American Educational Research Journal
Winter 1982, Vol. 19, No. 4, Pp. 598-622

Consequences in High School and College of Sex Differences in Mathematical Reasoning Ability: A Longitudinal Perspective

CAMILLA PERSSON BENBOW and JULIAN C. STANLEY
The Johns Hopkins University

Between 1972 and 1974 the Study of Mathematically Precocious Youth (SMPY) identified over 2,000 7th and 8th graders who scored as well as a national sample of 11th and 12th grade females on the College Board's Scholastic Aptitude Test (SAT) Mathematics or Verbal tests. A substantial sex difference in mathematical reasoning ability was found (Benbow & Stanley, 1980b, 1981). The consequences and development of this sex difference over the following 5 years were investigated longitudinally. Over 91 percent (1,996 out of 2,188 SMPY students) participated. This study established that the sex difference persisted over several years and was related to subsequent sex differences in mathematics achievement. The sex difference in mathematics did not reflect differential mathematics course taking. The abilities of males developed more rapidly than those of females. Sex differences favoring males were found in participation in mathematics, performance on the SAT-M, and taking of and performance on mathematics achievement and Advanced Placement Program examinations. SMPY females received better grades in their mathematics courses than SMPY males did. Few significant sex differences were found in attitudes toward mathematics.

The Study of Mathematically Precocious Youth (SMPY) was begun in 1971 to study and facilitate the education of mathematically precocious junior high school students. It was designed to be longitudinal, involving successive follow-ups throughout the adult lives of the students identified and helped by SMPY.

We thank Robert A. Gordon, Lynn H. Fox, David P. Baker, and the anonymous reviewers for helpful comments and suggestions concerning earlier drafts of this article.

One of SMPY's most controversial and unexpected findings was the large and consistent sex difference in mathematical reasoning ability favoring boys (Benbow & Stanley, 1980b, 1981). In six talent searches conducted over an 8-year interval in the mid-Atlantic states and involving almost 10,000 students, the gifted seventh or eighth grade boys, who had previously been matched with gifted seventh or eighth grade girls in mathematical ability on standardized achievement tests, scored substantially better than their female counterparts on a difficult test of mathematical reasoning ability. No sex differences were seen, however, on the equally difficult verbal reasoning test.

This finding was consistent with numerous other studies of sex differences in mathematical ability and achievement (Backman, 1972; Bieri, Bradburn, & Galinsky, 1958; Ernest, 1976; Fennema, 1974; Fox, 1976; Fox, Brody, & Tobin, 1980; Garai & Scheinfeld, 1968; Glennon & Callahan, 1968; Keating, 1974; Maccoby & Jacklin, 1974; National Assessment of Educational Progress, 1975; Suydam & Weaver, 1970; Very, 1967; Wilson, 1972). The significance of the Benbow and Stanley (1980b, 1981) results was that they found the sex difference at the seventh-grade level, when formal mathematics is similar for boys and girls, and were able to rule out the hypothesis (e.g., see Fennema & Sherman, 1977) that differential course taking by boys and girls in mathematics causes the sex difference in mathematical reasoning ability. The expected sex differences in attitudes toward mathematics were also not found by these investigators (Benbow & Stanley, 1980b, 1982), and now others (Fox, Brody, & Tobin, 1982). This, along with other observations, prompted the statement that it "seems likely that putting one's faith in boy-versus-girl socialization processes as the only permissible explanation of the sex difference in mathematics is premature" (Benbow & Stanley, 1980b, p. 1264).

This investigation studied the development and consequences during junior and senior high school of the sex difference found in the seventh grade among a certain subpopulation of the students studied by Benbow and Stanley (1980b). Of special interest was the determination of what additional sex differences emerged over this 5-year interval and how they related to the initial sex difference in mathematical reasoning ability at seventh grade. We hypothesized that sex differences in mathematics participation and achievement favoring boys would develop during the high school years and that they could be partly accounted for by the sex difference in mathematical reasoning ability found several years earlier. Moreover, it was hypothesized that the sex difference on SAT-M favoring males would increase during the high school years, partly because of environmental influences and because the mathematical reasoning ability of the boys might have been increasing at a faster rate all along than the girls' was.

