what the Scholastic Aptitude Test-Mathematics (SAT-M) measures. Mayer, McGuinness, and Smothergill, in particular, make the case that we need a theory to define mathematical reasoning ability, the ability the SAT-M was designed to measure. I can only agree, but especially with Smothergill, who adds that the state of the art does not yet permit that kind of review, being the result of an American research style that gives priority to investigating causes rather than things, as he and McGuinness point out. Jackson takes this reasoning a bit too far, however, when she suggests that we have to develop a theory before we study and report sex differences on the SAT-M. This is analogous to suggesting that we must thoroughly understand sickle cell anemia before we determine that blacks suffer from it more frequently than whites.

After having reviewed the scant literature on the nature of mathematical reasoning ability, I concluded that the ability to handle long chains of reasoning, singled out by Gardner (1983), best summarized the themes running through the various articles (Benbow 1988). **Jackson** criticizes this conclusion because it implies that problem-solving is a mechanical process, which is not consistent with the fact that steps or rules change as problem-solving progresses, or that there are individual differences in preferred strategies. I would have thought, however, that this was exactly what handling long chains of reasoning *did* imply. It is apparent then that my summary definition was unclear, being perhaps as **Sternberg** suggests, both too limited and too broad.

I must also agree with **Sternberg** that the SAT-M is a narrow operationalization of mathematical reasoning ability and mathematical talent. Stanley and Benbow (1986) carefully delineated what aspect of mathematical reasoning ability we felt the SAT-M measures in these gifted seventh graders. The test was designed to measure developed mathematical reasoning ability of high school students, most of whom have studied rather abstract mathematics for several years. We feel that because most of the seventh graders in our talent searches were demonstrably unfamiliar with mathematics from algebra onward yet many were able to get high scores on the SAT-M, they had to have exercised extraordinary ability at the analytical level of Bloom's (1956) taxonomy. We accordingly concluded that the SAT-M must draw more on analytical reasoning in the SMPY test-takers than in high school juniors and seniors, a conclusion for which we have some evidence (Minor & Benbow 1986).

McGuinness asks why we dismiss the Senk and Usiskin (1983) findings, which showed that instruction affects the ability to perform geometric proofs. In fact, we do not dismiss that conclusion. It is actually used to formulate the basis for my statement that "practically speaking, we must be environmentalists" and that intervention strategies ought to be directed toward girls, a stance to which **Eysenck** objects and **Jensen** feels is a "farfetched" fantasy. What I did reject was Senk and Usiskin's conclusion that they had refuted the argument that there are sex differences in mathematical reasoning ability. They felt that the formulation of proofs after prolonged instruction was a better measure of mathematical reasoning than ours. I disagree; it is merely a different one. SMPY focused on the ability to solve problems without having been explicitly taught how to do so. For this reason, too, the interesting data that Jensen provides on ethnic and sex differences on the mathematics parts of the Stanford Achievement Test address another issue.

Halpern provides a useful clarification of our data. She states that the SAT-M is an aptitude test designed to predict how much an individual will benefit from instruction, and that the SMPY girls benefit from instruction at least as much as the boys. I largely agree with that conclusion, but it fails to take into account the finding that SMPY males scored higher on standardized achievement tests at the end of high school even though their course grades were slightly lower than the SMPY females. Thus, by Halpern's definition, SMPY males have greater potential to benefit from early instruction than SMPY females do, and on some measures (standardized tests) they do exhibit higher levels of achievement. This conclusion is not limited to the SMPY population. Stanley (1987a; 1987b) studied scores on a wide range of achievement tests (not just in the math/science areas). He found large gender discrepancies, especially in science, mathematical sciences, and history, among college-bound high school students, who also exhibit gender differences in SAT scores.

Mayer and **Farmer** find an inconsistency in my report that boys are better on SAT-M but girls receive better course grades. Moreover, Mayer feels that grades are a less biased indicator of mathematical ability. Few people would agree with the latter point. It is well known that grades frequently depend on how often homework is turned in, how neat the work is, how quiet and attentive the student is, and so forth. In studying gifted children, we encounter many students who get A's on all tests but poor overall grades because they fail to turn in homework. Moreover, there are many types of mathematical ability (a factor of which Mayer appears to be unaware). As **Hunt** indicates, tests usually tap higher-level abilities whereas grades tap lower-level ones. We focus on mathematical *reasoning*, which most would agree is a higher ability. Mayer's confusion regarding the several types of mathematical ability is apparent when he discusses the age effect on sex differences in mathematics. In elementary school few differences are found. This is probably because the curriculum emphasizes computation and the learning of basic concepts and few tests have been designed to measure any other abilities in that age group. I am not aware of any tests of mathematical *reasoning* ability designed for elementary students. Our work is restricted to the latter ability, as indicated in the target article.

Another misapprehension is that the SAT-M consists primarily of word problems. That is not the case. Thus, when **Chipman** argues that there is sex bias in its content as a result of the word problems, she is misinformed. **Braine** argues that the SAT-M does not measure mathematical reasoning ability. The Educational Testing Service's technical guide to this test states explicitly that the SAT-M is designed to measure *the developed mathematical reasoning ability* of college-bound eleventh and twelfth graders (Angoff 1971). Above we described how

SMPY interprets the test when given to gifted seventh graders. Finally, **Bloom** suggests that perhaps there are no sex differences in mathematical reasoning ability but rather in the *process* of mathematical reasoning. Although this serves to narrow the interpretation of the finding, it certainly does not change the fact that there are many more boys than girls who score extremely well on the SAT-M (Stanley's and my central finding).

I conclude that these commentators are indeed right in realizing that we need to know the nature of mathematical reasoning ability. In collaboration with Veronica Dark, I have begun such work (Dark & Benbow, in preparation).

Magnitude of sex differences

Sternberg questions the practical significance of a difference of only about .5 standard deviations in magnitude. That ignores the main point of my target article and that of Benbow and Stanley (1983b). We place significance on the ratios of boys to girls with high scores on the SAT-M. The higher the cutting score on the SAT-M, the more males than females there are, such that at a SAT-M score of 700 or more we have identified approximately 13 boys for every one girl, even though equal numbers of boys and girls took the test. This is a difference that is indeed practically significant.

Kornbrot feels that citing such ratios is misleading and inappropriate; she provides effect sizes for the male: female ratios at various score levels (i.e., for the 2 : 1 ratio at 500, for the 4 : 1 ratio at 600, and for the 13 : 1 ratio at 700 SAT-M). She comes up with effect sizes of 0.25, 0.40, and 0.40, respectively, and concludes that the differences are not important. These calculations seem misleading. Out of the 292 students achieving scores of 700 on the SAT-M before age 13 over a three-year period, where essentially equal numbers of boys and girls were tested, only 23 were females (i.e., less than 10%). We do not need statistics to tell us that that difference is important. Moreover, my effect size for that difference is closer to 2.0 than to the 0.4 figure reported by Kornbrot. Kornbrot also states that citing group A to group B ratios is misleading since ratios always increase with the severity of the criterion when effect size is constant. This is correct, but what Kornbrot does not seem to consider is the difference in variability and the distribution of SAT-M scores for males and females (see Figure 1). The male scores are skewed. This skewness is found only for SAT-M scores, not for the SAT-V scores of the same group (see Figure 2). The SAT-V distributions are the same for males and females. Moreover, the SAT-M distribution for females has the same shape as their SAT-V distribution. Thus, in essence, our research should be focused on trying to discover *why* the male SAT-M score distribution is so different rather than why SMPY girls do not score as well as SMPY boys do, as **Becker & Hedges**, **Humphreys**, and **Mackenzie** point out.

Bloom states that 15 years of finding sex differences on the SAT-M may not be long enough to warrant the conclusion that there are many more mathematically talented males than females. The reader might contrast this with the view of **Kornbrot**, who states that the results are impressive because they span and are consistent over a 15-year period, which is long in the life of a researcher and in the history of feminism. Finally, **Rosenthal** argues that the sex difference is decreasing, but in so doing he assumes that all the samples are equivalent; they are not. From 1980 on, the samples have indeed been consistently selected by the same criteria. During this time period there is no evidence for a decrease, rather the opposite. The r of the magnitude of the sex difference with time (1980 through 1986) is +0.66. For example, the size of the sex difference in the Johns Hopkins 1987 talent search (N = 26,870) was 33 points, up one point from 1986.

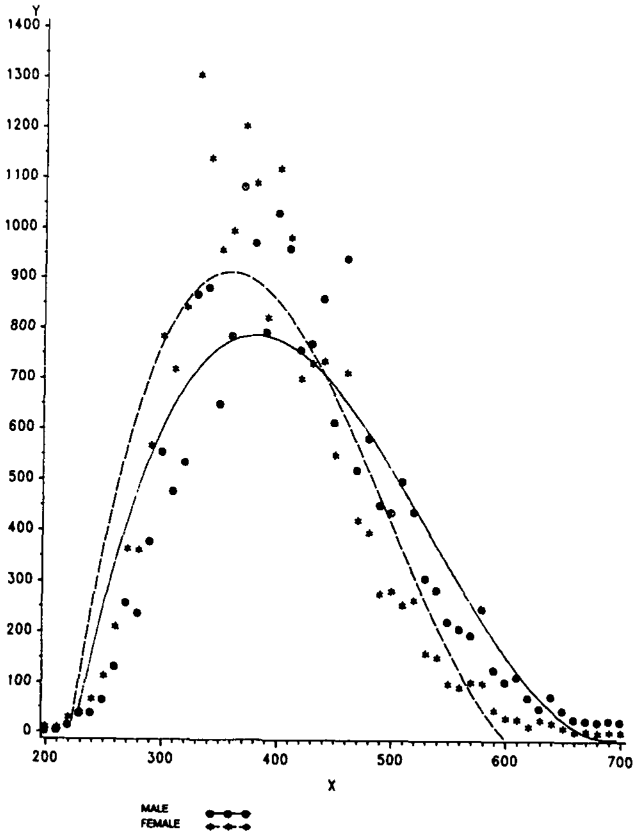

Figure 1. The distribution of SAT-M scores by sex for intellectually talented seventh grade students who took the test from 1980 through 1983 as part of the Johns Hopkins CTY talent searches.

In sum, the sex difference in SAT-M scores among the intellectually talented is large, consistent, and of practical significance.

Does the sex difference have consequences?

Sternberg states that my target article is silent on whether the sex difference has any important consequences and he suggests that we are trying to address an issue that is similar to caring about differences in the size of appendices or feet. As indicated in Section 8, our sex difference in ability does indeed predict later sex differences in achievement in high school (Benbow & Minor 1986; Benbow & Stanley 1982a) and at the end of college (Benbow 1987a).

Mackenzie feels we cannot claim that the sex difference has important consequences on the basis of fewer females

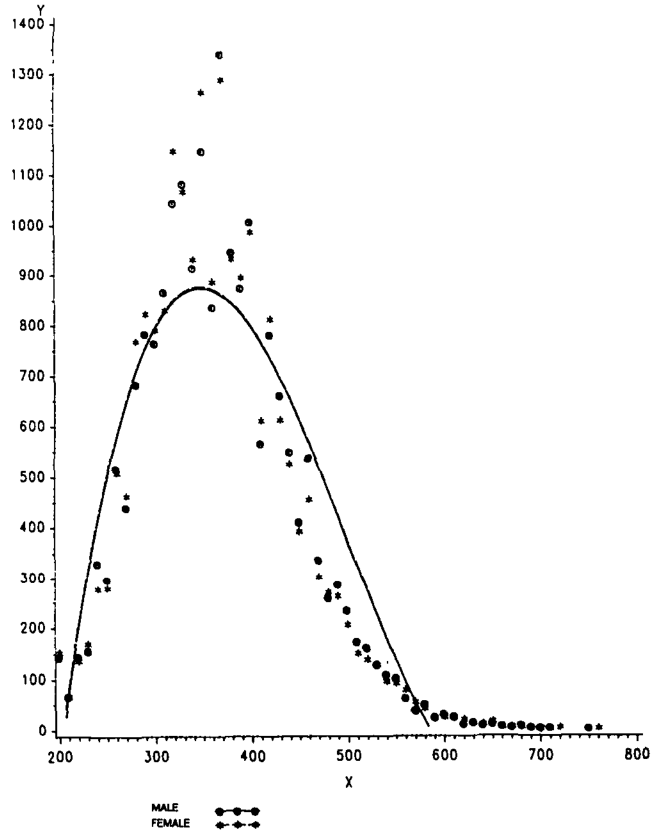

Figure 2. The distribution of SAT-V scores by sex for intellectually talented seventh grade students who took the test from 1980 through 1983 as part of the Johns Hopkins CTY talent searches.

than males participating in the sciences. Rather, we should look at performance in university mathematics courses and Ph.D. dissertations. We do know that females take many fewer mathematics courses than males do, but when they do take the course, their grades do not differ. Interpreting grades, however, is extremely problematic. As for Ph.D. dissertations, most of our students are not old enough to have reached that level yet. That is a future analysis, which I agree is necessary. Thus, the only important available indicator to date is participation in science. Many more males than females are participating at a high level and, thus, they will undoubtedly produce more Ph.D. dissertations.