SUBJECTS

The special subpopulation of subjects consisted of students from SMPY's first three talent searches. In those, seventh and eighth grade students (some *accelerated* 9th and 10th-graders were eligible also) attending schools in Maryland were eligible to participate if they scored in the upper 5 percent (March 1972) or the upper 2 percent (January-February 1973 or January 1974) in mathematical ability on the national norms for a standardized achievement test administered in the regular testing program of the students' schools. As part of the talent search itself, qualifying students then took the College Board's *Scholastic Aptitude Test-Mathematics* (SAT-M) and also, in 1973, the SAT-Verbal. These tests measure mathematical and verbal reasoning ability, respectively (Angoff, 1971; Messick & Jungeblut, 1981; also see Benbow & Stanley, 1981). The participants in the talent search also completed a background questionnaire. It has been shown that these students tend to come from homes where the parents had been rather highly educated (Benbow & Stanley, 1980a; Keating, 1974).

To be part of this study, which followed up the talent search participants at high school graduation, the student had to have scored at least 390 on SAT-M *or* 370 on SAT-V during the talent search as a 7th or 8th grader. These SAT criteria selected students who as 7th or 8th graders scored as well as the average 11th and 12th grade female does on SAT-M and SAT-V (Admission Testing Program [ATP], 1979).

A sample size of 2,188 was obtained of which approximately 61 percent were males. (In the initial talent searches, approximately 57 percent were males.) When the subjects were contacted (between 1976 and 1980), most were college freshmen. This was approximately 4 to 5 years after participation in one or more of SMPY's talent searches.

INSTRUMENTATION

The initial talent-search questionnaire of the students, their talent-search SAT scores, and an eight-page questionnaire assessing their achievements in high school are the three main sources for the results of this study. The initial talent-search questionnaire was designed to assess the characteristics of the students at the time of talent search participation. The purpose of the follow-up questionnaire was to determine this group's achievement in high school, particularly in mathematics and science.

PROCEDURES

The subjects were mailed the eight-page follow-up questionnaire with an offer of monetary compensation ($5.00, or in some cases $6.00) as an incentive to return the questionnaire. The questionnaires were mailed to students at a time when they should have completed high school without educational acceleration or deceleration since talent search participation. Usually, the questionnaire was completed by the students while they were freshmen in college. Because the students were sampled from the three talent searches held in 1972, 1973, and 1974, and they could have participated in the talent searches as either seventh or eighth graders, the follow-up questionnaires were sent out in four different waves: in December 1976 (N = 214, Cohn, Note 1),[1] 1977 (594), 1978 (881), and 1979 (499). After 6 weeks, students who still had not completed the questionnaire were sent a reminder letter, including an additional questionnaire. Six weeks later a postal card reminder was sent. Finally, subjects were telephoned and urged to provide the questionnaire responses orally.

The response rates for each wave of the follow-up were 94 (Cohn, Note 1), 90, 93, and 90 percent, respectively, of the total sample. If unlocatable persons are omitted, the response rates become 98 (Cohn, Note 1), 94, 96, and 93 percent, respectively. Across all waves, the overall response rate exceeded 91 percent of the total sample of 2,188 students. There were 1,996 students in the analyses, 62 percent of whom were males. Because the response rate was so high, it seems likely that the findings accurately reflect the development in high school of the sex difference in mathematical reasoning ability among the talent search participants.

Nonrespondents: The nonresponse rate for males and females was approximately 9 percent in both cases. Nonrespondent males were not significantly less able mathematically or verbally than the males who did return their questionnaires. Nonrespondent females were, however, significantly less able mathematically, but not verbally, than those who returned questionnaires. The effect size (Cohen, 1977) for the difference on SAT-M for the girls was considered small. Thus, this difference was not judged important. (The meaning of effect size will be discussed below.)

Data Analysis

Statistical analyses, performed with the SPSS program (Nie et al., 1975), were done separately for the first wave, second wave, and combined third and fourth waves of the follow-ups.[2] Analyses using talent-search SAT scores were also performed separately by grade at talent search.[3]

Because the numbers for all the tests were large, effect sizes (Cohen, 1977) were computed to test whether a significant difference was important. Effect sizes are computed differently, depending on which statistic is used. Thus, the values for the various effect sizes are not equivalent. Cohen (1977) arbitrarily classified effect sizes as being either small, medium, or large. In this study a medium or large effect size was considered important.

[1] It was Cohn and Stanley's (Cohn, Note 1) responsibility for conducting the first wave of the follow-ups, with 214 students who had either met a science criterion or had scored greater than 420 on SAT-M as an eighth grader. The data collection and analysis for the remaining three waves, with N = 1,974, was our responsibility.

[2] The first and second wave data could not be combined with the third and fourth waves, because the questions on the questionnaire were slightly different in some cases.

[3] This was done to reduce confounding; in the talent search, most eighth-grade participants received higher scores on the SAT than did the seventh-grade participants.

TABLE I

SAT Scores at the Time of the Talent Search and in High School of the Participants in Follow-up by Wave, versus High School Performance of a National Sample of College-Bound Seniors

Talent Search

	First Wave[a]		Second Wave		Third and Fourth Waves	
	M	*SD*	*M*	*SD*	*M*	*SD*
SAT-M						
Boys	567	91	549	74	526	76
Girls	505	58	510	58	498	61
t of mean difference	5.1 $p < .001$		6.7 $p < .001$		6.9 $p < .001$	
SAT-V[b]						
Boys	—		443	86	400	65
Girls	—		468	86	411	74
t of mean difference			−3.1 $p < .01$		*ns*	

High School

	First Wave[a]		Second Wave		Third and Fourth Waves		National Sample of College-Bound Seniors	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
SAT-M								
Boys	691	75	693	72	695	67	493	121
Girls	652	72	643	68	650	75	443	109
t of mean difference	3.5 $p < .001$		7.9 $p < .001$		10.6 $p < .001$			
SAT-V								
Boys	596	100	602	82	590	88	431	110
Girls	594	115	612	83	592	91	423	110
t of mean difference	*ns*		*ns*		*ns*			

Note. N = 1,996

[a] Taken from Cohn (Note 1).

[b] SAT-V was administered only in the 1973 Talent Search. Thus SAT-V scores were available for the 1973 Talent Search eighth graders, all in the second wave of the follow-ups, and for the 1973 Talent Search seventh graders, all in the third wave of the follow-ups.

RESULTS

Initial Sex Difference

Mean SAT scores of the follow-up group at the time of the talent search are shown in Table I. Because of the additional selection criteria, mean scores are much higher than the average from SMPY's six talent searches.

The groups' scores were also far superior to the means of a national sample of college-bound seniors (ATP, 1979). On SAT-M, boys in each wave scored significantly higher than the girls by at least 28 points, whereas girls scored higher on the SAT–V—significantly so for the second-wave.[4] The effect size for the sex difference on SAT-M in the talent search on the average was medium (i.e., .82, .59, and .41, respectively), while for the difference on SAT-V it was small (i.e., .29 and .16, respectively). Thus, the sex difference on SAT-M was considered important, but not the difference on SAT-V. Below we investigate the consequences in high school of this important difference on SAT-M in the seventh or eighth grade.

SAT Scores in High School

In the follow-up questionnaire the students were asked to report the SAT scores they received in high school. By the end of high school the boys' and girls' mean score on SAT-M had increased an average of 155 and 145 points, respectively, from the time of talent search participation (see Table I). Both boys and girls in the follow-up scored approximately 200 points better than their respective norm group of college-bound seniors (see the lower half of Table I). The sex difference on SAT-M in high school was significant beyond the $p < .001$ level. Effect sizes ranged from .52 to .71, all of which are in the medium range. Thus, the sex difference found in the seventh grade clearly persisted.

On SAT-V in the second wave of the follow-up males improved by 159 points and females by 144 points (see Table I). For the combined third and fourth waves males increased their scores by 190 points and females by 181 points. The mean scores on SAT-V were approximately 170 points above the mean for a national sample of college-bound seniors. It can be seen in Table I that the initial sex difference on SAT-V favoring girls at the time of talent search participation diminished in high school, and for the second wave was no longer statistically significant. On both SAT-M and SAT-V SMPY males improved more than SMPY females. The sex differences in the two-point growth curve on SAT-M and SAT-V were found to be statistically significant ($p < .05$ for SAT-M and $p < .01$ for SAT-V) by two separate multivariate analysis of variance procedures (Finn & Bock, 1981). Relative to Johns Hopkins undergraduates, whose means are 626V and 677M, SMPY students at the end of high school score somewhat higher on SAT-M but somewhat lower on SAT-V.