Farmer states that less than 25% of SMPY subjects obtain graduate training in mathematical and scientific fields and wonders, therefore, who our future scientists and mathematicians will be. Farmer and I seem to differ in our interpretation of this 25%. I found it to signal very high participation in graduate training in the math/sciences for a group of students selected on only one variable, SAT-M scores in the seventh grade. Furthermore, about half of our males and 25% of our females had career goals in the math/sciences (not including social sciences). Since most of the students identified had talents in more than just one area, that seems to be strong evidence that students with relatively high scores on the SAT-M at an early age comprise a substantial part of our future pool of mathematicians and scientists. It would be unrealistic to expect that all or almost all students with high scores on the SAT-M in the seventh grade would be attending graduate school in the math/sciences or even pursuing careers in those areas. This reasoning seems analogous to some researchers' efforts to show that the Terman gifted group was not so special because no Einstein was identified. How many Einsteins would you expect out of a sample of 1,528 gifted kids (≥ 135 IQ) living in California and identified over a short period of time? My prediction would be zero.

Socialization factors

Sternberg's and **Hunt**'s remark that I and my colleagues, most notably Stanley, have not fully tested the environmental factors summarizes many of the commentators' opinions and our own. Our conclusion was that in 15 years we have not been able to identify a clear environmental explanation of our sex differences findings, *not,* as **Wahlsten** asserts, that sex differences in SAT-M scores do not result from different experiences.

Mills questions the appropriateness of some of our measures. It is true that in some studies we have used only responses to one item to test a socialization variable. When the studies were repeated by us or by others, however, using more sophisticated measures, the results did not differ.

Unfortunately, **Mills** is selective in the studies she cites in her commentary. For example, she found some evidence for a relationship between cross-sex type personality characteristics and measures of intellectual ability for females. With Mills's assistance, Zimmerman (1984) expanded the study and found that these cross-sex type personality characteristics (those who have some personality characteristics of the opposite sex), although consistent over time, could not predict later achievement in mathematics. Zimmerman (1984) and Fox et al. (1979) both rejected the masculine identification hypothesis. In addition, Mills states that most of our studies involve males and females of the same high mathematical ability (i.e., top 1 in 10,000) and wonders why we would expect to find any differences. Her statement is not accurate. The first work using such subjects to study socialization variables appeared in print in 1986 and is thus very recent. Most of our investigations deal with talent search participants; nearly all of these are in the top 3%, but with a wide distribution of SAT-M scores. On the SAT-M our girls and boys are *not* matched and thus we can expect to find differences, a concern also raised by **Newcombe & Baenninger.**

Mills claims that we have test bias because our measures were administered after the students were identified as being mathematically talented. Fortunately, we administered talent search questionnaires to the 10,000 students in the SMPY talent searches *before* they were tested, and some measures were administered at the same time as the aptitude testing. Mills also points out, correctly, that some of our studies on socialization variables are retrospective and that this introduces possible bias, distortion, and selective remembering. We agree, but see no practical alternative when dealing with students of this rare frequency, who cannot yet be identified at a much earlier age. The real issue is whether or not the nature and direction of such biasing is gender-related in ways that obscure the effects of sociological variables. Moreover, **Wahlsten** suggests that we study subtle environmental influences not separately but in a unified model. We believe that is an excellent idea and have actually begun such investigations using path analysis (i.e., LISREL).

Mills, Halpern, and **Becker & Hedges** discuss the possibility of sampling bias. In 1980 Stanley and I addressed this issue. We found, contrary to their expectations, that the females participating in the SMPY talent searches comprised a group that was higher in ability, relative to norms for their gender, than the boys (Benbow & Stanley 1980). Moreover, since as many girls as boys participate in the talent searches, Halpern's concern that only females (but not males) with positive attitudes toward math/science participate seems misplaced. As **McGuinness** notes, it seems illogical to reason that society encourages girls to excel in computation, where they are superior, but not in mathematical reasoning ability, or that society encourages girls to excel in other subjects but not in mathematics. **Tiger**'s comments are interesting in that regard too. He has found that in the Kibbutz, females do better in school than males do until adolescence. Coinciding with puberty there is a big change, even though the Kibbutz ideology is one of equality.

Moreover, a frequent mistake found in research on socialization variables is to assume that any gender differences found must be environmental in origin. For example, **Mills** mentions that gender differences in cognitive style should be considered. Is cognitive style totally the result of socialization, as she implies? As **Eysenck** points out, the same gender differences in attitudes and other socialization variables would be expected if the sex difference in mathematical reasoning ability were due entirely to genetic causes or entirely to environment. Thus, the existence of gender differences in variables such as attitudes or cognitive style does not provide conclusive evidence for an environmental explanation of the sex differences. Furthermore, **McGuinness** is quite right in stating that socialization research is correlational and therefore cannot shed light on causality.

Kornbrot reports that in Britain there are no gender differences in top grades earned or in first-class honors in pure mathematics at the University of Edinburgh; she accordingly infers that the environment must be responsible for gender differences in aptitude. This confuses grades with mathematical reasoning ability, the ability we studied. Kornbrot is not alone in confusing achievement with ability. Others, such as **Bleier,** do so as well. Socialization probably has far greater effects on achievement than on reasoning ability. In addition, Kornbrot reasons that because students who take science courses do better in degrees unrelated to science, that is evidence for the general importance of course-taking in the math/sciences. Alternatively, it is in fact the most able students who probably take extensive courses in the math/sciences.

Kornbrot suggests that there may be an interaction with sex and other, perhaps larger, predictors that are not part of the analyses; we then wrongly attribute to sex the share of the explained variance that really arises from these other variables. Kornbrot suggests that parental education and SES (socioeconomic status) may be two such interacting variables. We have performed analyses with such variables included. Generally, SES and parental education were not very useful in explaining achievement differences in high school (Benbow 1981).

Chipman and I obviously disagree about the interpretation of her recently co-edited book (Chipman et al. 1985). When reviewing the book (Benbow 1986c), the clear message I picked up concerned the importance of parents in influencing female achievement in the math/sciences. I still draw the same conclusion in rereading it. Not all of the chapters in her book dealt with the role of parents. When they did, there was clear support for the important role that parents play in influencing mathematics achievement (see, for example, pp. 93, 119, 175, 183, 195, 222, and 245). Chipman found it strange that I could use Wise's (1985) chapter in her volume to argue that there are small sex differences in course-taking in mathematics. Actually, I used both Wise (1985) and the Armstrong (1985) chapters to make that point. Armstrong showed that the sex difference in math participation is indeed very small now. Wise showed that when prior ability is taken into consideration the difference in course-taking was small even in 1963.

Chipman argues that differential course-taking can indeed account for the sex difference in SAT-M scores in my sample, despite Armstrong's chapter in her own book. She reports that more boys than girls take physics, computer science, calculus, and so forth, and therefore receive more practice in problem-solving. I do not see the relevance of these figures to Stanley's and my data. Our sex differences were found in the seventh grade, before any of these courses are taken and before differential course-taking in mathematics has begun for boys and girls (Benbow & Stanley 1980; 1982a). **Burnett**'s commentary also provides compelling evidence against differential course-taking as a likely explanation for sex differences in SAT-M scores among high school students.

Chipman also reports that spatial ability has no influence on mathematical performance. A chapter in her own book reported the contrary (Connor & Serbin 1985), as have others (e.g., Burnett et al. 1979). **Burnett** provides much evidence on that point, suggesting that the cognitive processes shared by mathematical reasoning and spatial ability tests probably involve (1) the quality of the visual image and (2) the speed with which one manipulates a coded image, both of which exhibit sex differences. In that regard, **Mills**'s statement that spatial ability is influenced more by environmental than biological factors becomes relevant. She cites no supporting reference for this statement, and I am not aware of any.

Farmer found my review of the socialization literature far too cursory to instill reader confidence in my conclusions. I must agree with her, but noting at the same time that the vote was not unanimous in that regard. **McGuinness,** for example, felt the socialization review was too tedious and wished it was not necessary for me to include it. More than one referee of my target article noted a desire to shorten that aspect of the paper. (My most recent unpublished findings are consistent with Farmer's conclusion that parental supportiveness is more predictive of female success [or achievement, not aptitude] than male success.)

Newcombe & Baenninger state that my target article is flawed because I do not provide a rationale for why the mathematically gifted should be studied as a special population, given that the distribution of SAT-M scores shows a continuous, normal, or near-normal distribution. But as my Figure 1 shows, the SAT-M distribution is not normally distributed for the males. Moreover, Figures 1 and 2 provide strong evidence for lack of restriction in range, another possibility mentioned by Newcombe & Baenninger. They also suggest that I am harder on environmental hypotheses than on biological ones, and that my target article is therefore biased. If I am harsher on the environmental studies it is because I had different purposes in mind in the portions of my review concerned with environmental than with biological hypotheses. In the case of the environmental hypotheses I tried to show that there is not enough evidence to conclude that they can account for *all* the sex difference in SAT-M scores. In the case of the biological factors I was trying to develop a case for why they may contribute *in part* to the sex difference and why such research should be considered relevant, as **Witelson** and others seem to have grasped.

Kenrick comments on the problem of the interaction between ideology and the psychology of sex differences: There are those who feel that feminist ideology should take precedence over data on gender differences. Advocates of this approach seem to use a two-front denial strategy (to deny that gender differences exist and assume, a priori, that they do not have a biological basis).

Bleier, Kornbrot, Newcombe & Baenninger, and **Wahlsten** state that I am committed to a biological explanation of the sex difference; **Chipman** cites a variety of newspaper articles to show my biological inclination. Then they proceed to argue that the sex difference in SAT-M scores among the intellectually talented cannot be due entirely to biological factors. However, I never stated that sex differences are due entirely to biology; in fact, I agree with Bleier and Halpern that they undoubtedly arise from an interaction of biology with the environment.

Braine states (and **Halpern** suggests) that more boys than girls took our test and therefore it is not surprising that we find sex differences. This is incorrect. The data reported from 1980 on, which include the majority (over 95%) of the cases, are based on equal numbers of boys and girls taking the test. Braine also states that my cross-cultural data are not based on very different cultures and are therefore not all that informative. I disagree. The Chinese and U.S. cultures are radically different. **Tiger's** Kibbutz data are relevant in this context too. Finally, Braine's plea that we focus on the factors keeping women out of some occupations rather than on socialization practices that may influence occupational decisions again seems ideological rather than scientific.

Sanders objects to my position that intervention strategies for females ought to be tried. Instead, she states, remediation should be based on ability, not gender. This means we would never increase the number of girls who are extremely talented mathematically because those targeted would not be performing at levels at which remediation would be necessary. As **Walberg** describes, there are specific programs and methods that can enhance achievement; they are of little use if not put into practice.

In sum, I agree that the environmental question is still open and that socialization is extremely important, especially for achievement. Any possible difference resulting from biology can be diminished or magnified, depending on the environment, as **Symons** and **Money** aptly point out. Much can be done to enhance the achievement of boys and girls and ought to be tried.

The importance of biology

We began studying biological factors as possible influences on the sex difference in SAT-M scores when years of work with environmental factors proved they were unable to account *entirely* for the differences. **Tiger** is puzzled by why we or anyone else would expect boys and girls to be the same. **Smothergill** wonders why we had such faith in the environment for so long when children raised under conditions traditionally regarded as extremely similar (i.e., those in the same family) routinely turn out to be so different (Plomin & Daniels 1987).

Zohar & Guttman provide a further rationale for studying biological factors. They suggest that if the major operating force is social, we would expect to see the greatest change in adolescence when sex-role pressures peak. Since the differences are established early, this argues for the importance of other factors. One possibility is **Kenrick's** suggestion that testosterone affects primate male's tendency toward hyperactive competition, which leads to better mathematical ability and achievement. A test would be whether girls completed as many items as the boys did on the SAT. A sex difference in the number of omissions may also be relevant. For the talent search students studied by Stanley and me, there were significant sex differences in both indices but they were not large (Wolins & Benbow, in preparation). At this point **Goldman-Rakic & Clark's** caution is appropriate. We must recognize the complexity of the neural networks underlying complex behaviors before we can begin to understand their sexual dimorphisms.

Physiological correlates

Our search for physiological correlates was begun in order to provide a rationale for why biology may be an influence on the sex difference in mathematical reasoning ability and was prompted by Geschwind and Behan's (1982) article. Geschwind and Behan's prenatal testosterone hypothesis appeared to be one way to explain our results. We have always been fully aware, however, that our data are only consistent with and do not constitute strong evidence for that hypothesis and that, as **McGuinness, Money,** and **Sanders** point out, the Geschwind-Behan hypothesis is far from established and is probably not entirely correct. **Denenberg, Berrebi & Fitch** do, however, provide some fascinating support from animal studies for my brain/behavior hypothesis and the importance of hormones in influencing brain development. **Goldman-Rakic & Clark** also feel that physiological factors may indeed contribute to sex differences observed in cortical function and the expression of cortical skills. They provide evidence that the cerebral cortex is responsive to androgens during gestation and postnatal life and that it might undergo sexual differentiation, as **Swanson** also believes. Finally, **Nyborg** describes his General Trait Covariance Androgen/Estrogen model, which provides another possible framework for how hormones may affect cognitive and affective development in humans.