[4] The first wave of the follow-up consisted only of students who had been at least eighth graders in the talent search, the second wave consisted mainly of former eighth graders but of some former seventh graders in the talent searches, and the combined third and fourth waves consisted mainly of former seventh graders but also of some former eighth graders. The talent search mean score difference on SAT-M and SAT-V for the waves is probably accounted for by this difference in composition of the groups.

Mathematics Course Taking

SMPY males took significantly more semesters of high school mathematics than did SMPY females (i.e., 9.2 vs. 8.4). By a *t*-test this difference was statistically significant beyond the .001 level. The effect size, *d,* was approximately .33. Thus, the effect was considered to be small and therefore not important.

Self-reported achievement in mathematics courses taken is shown by sex in Table II. The mean grades were high, with girls receiving marginally better grades. Although the overall mean mathematics grades for boys and girls were not much different (i.e., on a scale where A = 4, B = 3, etc., boys had a mathematics grade point average of 3.5 and girls 3.6), it becomes apparent from Table II that in almost every comparison by course and sex girls receive slightly better grades. The difference resulted from more SMPY girls than boys reporting that they received A's in their coursework. A sign test was employed to test the statistical significance of this difference in grades by sex. It yielded chi-square = 20.5, $p < .001$, with a large effect size, $g = .41$. Thus, it was accepted that SMPY girls reported receiving somewhat better grades in their mathematics classes than their male counterparts reported.

The correlations between talent search SAT-M score and overall grade point average in high school mathematics courses, computed separately for follow-up wave and grade, ranged between .31 and .41 (medium effect size range) for the boys and .17 and .27 (small effect size range) for the girls. Verbal ability did not relate much to mathematics grades. Thus, for boys, mathematical ability exhibited in the seventh grade does relate importantly to grades received in their mathematics coursework.

SMPY males reported taking their mathematics in a significantly earlier school grade than SMPY females. The mean school grades when SMPY students took each of their mathematics courses are also shown in Table II. In almost every comparison by sex, SMPY males took the course in a slightly earlier grade. A sign test indicated that the difference was significant (chi-square = 22.1, $p < .001$). The effect size was large ($g = .42$).

It is apparent from Table II that differential course taking in mathematics by the SMPY males and females occurred at the upper levels. About the same percentage of girls and boys took each mathematics course up through trigonometry. But then approximately 10 percent more boys than girls took college algebra and analytic geometry. This difference in proportions was significant at the $p < .05$ level, but the computed effect size was small. With respect to calculus, approximately two-thirds of SMPY boys took at least one calculus course, compared to 40 percent of the girls (see Table II). This difference in proportions was significant ($p < .01$), with a medium effect size ($h = .53$). We conclude that the gender difference in taking calculus in high school was important.

TABLE II

Reported Mathematics Course Taking in High School by SMPY Students, Their Mean Course Grades, and Their Mean School Grades When Enrolled, Shown by Sex and Follow-up Wave for Those Courses Where at Least 5% Had Been Enrolled[a]

Follow-up Wave	First		Second		Third & Fourth	
	Male (133)	Female (69)	Male (310)	Female (221)	Male (785)	Female (478)
Algebra I						
Mean course grade	3.74	3.85	3.65	3.70	3.69	3.75
SD	0.51	0.36	0.60	0.58	0.55	0.51
Mean school grade	8.10	8.16	8.13	8.19	8.11	8.11
SD	0.62	0.44	0.58	0.50	0.51	0.46
Percentage enrolled	96	100	93	94	92	94
Algebra II						
Mean course grade	3.61	3.67	3.57	3.60	3.60	3.62
SD	0.65	0.59	0.68	0.66	0.63	0.61
Mean school grade	9.39	9.49	9.49	9.53	9.48	9.58
SD	0.08	0.68	0.82	0.83	0.84	0.91
Percentage enrolled	94	100	92	94	92	95
Plane Geometry						
Mean course grade	3.66	3.72	3.64	3.64	3.68	3.65
SD	0.65	0.51	0.62	0.57	0.59	0.59
Mean school grade	9.61	9.82	9.54	9.74	9.34	9.47
SD	0.71	0.62	0.67	0.58	0.65	0.63
Percentage enrolled	93	99	93	94	92	94
College Algebra						
Mean course grade	3.61	3.67	3.60	3.53	3.49	3.53
SD	0.73	0.57	0.63	0.74	0.73	0.68
Mean school grade	10.87	11.33	10.68	10.97	10.70	10.89
SD	0.99	0.57	0.90	0.86	0.90	0.79
Percentage enrolled	53	35	49	40	43	35

(*Continued on next page*)

TABLE II
(continued)

Follow-up Wave	First		Second		Third & Fourth	
	Male (133)	Female (69)	Male (310)	Female (221)	Male (785)	Female (478)
Trigonometry						
Mean course grade	3.51	3.56	3.55	3.60	3.54	3.58
SD	0.75	0.65	0.66	0.66	0.67	0.66
Mean school grade	10.52	11.19	10.60	10.96	10.50	10.79
SD	0.77	0.54	0.75	0.64	0.77	0.67
Percentage enrolled	87	86	81	80	83	80
Analytical Geometry						
Mean course grade	3.51	3.65	3.49	3.62	3.49	3.55
SD	0.75	0.53	0.72	0.62	0.74	0.68
Mean school grade	10.93	11.15	10.87	11.08	10.74	10.92
SD	0.71	0.56	0.75	0.56	0.97	0.63
Percentage enrolled	80	67	71	61	70	60
Calculus I						
Mean course grade	3.47	3.62	3.44	3.55	3.40	3.59
SD	0.85	0.56	0.78	0.69	0.77	0.61
Mean school grade	11.42	11.93	11.65	11.82	11.60	11.82
SD	0.77	0.26	0.61	0.42	0.63	0.45
Percentage enrolled	61	42	69	34	66	43
Calculus II						
Mean course grade	3.56	3.59	3.42	3.63	3.39	3.50
SD	0.71	0.57	0.76	0.60	0.81	0.73
Mean school grade	11.67	11.93	11.73	11.87	11.73	11.85
SD	0.71	0.27	0.57	0.34	0.60	0.47
Percentage enrolled	55	39	61	30	53	29

Probability & Statistics						
Mean course grade	3.62	3.90	3.59	3.87	3.69	3.84
SD	0.56	0.32	0.53	0.34	0.54	0.42
Mean school grade	11.28	11.80	11.47	11.30	11.17	11.55
SD	0.96	0.63	0.93	1.19	1.05	0.63
Percentage enrolled	22	15	18	10	17	12
Elementary Functions						
Mean course grade	3.50	3.50	4.00	3.50	3.25	3.51
SD	0.55	0.55	0	0.71	0.90	0.66
Mean school grade	10.83	11.33	11.33	11.33	10.92	11.11
SD	0.40	0.52	0.58	0.58	0.57	0.47
Percentage enrolled	5	9	1	1	6	8
Computer Science						
Mean course grade	3.64	3.75	3.65	3.88	3.66	3.68
SD	0.50	0.50	0.58	0.34	0.63	0.59
Mean school grade	11.00	11.00	11.32	11.33	11.21	11.53
SD	0.95	1.41	0.96	1.11	0.93	0.80
Percentage enrolled	9	6	19	7	23	17

[a] The differences between males and females in course grades and school grades were significant. Females received better grades ($\chi^2 = 20.5$, $p < .001$, $g = .41$), and males took their mathematics in an earlier grade ($\chi^2 = 22.1$, $p < .001$, $g = .42$).

Mathematics "enrichment" courses also were taken more frequently by boys than by girls. The difference in proportions was significant ($p < .05$), but with a small effect size.

Because SMPY girls took their mathematics later than SMPY boys, they had less time to take calculus and other advanced mathematics courses in high school. This difference accounts for some of the disparity in the number of mathematics courses taken. For example, the same proportion of girls and boys took mathematics up to and including 11th grade; that is, in the 11th grade, 83 percent of both boys and girls took mathematics. In the 12th grade, however, more SMPY boys than girls took mathematics (68% vs. 60%). This difference was statistically significant, but its effect size was not large enough to be considered even small ($h = .17$). Thus, it was not judged important.