Moreover, **Harshman** provides strong evidence that individual differences in reasoning ability are related to differences in brain organization and that sex differences in cognitive abilities are at least partly neurological in origin. He has also found clear sex differences in the cognitive correlates of left-handedness, which may help explain some of the seeming inconsistencies in my data. Yet several problems in my physiological correlates data are pointed out. Some are valid, and some are based on misunderstandings. I discuss these below.

Kimura argues that no information is provided on nonallergy disorders and, therefore, we are unable to conclude that pathologies in special populations are limited to those disorders. (**Humphreys** makes a similar point.) We never reached such a conclusion and are indeed searching for further physiological correlates. Because the Geschwind and Behan (1982) hypothesis would predict the correlates I studied, those correlates were the logical first choice to pursue. It is gratifying to know, therefore, that **Humphreys** has found similar physiological correlates of intelligence.

Kimura also states that there is no evidence for dyslexics having better nonverbal skills. I must refer her to Geschwind (1982). (**Hardyck** also claims there is no evidence that dyslexics are more frequently left-handed; Geschwind and Behan [1982; 1984] and Long and Murray [1982] seem to provide just such evidence.) Kimura also writes that there is no evidence that left-handers are superior in mathematics. That is not my conclusion because left-handers represent a diverse group in terms of brain organization. I have only studied extreme precocity and I only reported an association of left-handedness with that. Actually, there was not much of an increase in the frequency of left-handedness in my above-average comparison group (i.e., about the top 1 in 20 in ability). Kimura also objects to my hypothesis that superior reasoning skills are associated with a diffuse representation of cognitive functions – not because the idea may not be reasonable, but because I have no evidence for it. Although I agree with her in spirit, we did have some supporting evidence (Benbow & Benbow 1987). Now we are completing a dichotic listening study (as well as a chimeric presentation and a palm-tracing experiment), as **Searleman** also suggested we perform, and our data appear to be consistent with my hypothesis. Although the traditional laterality patterns were found for our control group, indications of greater involvement of the right hemisphere for the extremely talented were found (Benbow et al., in preparation). This should also address the concerns voiced by **Hardyck** and **Bryden,** as should **Harshman's** finding of a three-way interaction among sex, handedness, and reasoning ability with several cognitive abilities and with his dichotic listening data. This pattern of results suggested to Harshman that variations in brain organization may underlie cognitive differences. In fact, he predicts that cognitive differences arise in part from handedness, sex, and reasoning-related differences in brain organization and that there may be a variety of normal brain organizations.

I am grateful to **Hardyck** for detecting an error in my manuscript. When I discuss the results of my tachistoscopic study with extremely precocious subjects, I did indeed mean that the extremely precocious had lower response times for the right hemisphere than the left (Benbow & Benbow 1987), which Hardyck points out is supportive of my laterality hypothesis.

Kimura and **Witelson** state that my hypothesis of diffuse representation of cognitive functions in extremely precocious students is contradictory to my invoking greater functional asymmetry to explain my sex-difference finding. In Benbow (1986b) I pointed out the difficulty my data posed for this hypothesis, but can see how my writing in the current target article created a misimpression for Kimura, Witelson, and perhaps others. If **Harshman** is right in postulating a variety of normal brain organizations that may be associated with variation in cognition, the more or less lateralized distinction is probably too simplistic anyway, as Witelson also points out. Thus, to provide correlations between degree of laterality with cognitive ability, as **Newcombe & Baenninger** suggest, would be misleading and uninformative.

Kimura also states, however, that there is no evidence for the role of sex hormones in altering brain asymmetry. I must refer her to the commentaries by **Denenberg et al., Swanson, Goldman-Rakic & Clark, Harshman** and **Hines. Money** discusses problems with Geschwind's hypothesis about prenatal testosterone exposure and its relation to learning disabilities and mathematical precocity. I agree; it is a hypothesis in need of supporting evidence and modification. Money and Hines also question my rationale for not providing a summary of the considerable body of clinical psychoendocrine and achievement data gathered developmentally from childhood to adulthood from people with a known prenatal history of hormonal deficiency or excess. I agree that these data are important and should have been included in the target article. Several years ago, I had realized the importance of such data and had asked **Money** whether I could test some of my hypotheses using his data bank. He turned me down. I must also agree with Money that my findings are only that mathematical precocity is associated with left-handedness, allergies, myopia, and being male. I have not unravelled an explanation for these connections. My data are consistent with Geschwind's conjectures. **Nyborg's** model may be an alternative explanation, for example.

Wahlsten and **Witelson** comment on the inconsistencies in the picture for possible sex difference in the corpus callosum. In my target article I too mentioned these difficulties: "It is not clear whether differences in the size of the corpus callosum could have any relation to sex differences" (sect. 12.1, para. 9). For me the finding of real significance was Witelson's (1985): corpus callosum size differences between right-handers and left- or mixed-handers and the complicated sex factor which interacted with hand preference – findings that Witelson describes more fully in her commentary. Thus, Wahlsten's comment, based on the above exchange, that I am not sufficiently critical of weak research on biological factors seems, at least in this context, misplaced.

Several individuals (**Bloom, Hunt, Kornbrot, Mackenzie,** and **Mayer**) commented that our physiological-correlates data did not follow the pattern needed to show they could explain our sex differences data. Although the commentators make some valid points in this regard, some of the objections are based on misunderstandings.

In Benbow (1986b; 1987b; 1988) I addressed the reason one *might* expect verbal reasoning ability to be associated with the same physiological correlates as mathematical reasoning ability. Moreover, females do not show an advantage in verbal reasoning ability. Among extremely precocious verbal reasoners there are also *slightly* more males than females. In addition, there are sex differences in our left-handedness results. **Harshman** also provides evidence as to why our pattern of results may not be as inconsistent as it may first appear. Finally, our rationale for studying physiological correlates was to provide a justification for the possibility that extreme mathematical reasoning ability does have a biological (in addition to an environmental) basis, one that might be related to brain organization. According to Geschwind and Behan (1982) and Schachter et al. (1987), these physiological correlates may be related to prenatal testosterone exposure. In my view the latter link, which remains to be established, provides a connection with our sex-differences finding. **Harshman, Goldman-Rakic & Clark,** and **Denenberg et al.,** who provide a strong counter to **Bleier**'s criticisms, have data consistent with this viewpoint. (By the way, I had intentionally called left-handedness, allergies, and so forth, *correlates* to avoid the confusion with causality that **Bloom** discusses. I have never stated that they cause the sex difference.)

Hunt indicates that I used different criteria for left-handedness depending on which group was being studied and that the most lenient criterion was used for the extremely precocious students. This was not the case. Extremely precocious students were *not* classified as left-handed if they performed only one action with their left hand. We used a handedness inventory and accepted as left-handed any who performed a majority of the tasks with their left hand. This handedness inventory was used with the parents and the comparison group, but only self-report of handedness was available for the siblings. Hunt implies that we left out the Asians in our sample because their data did not fit our hypothesis. This is incorrect. Asians were studied separately because of a well-documented lower frequency of left-handedness in that population (see Benbow 1986b). There was also an increased frequency of left-handedness among the Asians in our sample (see Benbow 1986b). Finally, Hunt states that there were no interactions between sex and handedness in our data. These are reported by Benbow (1986b) and by **Harshman.** Similarly **Mackenzie** does not achieve significance when testing for sex differences in our left-handedness data using z-scores. We used a median test (because the data were not normally distributed) and found the difference to be significant.

Searleman discusses environmental pressures on left-handedness; my results with the Asian mathematically precocious students are consistent with these. He then suggests that differential pressure to change handedness is given to average and high-ability students. Inconsistent with this speculation is the finding that siblings of the extremely precocious and the comparison group, who are also bright but less so, did not show a marked increase in frequency of left-handedness.

Rosenthal, Humphreys, and **Newcombe & Baenninger** object to our not providing effect sizes and r's for our biological correlates data. For left-handedness, for example, this is not possible, because its distribution is not normal. More important, however, we believe that these correlates are only associated with *extremes* in ability (a distinction few of the commentators grasped) and our main sample was, therefore, also restricted in range. Thus, **Searleman**'s finding that left-handed FS+ male college students were inferior in their SAT-M performance may not be germane, especially if one considers **Harshman**'s findings. Nevertheless, **Halpern**'s finding of an effect of sex and familial sinistrality on university mathematics placement seems inconsistent with Searleman's results.

Contrary to Hardyck's suggestion, I am indeed aware of the effects of familial sinistrality and recent theories of laterality suggesting that tasks are not hemisphere-specific (**Hardyck**): I discuss familial sinistrality in Benbow (1986b) and Benbow and Benbow (1987). Almost half of our students are either left-handed, mixed-handed, or right-handed with a familial history of left-handedness. Such individuals are more likely to exhibit diffuse representation of cognitive functions, as our subsequent investigations have shown (Benbow & Benbow 1987; Benbow et al., in preparation). Our left-handedness findings are interesting in what they reveal about brain lateralization. Thus, even though 85% of our students are right-handed, as **McGuinness** and **Bleier** point out, most of our subjects showed some signs of diffuse representation of cognitive functions. Moreover, even if Geschwind is right about testosterone and its effect on the brain, would one expect all individuals exposed to testosterone prenatally to be left-handed? Not all pregnant women who took thalidomide bore deformed babies. **Harshman**'s comments regarding a variety of normal brain organizations are also relevant in this context.

Although **Sanders** speculates as freely as I did, she is right in her main point that there is not much evidence that mathematical reasoning ability is strongly under the influence of the right hemisphere and related to sex differences as I proposed. There is a simple reason for this: Few have studied mathematical reasoning ability (rather than spatial ability) and laterality.

In sum, I agree that my physiological correlates are not straightforward and easy to interpret. Nonetheless, they provide suggestive evidence favoring a brain/behavior hypothesis and a biological basis for mathematical reasoning ability.

Genetic bases

Zohar & Guttman, Eysenck, and **Vandenberg** emphasize the importance of performing genetic analyses before effective intervention can occur. I agree that it should be done by well-qualified researchers, but I think intervention strategies should be designed using existing knowledge in the meantime. Vandenberg's report of greater genetic variance in mathematics achievement for males than females is consistent with preliminary data from our laboratory (Benbow 1987a).

The huge annual talent searches conducted at Duke, Johns Hopkins, Northwestern, and the University of Denver could yield excellent twin and sibling data (score on SAT-M, SAT-V, and the Test of Standard Written English) for able 12- and 13-year-olds.

Greater male variability

Becker & Hedges, Humphreys, and **Mackenzie** propose, as I have acknowledged, that SMPY's sex difference data may be due to variability. Figures 1 and 2 are certainly dramatic illustrations of that point, although they also indicate that that skewness of male scores ought to be considered as well. Reschly and Jipson (1976) have shown, however, that there are no more males than females who are mentally retarded.

Summary

The purpose of my target article was to show that (1) there are dramatic sex differences in extreme mathematical reasoning ability, as measured by the SAT-M, (2) these sex differences have long-term consequences, (3) various socialization hypotheses have so far proven unable to account for *all* the sex difference in SAT-M scores (note that we are not talking about achievement or occupational success), and (4) there may be a reason why endogenous factors have some influence on this sex difference in precocity. **Sternberg** criticizes me for not providing a specific hypothesis or theory to explain these sex differences. That was not my purpose. Rather, I tried to show that socialization factors are not the only ones affecting the sex difference in SAT-M scores. It is useful to keep in mind **McGuinness**'s statement that biology is not destiny and that changing the environment represents as hard an obstacle as genetically based sex differences.

ACKNOWLEDGMENTS

I should like to thank Michael O'Boyle and Julian Stanley for helpful comments and suggestions. Financial support was provided by the National Science Foundation (MDR–8470387).

References

Abramson, J. & Franklin, B. (1986) *Where they are now. The story of the women of Harvard Law 1974.* Doubleday.

Adelson, J. (1985) What we don't know about sex differences. *New Perspectives* (Spring): 9–13.

Aiken, L. R., Jr. (1972) Research on attitudes toward mathematics. *Arithmetic Teacher* 19:229–34.

Aiken, L. R., Jr. & Dreger, R. M. (1961) The effect of attitudes on performance in mathematics. *Journal of Educational Psychology* 52:19–24.

Anastasi, A. (1982) *Psychological testing.* Macmillan.

(1985) Reciprocal relations between cognitive and affective development with implications for sex differences. In: *Nebraska symposium on motivation, 1984: Psychology and gender,* ed. T. B. Sonderegger. University of Nebraska Press.

Anderson, D. K., Rhees, R. W. & Fleming, D. E. (1985) Effects of stress on differentiation of the sexually dimorphic nucleus of the preoptic area (SDN-POA) of the rat brain. *Brain Research* 332:113–8.

Angoff, W., ed. (1971) *The College Board Admissions Testing Program.* Princeton, N.J.: College Entrance Examination Board.

Annett, M. & Kilshaw, D. (1982) Mathematical ability and lateral asymmetry. *Cortex* 18:547–68.

Arbeiter, S. (1985) *Profiles, college-bound seniors, 1984.* Princeton, N.J.: College Entrance Examination Board.

Arjmand, O., Benbow, C. P. & Lorenz, F. (in preparation) Attitude and parental encouragement patterns among intellectually talened males and females.

Armstrong, J. M. (1981) Achievement and participation of women in mathematics: Results of two national surveys. *Journal for Research in Mathematics Education* 12:356–72.

(1985) A national assessment of participation and achievement of women in mathematics. In: *Women and mathematics: Balancing the equation,* ed. S. F. Chipman, L. R. Brush & D. M. Wilson. Erlbaum.

Arnold, A. P. & Gorski, R. A. (1984) Gonadal steroid induction of structural sex differences in the central nervous system. *Annual Reviews of Neuroscience* 7:413–42.

Ashton, G. C. (1983) Myopia and cognitive ability. *Behavior Genetics* 13:526.

(1986) Blood polymorphisms and cognitive abilities. *Behavior Genetics* 16:517–29.

Backman, M. E. (1972) Patterns of mental abilities: Ethnic, socioeconomic, and sex differences. *American Educational Research Journal* 9:1–12.

Barlow, F. (1969) *Mental prodigies.* Greenwod.

Bartsocas, C. S. & Kastrantas, A. D. (1981) X-linked form of myopia. *Human Hereditas* 31:199–200.

Basu, S. K. & Jindal, A. (1983) Genetic aspects of myopia among the Shia Muslim Dawoodi Bohras of Udaipur, Rajsthan. *Human Hereditas* 33:163–69.

Battle, E. S. (1966) Motivational determinants of academic competence. *Journal of Personality and Social Psychology* 4:634–42.

Beatty, W. W. (1979) Gonadal hormones and sex differences in nonreproductive behaviors in rodents: Organizational and activational influences. *Hormones and Behavior* 12:112–63.

Becker, B. J. (1978) The relationship of spatial ability to sex differences in the performance of mathematically precocious youths on the mathematical sections of the Scholastic Aptitude Test. Master's thesis, The Johns Hopkins University.

Becker, B. J. & Hedges, L. V. (1984) Meta-analysis of cognitive gender differences: A comment on an analysis by Rosenthal and Rubin. *Journal of Educational Psychology* 76:587–87.

Bell, A. D. & Variend, S. (1985) Failure to demonstrate sexual dimorphism of the corpus callosum in childhood. *Journal of Anatomy* 143:143–47.

Bem, S. L. (1974) The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology* 42:155–62.

Benbow, C. P. (1978) Sex-related differences in spatial ability. Master's thesis, The Johns Hopkins University.

(1981) Development of superior mathematical ability during adolescence. Unpublished doctoral dissertation, The Johns Hopkins University.

(1983) Adolescence of the mathematically precocious: A five-year longitudinal study. In: *Academic precocity: Aspects of its development,* ed. C. P. Benbow & J. C. Stanley. The Johns Hopkins University Press.

(1986a) Home environments and toy preferences of extremely precocious students. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

(1986b) Physiological correlates of extreme intellectual precocity. *Neuropsychologia* 24:719–25.

(1986c) Book review of *Women and mathematics: Balancing the equation,* ed. S. F. Chipman, L. R. Brush & D. M. Wilson. In: *Child Development Abstracts and Bibliography* 60:1.

(1987a) Factors associated with high achievement in mathematics and science among mathematically talented students: A longitudinal study. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

(1987b) Possible biological correlates of precocious mathematical reasoning ability. *Trends in Neuroscience* 10:17–20.

(1988) Neuropsychological perspectives on mathematical talent. In: *The exceptional brain: Neuropsychology of talent and special abilities,* ed. L. K. Obler & D. Fein. Guilford Press.

(in preparation) Factors associated with high achievement in mathematics and science among mathematically talented students: A longitudinal study.

Benbow, C. P. & Benbow, R. M. (1984) Biological correlates of high mathematical reasoning ability. In: *Sex differences in the brain – The relation between structure and function. Progress in Brain Research,* vol. 61, ed. G. J. DeVries, J. P. C. DeBruin, H. B. M. Uylings & M. A. Corner. Elsevier Press.

(1987) Extreme mathematical talent: A hormonally induced ability? In: Duality and unity of the brain (Vol. 47 of *Wenner-Gren Center international symposium series*), ed. D. Ottoson. Macmillan.

Benbow, C. P., Hoff, E. & O'Boyle, M. (in preparation) Difference in brain organization between extremely intellectually talented students and average-ability students.

Benbow, C. P. & Minor, L. L. (1986) Mathematically talented students and achievement in the high school sciences. *American Educational Research Journal* 23:425–36.

Benbow, C. P. & Stanley, J. C. (1980) Sex differences in mathematical ability: Fact or artifact? *Science* 210:1262–64.

(1981) Mathematical ability: Is sex a factor? (Letter) *Science* 212:118–19.

(1982a) Consequences in high school and college of sex differences in mathematical reasoning ability: A longitudinal perspective. *American Educational Research Journal* 19:598–622.

(1982b) Intellectually talented boys and girls: Educational profiles. *Gifted Child Quarterly* 26:82–88.

(1983a) Differential course taking hypothesis revisited. (A commentary) *American Educational Research Journal* 20:469–73.

(1983b) Sex differences in mathematical reasoning ability: More facts. *Science* 222:1029–31.

(1984) Gender and the science major. In: *Advances in motivation and achievement*, vol. 2, ed. M. W. Steinkamp & M. L. Maehr. JAI Press.

Benbow, C. P., Stanley, J. C., Zonderman, A. B. & Kirk, M. K. (1983) Structure of intelligence of intellectually precocious children and of their parents. *Intelligence* 7:129–52.

Berrebi, A. S., Fitch, R. H., Ralphe, D. L., Denenberg, J. O., Friedrich, V. L., Jr. & Denenberg, V. H. (1988) Corpus callosum: Region-specific effects of sex, early experience, and age. *Brain Research* 438:216–24

Besag, F. D. (1987) Gender, race, and ethnicity and math avoidance. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

Bieri, J., Bradburn, W. & Gallinsky, M. (1958) Sex differences in perceptive behavior. *Journal of Personality* 26:1–12.

Block, N. J. & Dworkin, G. (1974) IQ: Heritability and inequality, Part 1. *Philosophy and Public Affairs* 3:331–408.

Bloom, B. S., ed. (1956) *Taxonomy of educational objectives. Handbook I: The cognitive domain.* David McKay.

Boli, J., Allen, M. L. & Payne, A. (1985) High-ability women and men in undergraduate mathematics and chemistry courses. *American Educational Research Journal* 22:605–26.

Boller, F. & Grafman, J. (1985) Acalculia. In: *Handbook of clinical neurology.* Vol. 45: *Clinical neuropsychology*, ed. P. J. Vinken, G. W. Bruyn & H. L. Klawans. Elsevier.

Bornstein, R. A. (1984) Unilateral lesions and the Wechsler Adult Intelligence Scale-Revised: No sex differences. *Journal of Consulting and Clinical Psychology* 52:604–8.

Boswell, S. L. (1985) The influence of sex-role stereotyping on women's attitudes and achievement in mathematics. In: *Women and mathematics: Balancing the equation*, ed. S. F. Chipman, L. R. Brush & D. M. Wilson. Erlbaum.

Bradshaw, J. L. & Nettleton, N. C. (1981) The nature of hemispheric specialization in man. *Behavioral and Brain Sciences* 4:51–91.

(1983) *Human cerebral asymmetry.* Prentice-Hall.

Brody, L. E. (1985) The effects of an intensive summer program on the SAT scores of gifted seventh graders. Unpublished doctoral dissertation, The Johns Hopkins University.

Broverman, D. M., Broverman, I. K., Vogel, W., Palmer, R. D. & Klaiber, E. L. (1964) The automatization cognitive style and physical development. *Child Development* 35:1343–59.

Brush, L. R. (1978) A validation study of the mathematics anxiety rating scale (MARS). *Educational and Psychological Measurement* 38:485–90.

(1980) *Encouraging girls in mathematics: The problem and the solution.* Abt Books.

Bryden, M. P. (1986) Dichotic listening performance, cognitive ability, and cerebral organization. *Canadian Journal of Psychology* 40:445–56.

Bryden, M. P., Hécaen, H. & DeAgostini, M. (1983) Patterns of cerebral organization. *Brain and Language* 20:249–62.

Burnett, S. A. (1986) Sex-related differences in spatial ability: Are they trivial? *American Psychologist* 41:1012–13.

Burnett, S. A., Lane, D. M. & Dratt, L. M. (1979) Spatial visualization and sex differences in quantitative ability. *Intelligence* 3:345–54.

(1982) Spatial ability and handedness. *Intelligence* 6:57–68.

Burnett, S. A. & Peters, D. (1985) Spatial ability and visual memory. Paper presented at the annual meeting of the Southwestern Psychological Association, San Antonio, Texas.

Byne, W., Bleier, R. & Houston, L. (1986) Individual, age and sex-related differences in human corpus callosum: A study using magnetic resonance imaging. *Society for Neuroscience Abstracts* 12:719.

Caplan, P. J., MacPherson, G. M. & Tobin, P. (1985) Do sex-related differences in spatial abilities exist? *American Psychologist* 40:786–89.

Caramazza, A., Gordon, J., Zuriff, G. E. & DeLuca, D. (1976) Right-hemispheric damage and verbal problem solving behavior. *Brain and Language* 3:41–46.

Carey, G. L. (1958) Sex differences in problem-solving performance as a function of attitude differences. *Journal of Abnormal and Social Psychology* 56:256–60.

Center for the Advancement of Academically Talented Youth (1983) The CTY 1983 Talent Search Report. Baltimore: CTY.

(1984) The CTY 1984 Talent Search Report. Baltimore: CTY.

(1985) The CTY 1985 Talent Search Report. Baltimore: CTY.

(1986) The CTY 1986 Talent Search Report. Baltimore: CTY.

Chi, J. G., Dooling, D. C. & Gilles, F. H. (1977) Gyral development of the human brain. *Annals of Neurology* 1:86–93.

Chi, M. T. H. (1978) Knowledge structures and memory development. In: *Children's thinking: What develops?* ed. R. S. Siegler. Erlbaum.

Chipman, S. F., Brush, L. R. & Wilson, D. M., eds. (1985) *Women and mathematics: Balancing the equation.* Erlbaum.

Chipman, S. F. & Thomas, V. G. (1985) Women's participation in mathematics: Outlining the problem. In: *Women and mathematics: Balancing the equation*, ed. S. F. Chipman, L. R. Brush & D. M. Wilson. Erlbaum.

(in press) The participation of women and minorities in mathematical, scientific, and technical fields. *Review of Research in Education*, vol. 14.

Chipman, S. F. & Wilson, D. M. (1985) Understanding mathematics course enrollment and mathematics achievement: A synthesis of the research. In: *Women and mathematics: Balancing the equation*, ed. S. F. Chipman, L. R. Brush & D. M. Wilson. Erlbaum.

Clark, A. S., MacLusky, N. J. & Goldman-Rakic, P. S. (1986) Steroid hormone binding and metabolism in the cerebral cortex and limbic structures of the developing rhesus monkey brain. *Society for Neuroscience* 12:1217 (Abstract).

Clark, R. W. (1971) *Einstein: The life and times.* World Publishing.

Cockroft, W. H. (1980) *Mathematics counts: Report of the committee of inquiry into teaching of mathematics.* Her Majesty's Stationery Office.

Cohen, J. (1977) *Statistical power analysis for the behavioral sciences.* Academic Press.

Cohn, S. J. (1977) Cognitive characteristics of the top-scoring participants in SMPY's 1976 Talent Search. *Gifted Child Quarterly* 22:416–21.

Cohn, S. J., Cohn, C. M. G. & Jensen, A. R. (in press) Speed of information processing and myopia in academically gifted youths and their siblings. *Intelligence.*

College Board (1985) *College-bound seniors, 1984–85.* Princeton, N.J.: College Board.

Connor, J. M. & Serbin, L. A. (1985) Visual-spatial skill: Is it important for mathematics? Can it be taught? In: *Women and mathematics: Balancing the equation*, ed. S. F. Chipman, L. R. Brush & D. M. Wilson. Erlbaum.

Cooper, H. M. (1984) *The integrative research review: A systematic approach.* Sage.

Coren, S., Searleman, A. & Porac, C. (1986) Rate of physical maturation and handedness. *Developmental Neuropsychology* 2:17–23.

D'Amico, C. & Kimura, D. (1987) Different subgroups of adextrals based on speech lateralization and ability patterns. *Journal of Clinical and Experimental Neuropsychology* 9:278 (Abstract).

Datta, L. (1985) Foreword. In: *Women and mathematics: Balancing the equation*, ed. S. F. Chipman, L. R. Brush & D. M. Wilson. Erlbaum.

Davis, J. A. (1965) Undergraduate career decisions: Correlates of academic choice. *NORC Monograph in Social Research*, No. 2.

Dawson, J.L.M.B. (1977) An anthropological perspective on the evolution and lateralization of the brain. In: *Evolution and lateralization of the brain*, ed. S. J. Dimond & D. A. Blizard. Annals of the New York Academy of Sciences, vol. 299.