It is of interest what the possible predictors of high school mathematics course taking are. We hypothesized that it was probably a combination of ability, family background variables, and attitudes. Our best measures of ability were the talent search SAT-M score and, if available, talent search SAT-V score. For family background variables our best available measures were parents' education, fathers' occupational status,[5] number of siblings, and sibling position. Finally, our best measures of attitudes were rated liking for mathematics in talent search, if available (otherwise in high school), rated importance of mathematics for future career (in seventh or eighth grade), and having rated mathematics as the favorite course in high school. The "dummy" variable sex was also added to the equation to see what weight it would have. Stepwise multiple regression analyses were performed separately by follow-up wave and grade. In the analyses, only relatively small amounts of variance in mathematics course taking (between 9 and 16%) could be accounted for by the predictor variables. The actual R^2 s, respectively by follow-up wave and grade, were .12 (N = 177), .16 (N = 329), .09 (N = 682), and .10 (N = 448). (Further data are available upon request from the authors.) The effect size values, f^2, ranged from .10 to .19. One was in the medium range, while the other three were in the small range.

Having rated mathematics as the favorite course in high school appeared to be the overall best predictor of mathematics course taking. This was followed by sex and mathematical ability in talent search. Least important among the set of variables appeared to be the family background characteristics. It can be argued that having rated mathematics as the favorite course in high school is not a predictor because this rating was done at the end of high school, not before. It becomes of interest, then, that liking for math in the seventh grade was not a stronger predictor of course taking than the ability measures.

---

[5] Occupational status was assessed by the National Opinion Research Center (NORC) transform of the Duncan Socioeconomic Index (Hodge, Siegel, & Rossi, 1964; Reiss, 1961).

Comparing the standard error of estimate with the standard deviation of the criterion variable revealed the numbers to be almost identical. On account of this and the effect sizes, it was concluded that the independent variables, which included sex and ability measures, could not accurately predict subsequent mathematics course taking in this population of able young students.

In a separate analysis it was found that talent search SAT-M and SAT-V scores, liking for mathematics, and sex could not discriminate between students who took mathematics in the 12th grade and the ones who did not (although the discriminant functions comprised of these variables were mostly significant). These same variables, however, could significantly and somewhat better discriminate between students who took calculus and those who did not. For the first set of analyses in which we tried to discriminate between students taking mathematics in the 12th grade and the ones not doing so, done separately by follow-up wave and grade, the first canonical correlation ranged in value from .11 to .32. Classification of students on whether or not they took mathematics in the 12th grade, based on the discriminant function, was correct approximately 59 percent of the time. The same analyses used to discriminate between students who took calculus and the ones who did not resulted in canonical correlations ranging from .30 to .47. Classification based on the discriminant function was correct approximately 65 percent of the time. Clearly, among a highly able group of students, other factors, such as availability of calculus in the high school, probably are also important in determining whether a student takes mathematics, especially calculus, in the 12th grade.

In a further analysis the difference between talent search SAT-M and SAT-V was computed. This variable was considered important because, independent of SAT-M level, it determines the hierarchy of preference for mathematics versus verbal tasks (Gordon, 1981). The difference was correlated with the number of semesters of mathematics taken. For girls a significant relationship was not found. For boys, one was found (r = .19, N = 270), with a small effect size. The larger V-M was, the (slightly) more mathematics the student tended to take.

AP-Level Mathematics Courses

Of the mathematics courses offered in high school, the most advanced and difficult are the ones that prepare students for taking the College Board's Advanced Placement Program (AP) examinations. Students can take either the AP Mathematics, Level AB, or the more advanced Level BC. A high grade on the Level AB examination usually yields credit for a one-semester college course in calculus, while two semesters of credit can usually be gained by means of a high grade on the BC examination. Grades reported on these examinations range from 1 to 5, where 3, 4, and 5 are considered high scores.

TABLE III

Reported Performance on the Advanced Placement Program (AP) Examinations and the College Board's High-School Level Achievement Tests in Mathematics by Follow-up Wave

AP Scores	First Wave (202)		Second Wave (531)		Third & Fourth Waves (1263)		AP Examination Distribution of Candidate Grades May 1980
	Male	Female	Male	Female	Male	Female	
Calculus AB							
M Score	4.1	3.0	3.8	3.0	3.7	3.7	3.0
SD	0.8	1.4	1.2	1.3	1.1	1.1	1.2
N	11	4	33	15	98	43	20,096
Calculus BC							
M Score	3.6	2.3	3.7	4.0	3.8	3.6	3.2
SD	1.2	0.6	1.2	0.9	1.2	1.1	1.3
N	26	3	51	6	132	26	7,783
Achievement Test Scores							National Sample of 1978 College-Bound H.S. Students (ATP, 1979)
Math Level I							
M Score	692	664	698	656	695	644	541
SD	81	99	74	70	65	76	99
N	34	19	60	58	149	100	146,426
Math Level II							
M Score	742	676	751	724	748	705	665
SD	67	93	60	57	59	71	95
N	46	7	91	29	281	99	32,743

Table III shows that, although many more boys than girls took these examinations, no significant differences in grades earned were seen.