Deaux, K. (1985) Sex and gender. *Annual Review of Psychology* 36:49–81.

de Lacoste-Utamsing, C. & Holloway, R. L. (1982) Sexual dimorphism in the human corpus callosum. *Science* 216:1431–32.

Demeter, S., Ringo, J. & Doty, R. W. (1985) Sexual dimorphism in the human corpus callosum? *Society for Neuroscience Abstracts* 11:868.

DeRenzi, E. (1982) *Disorders of space exploration and cognition.* Wiley.

Deutsch, D. (1980) Handedness and memory for tonal pitch. In: *Neuropsychology of left-handedness*, ed. J. Herron. Academic Press.

DeVries, G. J., DeBruin, J. P. C., Uylings, H. B. M. & Corner, M. A., eds. (1984) *Sex differences in the brain – the relation between structure and function. Progress in brain research*, vol. 26. Elsevier Press.

(1984a) *Sex differences in the brain. Progress in brain research*, vol. 61. Elsevier Press.

Diamond, M. C., Dowling, G. A. & Johnson, R. E. (1981) Morphologic cerebral cortical asymmetry in male and female rats. *Experimental Neurology* 71:261–68.

Diamond, M. C., Murphy, G. M., Jr., Akiyama, K. & Johnson, R. E. (1982) Morphological hippocampal asymmetry in male and female rats. *Experimental Neurology* 76:553–66.

Diamond, S. J. & Beaumont, J. G. (1974) *Hemisphere function in the human brain.* Paul Elek.

Dix, L. (1987) *Women: Their underrepresentation and career differentials in science and engineering.* National Research Council.

Donlon, T. F. (1973) Content factors in sex differences in test questions. *Educational Testing Service Research Monograph.*

Donlon, T. F., Ekstrom, R. B. & Lockheed, M. (1976) Comparing the sexes on achievement test items of varying content. Paper presented at the annual meeting of the American Psychological Association, Washington, D.C.

Donovan, B. T. & van der Werff ten Bosch, J. J. (1965) *Physiology of puberty*. Arnold.

Dornbusch, S. (1974) To try or not to try. *Stanford Magazine* 2:50–54.

Dougherty, K., Herbert, M., Edenhart-Pepe, M. & Small, A. (1980) Sex-related differences in mathematics, grades 2–5. Unpublished manuscript.

Dreger, R. M. & Aiken, L. R., Jr. (1957) The identification of number anxiety in a college population. *Journal of Educational Psychology* 48:344–51.

Dunphy, E. B.(1970) The biology of myopia. *New England Journal of Medicine* 283:796–800.

Dunteman, G. H., Wisenbaker, J. & Taylor, M. E. (1979) Race and sex differences in college science program participation. Report submitted to the National Science Foundation under contract NO. SED77-18728.

Dutton, W. H. (1956) Attitudes of junior high school pupils toward arithmetic. *School Review* 64:18–22.

Dutton, W. H. & Blum, M. P. (1968) The measurement of attitudes toward arithmetic with a Likert-type test. *Elementary School Journal* 68:259–64.

Dweck, C. S. (1986) Motivational problems affecting learning. *American Psychologist* 41:1040–48.

Dweck, C. S., Davidson, W., Nelson, S. & Enna, B. (1978) Sex differences in learned helplessness; II. The contingencies of evaluative feedback in the classroom; and III. An experimental analysis. *Developmental Psychology* 14:268–76.

Dwyer, C. A. (1974) Influence of children's sex-role standards on reading and arithmetic achievement. *Journal of Educational Psychology* 66:811–16.

Eaton, W. D. & Enns, L. R. (1986) Sex differences in human activity level. *Psychological Bulletin* 100:19–28.

Eccles, J. (1985) Sex differences in achievement patterns. In: *Nebraska symposium on motivation, 1984: Psychology and gender*, ed. T. B. Sonderegger. University of Nebraska Press.

Eccles, J., Adler, T. & Meece, J. (1984) Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology* 46:26–43.

Ehrhardt, A. A. & Baker, S. W. (1974) Fetal androgens, human CNS differentiation, and behavior sex differences. In: *Sex differences in behavior*, ed. R. C. Friedman, R. M. Richart & R. L. Van de Wiele. Wiley.

Eisenberg, L. (1986) Mindlessness and brainlessness in psychiatry. *British Journal of Psychiatry* 148:497–508.

Ellis, T. H. N. (1986) Restriction fragment length polymorphism markers in relation to quantitative characters. *Theoretical and Applied Genetics* 72:1–2.

Ernest, J. (1976) Mathematics and sex. *American Mathematical Monthly* 83:595–612.

Ethington, C. A. & Wolfle, L. M. (1984) Sex differences in a causal model of mathematics achievement. *Journal for Research in Mathematics Education* 15:361–77.

(1986) A structural model of mathematics achievement for men and women. *American Educational Research Journal* 23:65–75.

Fairweather, H. (1976) Sex differences in cognition. *Cognition* 4:231–80.

Farmer, H. S. (1987) A multivariate model for explaining gender differences in career and achievement motivation. *Educational Researcher* 16(2):5–9.

Fennema, E. (1974) Mathematics learning and the sexes: A review. *Journal for Research in Mathematics Education* 5:126–39.

(1976) Influences of selected cognitive, affective, and educational variables on sex-related differences in mathematics learning. Prepared for National Institute of Education.

Fennema, E. & Sherman, J. (1977) Sex-related differences in mathematics achievement, spatial visualization, and sociocultural factors. *American Educational Research Journal* 14:51–71.

(1978) Sex-related differences in mathematics achievement and related factors: A further study. *Journal for Research in Mathematics Education* 9:189–203.

Fennema, E. & Tartre, L. A. (1985) The use of spatial visualization in mathematics by girls and boys. *Journal for Research in Mathematics Education* 16:184–206.

Ferry, G. (1982) How women figure in science. *New Scientist* 94:10–13.

Fischbein, S. (1981) Heredity-environment influences of growth and development during adolescence. In: *Twin research 3- Intelligence, personality, and development*, ed. L. Gedda et al. Alan R. Liss.

Fitch, R. H., Berrebi, A. S. & Denenberg, V. H. (1987) Corpus callosum: Masculinized via perinatal testosterone. *Society for Neuroscience Abstracts* 13:688.

Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorham, W. A., Orr, D. B. & Goldberg, I. (1962) Design for a study of American youth. Houghton Mifflin.

Fleming, D. E., Anderson, R. H., Rhees, R. W., Kinghorn, E. & Bakaitis, J. (1986) Effects of prenatal stress on sexually dimorphic asymmetries in the cerebral cortex of the male rat. *Brain Research Bulletin* 16:395–8.

Fox, L. H. (1975) Career interests and mathematical acceleration for girls. Paper presented at annual meeting of American Psychological Association, Chicago.

(1976) Sex differences in mathematical precocity: Bridging the gap. In: *Intellectual talent: Research and development*, ed. D. P. Keating. The Johns Hopkins University Press.

(1977) The effects of sex-role socialization on mathematics participation and achievement. In: *Women and mathematics: Research perspectives for change*, ed. J. Schoemaker. National Institute of Education Papers in Education and Work, No. 8. Washington, D.C.: Education and Work Group, The National Institute of Education, United States Department of Health, Education, and Welfare.

(1982) Sex differences among the mathematically gifted. Paper presented at the annual meeting of the American Association for the Advancement of Science, Washington, D.C.

Fox, L. H., Benbow, C. P. & Perkins, S. (1983) An accelerated mathematics program for girls: A longitudinal evaluation. In: *Academic precocity: Aspects of its development*, ed. C. P. Benbow & J. C. Stanley. The Johns Hopkins University Press.

Fox, L. H., Brody, L. & Tobin, D. (1979) Women and mathematics: The impact of early intervention upon course-taking and attitudes in high school. Final report to the National Institute of Education, Grant No. NIE 6-77-0062.

eds. (1980) *Women and the mathematical mystique*. The Johns Hopkins University Press.

(1982) The study of social processes that inhibit or enhance the development of competence and interest in mathematics among highly able young women. Report to the National Institute of Education.

(1985) The impact of early intervention programs upon course-taking and attitudes in high school. In: *Women and mathematics: Balancing the equation*, ed. S. F. Chipman, L. R. Brush & D. M. Wilson. Erlbaum.

Fox, L. H. & Richmond, L. J. (1979) Gifted females: Are we meeting their counseling needs? *Personnel and Guidance Journal* 57:256–59.

Fox, L. H., Tobin, D. & Brody, L. (1979) Sex-role socialization and achievement in mathematics. In: *Sex-related differences in cognitive functioning: Developmental issues*, ed. M. Wittig & A. Petersen. Academic Press.

Fraser, E. J. P. & Cormack, S. (in press) Women and mathematics: A different picture. *Teaching mathematics and its applications*.

Freed, N. H. (1983) Forseeable equivalent math skill of men and women. *Psychological Reports* 52:334.

Freter, S. H., Mikolowski, A., Fitch, R. H., Berrebi, A. S., Yutzey, D. A. & Dennenberg, V. H. (1987) Handling stimulation in infancy and gender: Effects upon maze learning and behavioral asymmetry. *Society for Neuroscience Abstracts* 13.

Friedman, A. & Polson, M. C. (1981) Hemispheres as independent resource systems: Limited-capacity processing and cerebral specialization. *Journal of Experimental Psychology: Human Perception and Performance* 7:1031–58.

Frieze, I., Whiteley, B., Hanusa, B. & McHugh, M. (1982) Assessing the theoretical models for sex differences in causal attributions for achievement and failure. *Sex Roles: A Journal of Research* 8:333–44.

Garai, J. E. & Scheinfeld, A. (1968) Sex differences in mental and behavioral traits. *Genetic Psychology Monographs* 77:169–29.

Gardner, H. (1983) *Frames of mind*. Basic Books.

Gardner, H., Brownell, H. H., Wapner, W. & Michelow, D. (1983) Missing the point: The role of the right hemisphere in the processing of complex linguistic materials. In: *Cognitive processing in the right hemisphere*, ed. E. Perecman. Academic Press.

Garfinkel, A. S. (1982) Genetic and environmental influences on the development of Piagetian logico-mathematical concepts and other specific cognitive abilities: A twin study. *Acta Geneticae Medicae et Gemellologiae* 31:10–61.

Garvey, C. (1977) *Play*. Harvard University Press.

Geschwind, N. (1982) Why Orton was right. *Annals of Dyslexia* 32:13–30.

Geschwind, N. & Behan, P. (1982) Left-handedness: Association with immune disease, migraine, and developmental learning disorder. *Proceedings of the National Academy of Sciences* 79:5097–5100.

(1984) Laterality, hormones, and immunity. In: *Cerebral dominance: The biological foundations*, ed. N. Geschwind & A. Galaburda. Harvard University Press.

Geschwind, N. & Galaburda, A. M. (1986) *Cerebral lateralization*. MIT Press.

Gilligan, C. (1982) *In a different voice: Psychological theory and women's development*. Harvard University Press.

Ginsburg, H. P. (1983) (Ed.) *The development of mathematical thinking.* Academic Press.

Glass, G. V., McGaw, B. & Smith, M. L. (1981) *Meta-analysis in social research.* Sage.

Glennon, V. J. & Callahan, L. G. (1968) A guide to current research: Elementary school mathematics. Washington, D.C.: Association for Supervisional Curriculum Development.

Goldman, P. S. & Brown, R. (1975) The influence of neonatal androgen on the development of cortical function in the rhesus monkey. *Society for Neuroscience* 1:494 (Abstract).

Goldman, P. S., Crawford, H. T., Stokes, L. P., Galkin, T. W. & Rosvold, H. E. (1974) Sex-dependent behavioral effects of cerebral cortical lesions in the developing rhesus monkey. *Science* 186:540–42.

Goldman, P. S. & Rakic, P. T. (1979) Impact of the outside world upon the developing primate brain. *Bulletin of the Menninger Clinic* 43:20–28.

Goldman-Rakic, P. S. (1987) Circuitry of the primate prefrontal cortex and the regulation of behavior by representational memory. In: *Handbook of physiology*, ed. V. Mountcastle & F. Plum, and *American Physiological Society* 5:373–417.

Goleman, D. (1987) Girls and math: Is biology really destiny? *New York Times Education Supplement:* 42–43.

Gordon, H. W. & Galatzer, A. (1980) Cerebral organization in patients with gonadal dysgenesis. *Psychoneuroendocrinology* 5:235–44.

Gorski, R. A., Gordon, J. H., Shyrne, J. E. & Southam, A. M. (1978) Evidence for a morphological sex difference within the medial preoptic area of the rat brain. *Brain Research* 148:333–46.

Goy, R. W. & McEwen, B. S. (1981) *Sexual differentiation of the brain.* MIT Press.

Graf, G. & Ruddell, J. (1972) Sex differences in problem-solving as a function of problem content. *Journal of Educational Research* 65(10):451–52.

Green, P. R. (1980) Mechanical considerations in myopia: Relative effects of accommodation, convergence, intraocular pressure, and the extraocular muscles. *American Journal of Optometry and Physiological Optics* 57:902–14.

Hahn, W. K. (1987) Cerebral lateralization of function: From infancy through childhood. *Psychological Bulletin* 101:376–92.

Hall, K. R. L. & DeVore, I. (1965) Baboon social behavior. In: *Primate behavior: Field studies of monkeys and apes*, ed. I. DeVore. Holt, Rinehart & Winston.

Halpern, D. F. (1986a) *Sex differences in cognitive abilities.* Erlbaum.

(1986b) A different answer to the question "Do sex differences in spatial abilities exist?" *American Psychologist* 41:1014–15.

Hampson, E. & Kimura, D. (in press) Reciprocal effects of hormonal fluctuations on human motor and perceptuo-spatial skills. *Behavioral Neuroscience.*

Hardyck, C. (1977) A model of individual differences in hemispheric functioning. In: *Studies in neurolinguistics*, vol. 3, ed. H. Whitaker & H. A. Whitaker. Academic Press.

Hardyck, C., Petrinovich, L. F. & Goldman, R. F. (1976) Left-handedness and cognitive deficit. *Cortex* 12:266–79.

Harris, L. J. (1978) Sex differences in spatial ability: Possible environmental, genetic, and neurological factors. In: *Asymmetrical function of the brain*, ed. M. Kinsbourne. Cambridge University Press.

(1980) Left-handedness: Early theories, facts, and fancies. In: *Neuropsychology of left-handedness*, ed. J. Herron. Academic Press.

Harris, M. J. & Rosenthal, R. (1985) Mediation of interpersonal expectancy effects: Thirty-one meta-analyses. *Psychological Bulletin* 97:363–86.

Harshman, R. A. & Hampson, E. (1987) Normal variation in human brain organization: Relation to handedness, sex and cognitive abilities. In: *Duality and unity of the brain* (vol. 47 of *Wenner-Gren Center International Symposium Series*), ed. D. Ottoson. Macmillan.

Harshman, R. A., Hampson, E. & Berenbaum, S. A. (1983) Individual differences in cognitive abilities and brain organization. Part I: Sex and handedness differences in ability. *Canadian Journal of Psychology* 37:144–92.

Haven, E. W. (1971) Factors associated with the selection of advanced academic mathematics courses by girls in high school. Doctoral dissertation, University of Pennsylvania. *Dissertation Abstracts International* 32:1747A.

Heath, A. C., Berg, K., Eaves, L. J., Solast, M., Corey, L. A., Sundet, J., Magnus, P. & Nance, W. E. (1985) Education policy and the heritability of educational attainment. *Nature* 314:734–35.

Hedges, L. V. & Olkin, I. (1985) *Statistical methods for meta-analysis.* Academic Press.

Heller, K. A. & Parsons, J. E. (1981) Sex differences in teachers' evaluative feedback and students' expectancies for success in mathematics. *Child Development* 52:1015–19.

Herring, S. & Reitan, R. M. (1986) Sex similarities in verbal and performance IQ deficits following unilateral cerebral lesions. *Journal of Consulting and Clinical Psychology* 54:537–41.

Hier, D. B. & Crowley, W. F., Jr. (1982) Spatial ability in androgen deficient men. *New England Journal of Medicine* 306:1201–05.

Hilton, T. L., Beaton, A. E. & Bower, C. P. (1971) Stability and instability in academic growth – a compilation of longitudinal data. Educational Testing Service.

Hilton, T. L. & Berglund, G. W. (1974) Sex differences in mathematics achievement: A longitudinal study. *Journal of Educational Research* 67:231–37.

Hines, M. (1982) Prenatal gonadal hormones and sex differences in human behavior. *Psychological Bulletin* 92:56–80.

Hines, M. & Gorski, R. A. (1985) Hormonal influences on the development of neural asymmetries. In: *The dual brain*, ed. D. F. Benson & E. Zaidel. Guilford Press.

Hines, M. & Shipley, C. (1984) Prenatal exposure to diethylstilbestrol (DES) and the development of sexually dimorphic cognitive abilities and cerebral lateralization. *Developmental Psychology* 20:81–94.

Holden, C. (1987) Female math anxiety on the wane. *Science* 236:600–601.

Holloway, R. L. & De Lacoste, M. C. (1986) Sexual dimorphism in the human corpus callosum: An extension and replication study. *Human Neurobiology* 5:87–91.

Hubbard, R., Henifin, M. S. & Fried, B. (1982) *Biological woman: The convenient myth.* Schenkman.

Humphreys, L. G., Davey, T. C. & Kashima, E. (1986) Experimental measures of cognitive privilege/deprivation and some of their correlates. *Intelligence* 10:355–76.

Hungerman, A. D. (1967) Achievement and attitude of sixth-grade pupils in conventional and contemporary mathematics programs. *Arithmetic Teacher* 14:30–39.

Hunter, J. E., Schmidt, F. L. & Jackson, G. B. (1982) *Meta-analysis.* Sage.

Husén, T. (1959) *Psychological twin research 1: A methodological study.* Almquist & Wiskells.

(1960) Abilities of twins. *Scandinavian Journal of Psychology* 1:125–35.

ed. (1967) *International study of achievement in mathematics*, vol. 2. Wiley.

Hyde, J. S. (1981) How large are cognitive gender differences? A meta-analysis using w^2 and *d*. *American Psychologist* 36:892–901.

Jackson, R. (1980) The Scholastic Aptitude Test: A reply to Slack and Porter's "Critical Appraisal." *Harvard Educational Review* 50:282–391.

Jensen, A. R. (1969) How much can we boost IQ and scholastic achievement? *Harvard Educational Review* Reprint Series No. 2:1–123.

Johnson, M. M. (1963) Sex-role learning in the nuclear family. *Child Development* 34:319–24.

Johnson, N. L. & Kotz, S. (1970) *Distributions in statistics: Continuous univariate distributions-1.* Wiley.

Jones, L. V. (1987) The influence on mathematics test scores, by ethnicity and sex, of prior achievement and high school mathematics courses. *Journal for Research in Mathematics Education* 18:180–86.

Just, M. A. & Carpenter, P. A. (1985) Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review* 92:137–72.

Kail, R., Carter, P. & Pellegrino, J. (1979) The locus of sex differences in spatial ability. *Perception and Psychophysics* 26:182–86.

Karlsson, J. L. (1973) Genetic relationship between giftedness and myopia. *Hereditas* 73:85–86.

(1975) Influence of the myopia gene on brain development. *Clinical Genetics* 8:314–16.

Keating, D. P. (1974) The study of mathematically precocious youth. In: *Mathematical talent: Discovery, description, and development*, ed. J. C. Stanley, D. P. Keating & L. H. Fox. The Johns Hopkins University Press.

Keating, D. P. & Stanley, J. C. (1972) Extreme measures for the exceptionally gifted in mathematics and science. *Educational Researcher* 1(9):3–7.

Keeves, J. (1973) Differences between the sexes in mathematics and science courses. *International Review of Education* 19:47–63.

Keller, E. F. (1985) *Reflections on gender and science.* Yale University Press.

Kelly, A. (1979) *Girls and science: An international study of sex differences in school science achievement.* Stockholm: Almquist & Wiksell.

Kelly, D. D. (1985) Sexual differentiation of the nervous system. In: *Principles of neural science*, ed. E. R. Kandel & J. H. Schwartz. Elsevier Press.

Kenrick, D. T. (1987) Gender, genes, and the social environment: A biosocial interactionist perspective. In: *Review of personality & social psychology.* Vol. 8: *Sex and gender*, ed. P. C. Shaver & C. Hendrick. Sage Press.

Kessen, W. (1965) *The child.* Basic Books.

Kimura, D. (1987) Are men's and women's brains really different? *Canadian Psychology* 28:133–47.

Kinsbourne, M. (1974) Mechanisms of hemispheric interaction in man. In: *Hemispheric cerebral function*, ed. M. Kinsbourne & L. Smith. Charles Thomas.

Kolata, G. B. (1980) Math and sex: Are girls born with less ability? *Science* 210:1262.

(1983) Math genius may have hormonal basis. *Science* 222:1312.

Kornbrot, D. E. (in press) Degree performance as a function of discipline studied, parental occupation, and gender. *Higher Education*.

Lansman, M., Donaldson, G., Hunt, E. & Yantis, S. (1982) Ability factors and cognitive processes. *Intelligence* 6:347–86.

Lawrence, G. (1984) A synthesis of learning style research involving the MBTI. *Journal of Psychological Type* 8:2–15.

Levy, J. (1976) Cerebral lateralization and spatial ability. *Behavior Genetics* 6:171–88.

Levy, J. & Gur, R. C. (1980) Individual differences in psychoneurological organization. In: *Neuropsychology of left-handedness*, ed. J. Herron. Academic Press.

Levy, J. & Reid, M. (1978) Variations in cerebral organization as a function of handedness, hand posture in writing, and sex. *Journal of Experimental Psychology: General* 107:119–44.

Lewis, R. S. & Kamptner, N. L. (1987) Sex differences in spatial task performance of patients with and without unilateral cerebral lesions. *Brain and Cognition* 6:142–52.

Light, R. J. & Pillemer, D. B. (1984) *Summing up: The science of reviewing research*. Harvard University Press.

Linn, M. C. & Peterson, A. C. (1985) Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development* 56:1479–98.

Loehlin, J. C. & Nichols, R. C. (1976) *Heredity, environment, and personality: A study of 850 sets of twins*. University of Texas Press.

Long, R. L. & Murray, M. E. (1982) Review: Neurophysiological aspects of developmental dyslexia. *Developmental and Behavioral Pediatrics* 3:2–6.

Luchins, E. H. & Luchins, A. S. (1980) Female mathematicians: A contemporary appraisal. In: *Women and the mathematical mystique*, ed. L. H. Fox, L. Brody & D. Tobin. The Johns Hopkins University Press.

Lyon, M. F. (1961) Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* 190:S.372–3.

Maccoby, E. E. (1966) *The development of sex differences*. Stanford University Press.

Maccoby, E. E., Doering, C. H., Jacklin, C. N. & Kraemer, H. (1979) Concentrations of sex hormones in umbilical-cord blood: Their relation to sex and birth order of infants. *Child Development* 50:632–42.

Maccoby, E. E. & Jacklin, C. N. (1974) *The psychology of sex differences*. Stanford University Press.

MacLusky, N. J., Naftolin, F. & Goldman-Rakic, P. S. (1986) Estrogen formation and binding in the cerebral cortex of the developing rhesus monkey. *Proceedings of the National Academy of Sciences* 83:513–16.

Marshall, S. P. (1983) Sex differences in mathematical errors: An analysis of distractor choices. *Journal for Research in Mathematics Education* 14:325–36.

Martinez, D. & Halpern, D. F. (1987) Effects of sex and handedness on cognitive abilities. Paper presented at the Sixty-sixth Annual Convention of the Western Psychological Association, Long Beach, Calif.

Mateer, C. A., Polen, S. B. & Ojemann, G. A. (1982) Sexual variation in cortical localization of naming as determined by stimulation mapping. *Behavioral and Brain Sciences* 5:310–11.

Mayer, R. E. (1983) *Thinking, problem solving, cognition*. Freeman.

(1985) Mathematical ability. In: *Human abilities: An information-processing approach*, ed. R. Sternberg. W. H. Freeman.

McCall, J. R. (1955) *Sex differences in intelligence: A comparative factor study*. The Catholic University of America Press.

McGee, M. G. (1979) Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin* 86:889–918.

McGlone, J. (1980) Sex differences in human brain asymmetry: A critical analysis. *Behavioral and Brain Sciences* 3:215–63.

McKeever, W. F., Seitz, K. S., Hoff, A. L., Marino, M. F. & Diehl, J. A. (1983) Interacting sex and familial sinistrality characteristics influence both language lateralization and spatial ability in right-handers. *Neuropsychologia* 21:661–68.

McKeever, W. F., VanDeventer, A. D. & Suberi, M. (1973) Avowed-, assessed-, and familial-handedness and differential processing of brief sequential and nonsequential visual stimuli. *Neuropsychologia* 11:235–38.

Mebert, C. J. & Michel, G. F. (1980) Handedness in artists. In: *Neuropsychology of left-handedness*, ed. J. Herron. Academic Press.

Meece, J. L., Parsons, J. E., Kaczala, C. M., Goff, S. B. & Futterman, R. (1982) Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin* 91:324–48.

Millar, M. J. & Whitaker, H. A. (1983) The right hemisphere's contribution to language. In: *Language functions and brain organization*, ed. M. J. Millar & H. A. Whitaker. Academic Press.

Mills, C. J. (1981) Sex roles, personality, and intellectual abilities in adolescents. *Journal of Youth and Adolescence* 10:85–112.

(1984) Sex differences in self-concept and self-esteem for mathematically precocious adolescents. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

(1985) Psychological type and learning style differences in a gifted and normative population of adolescents. Paper presented at the Northeastern Educational Researchers Organization meeting, Rockport, Maine.

Minor, L. L. & Benbow, C. P. (1986) Construct validity of the SAT-M: A comparative study of high school students and gifted seventh graders. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Mokros, J. R. & Koff, E. (1978) Sex-stereotyping of children's success in mathematics and reading. *Psychological Reports* 42:1287–93.

Money, J. (1964) Two cytogenetic syndromes: Psychologic comparisons. 1. Intelligence and specific-factor quotients. *Journal of Psychiatry Research* 2:223–31.

Money, J. & Lewis, V. (1966) IQ, genetics and accelerated growth: Adrenogenital syndrome. *Johns Hopkins Hospital Bulletin* 118:365–73.

Moore, E. G. J. & Smith, A. W. (1987) Sex and ethnic group differences in mathematics achievement: Results from the National Longitudinal Study. *Journal for Research in Mathematics Education* 18:25–36.

Moore, M. A., Dubas, J. S. & Newcombe, N. (1987) *Biological and experiential factors in sex-related differences in spatial ability*. Paper presented to the Society for Research in Child Development, Baltimore.

Mullen, B. & Rosenthal, R. (1985) *BASIC meta-analysis: Procedures and programs*. Erlbaum.

Mura, R. (1987) Sex-related differences in expectations of success in undergraduate mathematics. *Journal for Research in Mathematics Education* 18:15–24.

Nasrallah, H. A., Andreasen, N. C., Coffman, J. A., Olson, S. C., Dunn, V. D., Ehrhardt, J. C. & Chapman, S. M. (1986) A controlled magnetic resonance imaging study of corpus callosum thickness in schizophrenia. *Biological Psychiatry* 21:274–82.

Nass, R., Baker, S., Speiser, P., Virdis, R., Balsamo, A., Cacciari, E., Loche, A., Dumic, M. & New, M. (1987) Hormones and handedness – Left-hand bias in female congenital adrenal hyperplasia patients. *Neurology* 37:711–14.

National Assessment of Educational Progress (1975) Males dominate in educational success. *NAEP Newsletter* 8:insert.

(1979) Mathematical knowledge and skills: Selected results from the second assessment of mathematics. Denver, Colorado: Education Commission of the States.

National Commission for Excellence in Education (1983) *A nation at risk: The imperative for educational reform*. U.S. Department of Education.

National Institute of Education (1977) *Grants for research on education and work*. Washington, D.C.: National Institute of Education.

Nesher, P. (1986) Learning mathematics: A cognitive perspective. *American Psychologist* 41:1114–21.

Newcombe, N. (1982) Sex-related differences in spatial ability: Problems and gaps in current approaches. In: *Spatial abilities: Development and physiological foundations*, ed. M. Potegal. Academic Press.

Newcombe, N. & Dubas, J. S. (1987) Individual differences in cognitive ability: Are they related to timing of puberty? In: *Biological-psychosocial interactions in early adolescence*, ed. R. M. Lerner & T. T. Foch. Erlbaum.

Nicholls, J. (1980) A re-examination of boys' and girls' causal attributions for success and failure based on New Zealand data. In: *Achievement motivation: Recent trends in theory and research*, ed. L. Fyans. Plenum.

Nielsen, J., Nyborg, H. & Dahl, G. (1977) *Turner's syndrome. A psychiatric-psychological study of 45 women with Turner's syndrome, compared with their sisters and women with normal karyotype, growth retardation, and primary amenorrhoea*. Acta Jutlandica, Medicine Series 21. Aarhus University Press.

Nochlin, L. (1972) Why are there no great women artists? In: *Woman in sexist society*, ed. V. Gornick & B. Moran. New American Library.

Nordeen, E. J. & Yahr, P. (1982) Hemispheric asymmetries in the behavioral and hormonal effects of sexually differentiating mammalian brain. *Science* 218:391–93.

Nyborg, H. (1979) Sex chromosome abnormalities and cognitive performance. V. Female sex hormones and discontinuous cognitive development. Paper presented at the fifth biennial meeting of the International Society for the Study of Behavioural Development, Lund, Sweden.

(1983) Spatial ability in men and women: Review and new theory. *Advances in Behavioural Research and Therapy* 5:89–140.

(1984) Performance and intelligence in hormonally-different groups. In: *Sex differences in the brain – the relation between structure and function. Progress in brain research*, vol. 61, ed. G. J. DeVries, J. P. C. DeBruin, H. B. M. Uylings & M. A. Corner. Elsevier Press.

(1986) Sex chromosomes, sex hormones, and developmental disturbances: In search of a model. Paper presented at the 152nd Annual National Meeting of the American Association for the Advancement of Science, Philadelphia.

(1987) Covariant trait development across races, and within individuals. Paper presented at the third meeting of the International Society for the Study of Individual Differences, Toronto, Canada.

(submitted) Principles of sex hormonal regulation of body, brain and behavioral development.

Nyborg, H. & Nielsen, J. (1981) Sex hormone treatment and spatial ability in women with Turner's syndrome. In: *Human behavior and genetics*, ed. W. Schmid & J. Nielsen. Elsevier.

O'Boyle, M. W. & Hoff, E. J. (1987) Gender and handedness differences in mirror-tracing random forms. *Neuropsychologia* 25:977–82.

Oldfield, R. C. (1971) The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia* 9:97–113.

Oppenheim, J. S., Lee, B. C. P., Nass, R. & Gazzaniga, M. S. (1987) No sex-related differences in human corpus callosum based on magnetic resonance imagery. *Annals of Neurology* 21:604–5.

Paige, J. R. & Simon, H. (1966) Cognitive processes in solving algebra word problems. In: *Problem solving: Research, method, and theory*, ed. B. Kleinmuntz. Wiley.

Pallas, A. M. & Alexander, K. L. (1983) Sex differences in quantitative SAT performance: New evidence on the differential coursework hypothesis. *American Educational Research Journal* 20:165–82.

Parsons, J. E., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. & Midgley, C. (1983) Expectancies, values, and academic behavior. In: *Achievement and achievement motives: Psychological and sociological approaches*, ed. J. T. Spence. W. H. Freeman.

Parsons, J. E., Adler, T. F. & Kaczala, C. (1982) Socialization of achievement attitudes and beliefs: Parental influences. *Child Development* 53:310–21.

Parsons, J. E., Kaczala, C. & Meece, J. (1982) Socialization of achievement attitudes and beliefs: Classroom influences. *Child Developmnt* 53:322–39.

Pattison, P. & Grieve, N. (1984) Do spatial skills contribute to sex differences in different types of mathematical problems? *Journal of Educational Psychology* 76:678–89.

Paulsen, K. & Johnson, M. (1983) Sex-role attitudes and mathematical ability in 4th, 8th, and 11th grade students from a high socioeconomic area. *Developmental Psychology* 19:210–14.

Pederson, K., Elmore, P. & Bleyer, D. (1986) Parent attitudes and student career interests. *Journal for Research in Mathematics Education* 17:49–59.

Pedro, J. D., Wolleat, P., Fennema, E. & Becker, A. D. (1981) Election of high school mathematics by females and males: Attributions and attitudes. *American Educational Research Journal* 18:207–18.

Pellegrino, J. W. & Kail, R., Jr. (1982) Process analyses of spatial aptitude. In: *Advances in the psychology of human intelligence*, vol. 1, ed. R. J. Sternberg. Erlbaum.

Petersen, A. C. (1976) Physical androgeny and cognitive functioning in adolescence. *Developmental Psychology* 12:524–33.

Peterson, P. L. & Fennema, E. (1985) Effective teaching, student engagement in classroom activities, and sex-related differences in learning mathematics. *American Educational Research Journal* 22:309–36.

Petrusic, W. M., Varro, L. & Jamieson, D. G. (1978) Mental rotation validation of two spatial ability tests. *Psychological Research* 40:39–148.

Plomin, R. & Daniels, D. (1987) Why are children in the same family so different from each other? *Behavioral and Brain Sciences* 10:1–59.

Poffenberger, T. & Norton, D. (1959) Factors in the formation of attitudes toward mathematics. *Journal for Educational Research* 52:171–76.

Poltrock, S. E. & Brown, P. (1984) Individual differences in visual imagery and spatial ability. *Intelligence* 8:93–138.

Pomerantz, S. M., Fox, T., Sholl, S. A., Vito, C. C. & Goy, R. (1985) Androgen and estrogen receptors in fetal rhesus monkey brain and anterior pituitary. *Endocrinology* 116:83–89.

Porac, C. & Coren, S. (1981) *Lateral preferences and human behavior*. Springer-Verlag.

Porac, C., Coren, S. & Searleman, A. (1986) Environmental factors in hand preference formation: Evidence from attempts to switch the preferred hand. *Behavior Genetics* 16:251–61.

Posner, M. I., Boies, S. J., Eichelman, W. H. & Taylor, R. L. (1969) Retention of visual and name codes of single letters. *Journal of Experimental Psychology* 79:1–16.

Potter, S. (1950) *Lifemanship*. Henry Holt.

Ray, W. J., Newcombe, N., Semon, J. & Cole, P. M. (1981) Spatial abilities, sex differences and EEG functioning. *Neuropsychologia* 19:719–22.

Raymond, C. L. & Benbow, C. P. (1986) Gender differences in mathematics: A function of parental support and sex-typing? *Developmental Psychology* 22:808–19.

(submitted) Educational encouragement by parents: Its relationship to children's aptitude and perceptions and gender differences.

Read, D. B. (1981) Solving deductive reasoning problems after unilateral temporal lobectomy. *Brain and Language* 12:116–27.

Reinisch, J. M. (1984) Prenatal gonadal steroid influences on gender-related behavior. In: *Progress in brain research*, vol. 61, ed. G. J. De Vries, J. P. C. DeBruin, H. B. M. Uylings & M. A. Corner. Elsevier Press.

Reinisch, J. M. & Gandelman, R. (1978) Human research in behavioral endocrinology: Methodological and theoretical considerations. In: *Hormones and brain development*, ed. G. Dorner. Elsevier Press.

Reinisch, J. M., Gandelman, R. & Spiegel, F. S. (1979) Prenatal influences on cognitive abilities: Data from experimental animals and human and endocrine syndromes. In: *Sex-related differences in cognitive functioning: Developmental issues*, ed. M. Wittig & A. C. Petersen. Academic Press.

Reinisch, J. M. & Sanders, S. A. (1984) Prenatal gonadal steroidal influences on gender-related behavior. *Progress in Brain Research* 61:407–16.

Reis, S. M. (1987) We can't change what we don't recognize: Understanding the special needs of gifted females. *Gifted Child Quarterly* 31(2):83–89.

Reschly, D. J. & Jipson, F. J. (1976) Ethnicity, geographic locale, age, sex, and urban-rural residence as variables in the prevalence of mind retardation. *American Journal of Mental Deficiency* 81:154–61.

Resnick, L. B. & Ford, W. W. (1981) *The psychology of mathematics for instruction*. Erlbaum.

Resnick, S. M., Berenbaum, S. A., Gottesman, I. I. & Bouchard, T. J., Jr. (1986) Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology* 22:191–98.

Resnick, S. M. & Gottesman, I. I. (1978) Prenatal hormones and IQ: Need for sibling controls. *British Journal of Psychiatry* 132:205–6.

Robinson, H. B. (1983) A case for radical acceleration: Programs of The Johns Hopkins University and the University of Washington. In: *Academic precocity: Aspects of its development*, ed. C. Benbow & J. Stanley. The Johns Hopkins University Press.

Robitaille, D. F. (1977) A comparison of boys' and girls' feelings of self-confidence in arithmetic computation. *Canadian Journal of Education* 2:15–22.

Roland, P. E. & Friberg, L. (1985) Localization of cortical areas activated by thinking. *Journal of Neurophysiology* 53:1219–43.

Rose, R. M., Holaday, J. W. & Bernstein, I. (1971) Plasma testosterone, dominance rank, and aggressive behavior in male rhesus monkeys. *Nature* 231:366–68.

Roselli, C. & Resko, J. (1986) Effects of gonadectomy and androgen treatment on aromatase activity in the fetal monkey brain. *Biology of Reproduction* 35:106–12.

Rosenthal, K. & Kimura, D. (1987) Hormonal influences on cognitive ability patterns. *Department of Psychology Research Bulletin* No. 653, University of Western Ontario, London, Canada.

Rosenthal, R., ed. (1980) *Quantitative assessment of research domains*. (New Directions for Methodology of Social and Behavioral Science, No. 5.) Jossey-Bass.

(1984) *Meta-analytic procedures for social research*. Sage.

(1987) *Judgment studies: Design, analysis, and meta-analysis*. Cambridge University Press.

Rosenthal, R. & Rosnow, R. (1985) *Essentials of behavioral research: Methods and data-analysis*. McGraw-Hill.

Rosenthal, R. & Rubin, D. B. (1982) Further meta-analytic procedures for assessing cognitive gender differences. *Journal of Educational Psychology* 74:708–12.

Rossiter, M. W. (1982) *Women scientists in America*. The Johns Hopkins University Press.

Rovet, J. & Netley, C. (1982) Processing deficits in Turner's Syndrome. *Developmental Psychology* 18:77–94.

Rushton, J. P. (in press) Race differences in behaviour: A review and evolutionary analysis. *Personality and Individual Differences*.

Sanders, B., Soares, M. P. & D'Aquila, J. M. (1982) The sex difference on one test of spatial visualization: A nontrivial difference. *Child Development* 53:1106–10.

Sapolsky, R. M. (1987) The case of the falling nightwatchman. *Discover* 8:42–45.

Satz, P. & Soper, H. (1986) Left-handedness, dyslexia, and autoimmune disorder: A critique. *Journal of Clinical and Experimental Neuropsychology* 8:453–58.

Scarr, S. & Saltzman, L. C. (1982) Genetics and intelligence. In: *Handbook of human intelligence*. Cambridge University Press.

Schachter, S. C., Ransil, B. J. & Geschwind, N. (1987) Association of handedness with haircolor and learning disabilities. *Neuropsychologia* 25:269–76.

Schaub, W. (1971) Die genealogische Datenbank im Dienste der Wissenschaft. *Genealogie* 20:577–85, 630–38.

Schoenfeld, A. H. (1985) *Mathematical problem solving.* Academic Press.

Searleman, A., Herrman, D. J. & Coventry, A. K. (1984) Cognitive abilities and left-handedness: An interaction between familial sinistrality and strength of handedness. *Intelligence* 8:295–304.

Segalowitz, S. J. & Bryden, M. P. (1983) Individual differences in hemispheric representation of language. In: *Language functions and brain organization*, ed. S. J. Segalowitz. Academic Press.

Senk, S. & Usiskin, Z. (1983) Geometry proof writing: A new view of sex differences in mathematics ability. *Journal of Educational Research* 91:187–201.

Sergent, J. (1982) The cerebral balance of power: Confrontation or cooperation. *Journal of Experimental Psychology: Human Perception and Performance* 8:253–72.

Sherman, J. A. (1967) The problem of sex differences in space perception and aspects of individual functioning. *Psychological Review* 74:290–99.

(1974) Field articulation, sex, spatial visualization, dependency, practice, laterality of the brain and birth order. *Perceptual and Motor Skills* 38:1223–35.

(1977) Effects of biological factors on sex-related differences in mathematics achievement. *Women and mathematics: Research perspectives for change.* NIE papers in Education and Work: No. 8.

(1980) Mathematics, spatial visualization and related factors: Changes in girls and boys, grades 8–11. *Journal of Educational Psychology* 72:476–82.

Sherman, J. A. & Fennema, E. (1977) The study of mathematics by high school girls and boys: Related variables. *American Educational Research Journal* 14:159–68.

Siegler, R. J. (in press) Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development.*

Signorella, M. L. & Jamison, W. (1986) Masculinity, femininity, androgyny, and cognitive performance: A meta-analysis. *Psychological Bulletin* 100:207–28.

Smith, I. M. (1964) *Spatial ability.* The University of London Press.

Sofaer, J. & Emery, A. (1981) Genes for super-intelligence. *Journal of Medical Genetics* 18:410–13.

Sprafkin, C., Serbin, S. A., Denier, C. & Connor, J. M. (1983) Sex-differentiated play: Cognitive consequences and early interventions. In: *Social and cognitive skills*, ed. M. B. Liss. Academic Press.

Stafford, R. E. (1964) Sex differences in spatial visualization as evidence of sex-linked inheritance. *Perceptual and Motor Skills* 13:4–28.

Stanley, J. C. (1982) Identification of intellectual talent. In: *Measurement, guidance, and program improvement.* New directions for testing and measurement series, No. 13: 97–109.

(1987) Sex differences on cognitive tests: A summary. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

(1987a) Gender differences on the College Board Achievement Tests and the Advanced Placement Examinations: Effect sizes versus some upper-tail ratios. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

(1987b) Summary of points made in the symposium. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

Stanley, J. C. & Benbow, C. P. (1983) 15:1 isn't "catching up"! *Psychological Reports* 52:656.

(1986) Youths who reason exceptionally well mathematically. In: *Conceptions of giftedness*, ed. R. J. Sternberg & J. Davison. Cambridge University Press.

Stanley, J. C., Brody, L. E., Dauber, S. L., Lupkowski, A. E. & Benbow, C. P. (1987) Sex differences in cognitive abilities and achievements. A symposium presentation at the annual meeting of the American Educational Research Association, Washington, D.C.

Stanley, J. C., Huang, Jia-fen & Zu, Xue-min (1986) SAT-M scores of highly selected students in Shanghai tested when less than 13 years old. *College Board Review* 140:10–13, 28–29.

Steen, L. A. (1987) Mathematics education: A predictor of scientific competitiveness. *Science* 237:251

Stein, A. H. (1971) The effects of sex-role standards for achievement and sex-role preference on three determinants of achievement motivation. *Developmental Psychology* 4:219–31.

Stein, A. A. & Smithells, J. (1969) Age and sex differences in children's sex role standards about achievement. *Developmental Psychology* 1:252–59.

Stevenson, H. W., Lee, S.-Y. & Stigler, J. W. (1986) Mathematics achievement of Chinese, Japanese, and American children. *Science* 231:693–99.

Stright, V. M. (1960) A study of the attitudes toward mathematics of students and teachers in the third, fourth, and sixth grades. *Arithmetic Teacher* 7:280–86.

Sturt, E. & McGuffin, P. (1985) Can linkage and marker associations resolve the genetic aetiology of psychiatric disorders? Review and argument. *Psychological Medicine* 15:455–62.

Suydam, M. N. & Weaver, J. F. (1975) *Using research: A key to elementary school mathematics.* Columbus, Ohio: ERIC Science, Mathematics, and Environmental Education Clearinghouse in cooperation with the Center for Science and Mathematics.

Symons, D. (1979) *The evolution of human sexuality.* Oxford University Press.

(1987) The evolutionary approach: Can Darwin's view of life shed light on human sexuality? In: *Theories of human sexuality*, ed. J. H. Geer & W. T. O'Donohue. Plenum.

Tapley, S. M. & Bryden, M. P. (1977) An investigation of sex differences in spatial ability: Mental rotation of three-dimensional objects. *Canadian Journal of Psychology* 31:122–30.

Terman, L. M. (1954) Scientists and nonscientists in a group of 800 gifted men. *Psychological Monographs* 68(7):1–44.

Terman, L. M. & Oden, M. H. (1959) The gifted group at midlife. *Genetic studies of genius*, vol. 5. Stanford University Press.

Tiger, L. (1987) *The manufacture of evil: Ethics, evolution, and the industrial system.* Harper and Row.

Tiger, L. & Shepher, J. (1975) *Women in the kibbutz.* Harcourt, Brace, and Jovanovich.

Tobias, S. (1978) *Overcoming math anxiety.* Norton.

Tobias, S. & Weissbrod, C. (1980) Anxiety and mathematics: An update. *Harvard Educational Review* 50:63–70.

Tobin, D. (1986) A longitudinal study of values and career interests of mathematically gifted students. Unpublished doctoral dissertation, The Johns Hopkins University.

Toran-Allerand, D. (1986) Sexual differentiation of the brain. In: *Developmental neuropsychobiology*, ed. W. T. Greenough & J. M. Juraska. Academic Press.

Troup, G. A., Bradshaw, J. L. & Nettleton, N. C. (1983) The lateralization of arithmetic and number processing: A review. *International Journal of Neuroscience* 19:231–42.

Trune, D. (1982) Influence of neonatal cochlear removal on the development of mouse cochlear nucleus. I. Number, size, and density of its neurons. *Journal of Comparative Neurology* 209:409–24.

Turner, H. (1938) A syndrome of infantilism, congenital webbed neck, and cubitus valgus. *Endocrinology* 23:566–74.

University of London (1985) General Certificate of Education Examination: Statistics. University of London Press.

U.S. Department of Education (1986) *What works: Research about teaching and learning.* Washington, D.C.

U.S. Department of Justice, Federal Bureau of Investigation (1979) *Uniform Crime Reports of the United States, 1979.* U.S. Government Printing Office.

Vandenberg, S. G. & Kuse, A. R. (1979) Spatial ability: A critical review of the sex-linked major gene hypotheses. In: *Determinants of sex-related differences in cognitive functioning*, ed. M. C. Wittig & A. C. Petersen. Academic Press.

van Haaren, F. & van de Poll, N. E. (1984) The effect of a choice alternative on sex differences in passive avoidance behavior. *Physiology and Behavior* 32:211–16.

Van Hest, A., van Haaren, F. & van de Poll, N. E. (1987) Behavioral differences between male and female Wistar rats on DRL schedules: Effects of stimuli providing collateral activities. *Physiology and Behavior* 39:255–61.

Vellutino, F. R. (1979) *Dyslexia: Theory and research.* MIT Press.

Very, P. S. (1967) Differential factor structures in mathematical ability. *Genetic Psychology Monographs* 75(2):169–207.

Villardita, C. (1987) Verbal memory and semantic clustering in right brain-damaged patients. *Neuropsychologia* 25:277–80.

Waber, D. P. (1976) Sex differences in mental abilities: A function of maturation rate? *Science* 192:572–74.

(1977) Sex differences in mental abilities, hemispheric lateralization, and rate of physical growth at adolescence. *Developmental Psychology* 13:29–38.

Waber, D. P., Mann, M. B., Merola, J. & Moylan, D. M. (1985) Physical maturation rate and cognitive performance in early adolescence: A longitudinal evaluation. *Developmental Psychology* 21:666–81.

Walberg, H. J. (1983) Scientific literacy and economic productivity in international perspective. *Daedalus* 112:1–28.

(1984) Improving the productivity of America's schools. *Educational Leadership* 41(8):19–27.

Walberg, H. J. & Tsai, S.-L. (1983) Matthew effects in education. *American Educational Research Journal* 20:359–73.

Warrington, E. K. (1982) The fractionation of arithmetic skills: A single case study. *Quarterly Journal of Experimental Psychology* 34A:31–51.

Weber, G. & Weis, S. (1986) Morphometric analysis of the human corpus callosum fails to reveal sex-related differences. *Journal für Hirnforschung* 2:237–40.

Webster, D. & Webster, M. (1977) Neonatal sound deprivation affects brain stem auditory nuclei. *Archives of Otolaryngology* 103:392–96.

Weiss, V. (1982) *Psychogenetik*. Jena: Gustar Fischer.

Werts, C. E. (1967) Career changes in college. *Sociology of Education* 40:90–95.

Wiesel, T. N. (1982) Postnatal development of the visual cortex and the influence of environment. *Nature* 299:583–92.

Williams, D. A. (1980) Do males have a math gene? *Newsweek*, December 15.

Wilson, J. W. (1972) Patterns of mathematics achievement in grade 11: Z population. National Longitudinal Study of Mathematical Abilities, Report No. 17. Stanford, Calif: School Mathematics Study Group.

Wilson, M. & Daly, M. (1985) Competitiveness, risk-taking, and violence: The young male syndrome. *Ethology and Sociobiology* 6:59–73.

Wise, L. L. (1985) Project Talent: Mathematics course participation in the 1960s and its career consequences. In: *Women and mathematics: Balancing the equation*, ed. S. F. Chipman, L. R. Brush & D. M. Wilson. Erlbaum.

Wise, L. L., Steel, L. & MacDonald, C. (1979) *Origins and career consequences of sex differences in high school mathematics achievement*. American Institute for Research.

Witelson, S. F. (1976) Sex and the single hemisphere: Specialization of the right hemisphere for spatial processing. *Science* 193:425–27.

(1985) The brain connection: The corpus callosum is larger in left-handers. *Science* 229:665–68.

(1985a) An exchange on "gender" with R. C. Lewontin. *New York Review of Books* 32(October 24):53–55.

(1986) Wires of the mind: Anatomical variation in the corpus callosum in relation to hemispheric specialization and integration. *Neurology and Neurobiology* 17:117–38.

(1987a) Neurobiological aspects of language in children. *Child Development* 58:653–88.

(1987b) Hand preference and sex differences in the isthmus of the corpus callosum. *Society for Neuroscience Abstracts* 13:48.

Witelson, S. F. & Kigar, D. L. (1987) Individual differences in the anatomy of the corpus callosum: Sex, hand preference, schizophrenia and hemisphere specialization. In: *Individual differences in hemisphere specialization*, ed. A. Glass. NATO ASI Series, 130. Plenum.

(1988) Anatomical development of the human corpus callosum: A review with reference to sex and cognition. In: *Developmental implications of brain lateralization*, ed. D. L. Molfese & S. J. Segalowitz. Guilford Press.

Wolf, F. M. (1986) *Meta-analysis: Quantitative methods for research synthesis*. Sage.

Wolins, L. & Benbow, C. P. (in preparation) Item performance on the SAT-M: Differences by gender and between seventh graders and high school students.

Woolf, V. (1929) *A room of one's own*. Harcourt, Brace, Jovanovich.

Zaidel, E. (1978) Lexical organization in the right hemisphere. In: *Cerebral correlates of conscious experience*, ed. Buser and Rougeul-Buser. INSERM Symposium No. 6. Elsevier Press.

Zimmerman, W. (1984) The reliability and validity of the Bem Sex Role Inventory. Unpublished doctoral dissertation, The Johns Hopkins University.

Zolog, A. (1983) Speech disturbances in parietal lobe lesions. *Neurologie et Psychiatrie* 21(3):165–67.