Approximately 12 percent of the males and 8 percent of the females took the AP Calculus AB examination. Although the difference in proportions was significant at the $p < .01$ level, the effect size did not even reach the criterion for being considered small. The more difficult BC examination was taken by 17 percent of the males and 5 percent of the females, a difference significant at the $p < .01$ level. The effect size was small ($h = .40$), even though the ratio of boys to girls taking the test is more than 3-to-1. We believe that these ratios are not negligible. Moreover, the harder the mathematics criterion becomes, the greater the ratio of boys to girls.

Mathematics Achievement Tests

Significantly more SMPY males than females reported that they took the College Board's mathematics achievement tests (see Table III). The difference in proportions between males and females taking the Math Level I achievement test, which is less advanced than Math Level 2, was not significant (i.e., 20% of the boys vs. 23% of the girls). The difference on Math Level 2 was, however, significant ($p < .01$). Approximately 34 percent of the boys took this test, while only 18 percent of the girls did. The effect size was in the small range ($h = .37$).

SMPY males did score higher than the females on the mathematics achievement tests (Table III). On the Math Level 1 boys scored on the average 695, while girls scored around 650. The difference was significant beyond the $p < .01$ level, except for the first wave of the follow-up (see Table III). The associated effect sizes equaled .58 for the second wave and .73 for the combined third and fourth waves, which are considered to be in the medium range.

On the more difficult Math Level 2 the boys scored on the average 748 and the girls approximately 708, significantly different beyond the $p < .05$ level. The effect sizes ranged from .47 to .86, which is in the medium to large range.

It was hypothesized that perhaps the sex difference in mathematical reasoning ability, detected as early as the 7th grade, may be related to why SMPY females received lower scores than SMPY males on the mathematics achievement tests. An analysis of covariance was, thus, computed on Math Level 1 and 2 achievement test scores by follow-up wave and grade. The effect of sex was tested after controlling for talent search SAT-M. Because there was no significant sex difference on Math Level 1 for the first wave of the follow-up, the first wave was excluded from this analysis. When mathematical reasoning ability was controlled for, the effect of sex was reduced (except in one of the seven analyses). The ANCOVA F's ranged in value from .3 to 15.9, while the ANOVA F's ranged from 3.5 to 19.2. Yet sex still remained a significant effect for four of the seven analyses performed. Because the effect of sex was reduced, a relationship may exist between the sex difference on the Math Level 1 and 2 achievement tests and the less well-developed mathematical reasoning ability of SMPY girls compared to SMPY boys. But other factors also seem to be operating that make girls perform less well on the mathematics achievement test. Obviously, it could hardly be performance in class, because girls reported receiving better grades for their course work. Furthermore, it could not be that boys had taken more mathematics. The big gender difference in mathematics course taking occurs in calculus, which is not covered on the achievement tests and was taken by the students only after these tests were completed. More research is needed to discover possible causes for this discrepancy. Our present work can tell us only which factors are not likely to be involved.

Stepwise multiple regression analyses performed separately by follow-up wave and grade were also employed to see if talent search SAT scores and sex can predict the number of science and/or mathematics achievement or AP examinations taken in high school. There was a significant sex difference

($p < .001$) in the total number taken (1.45 for boys vs. 0.81 for girls) and the effect sizes were in the small to medium range (.68, .47, and .43 for the three waves).

Between 16 and 23 percent of the variance in the taking of these examinations could be accounted for by the talent search SAT scores and sex. The associated effect sizes ranged in value from .19 to .30, which are considered to be medium.

By itself talent search SAT-M, the overall best predictor, could account for between 15 and 21 percent of the variance. Sex was a significant predictor in some but not all equations, accounting for .6 to 3.6 percent additional variance. Therefore, it was accepted that talent search SAT scores could predict fairly accurately the number of science and/or mathematics achievement or AP examinations taken, and the contribution of sex was negligible.

The difference between talent search SAT-M and SAT-V scores (i.e., SAT-M minus SAT-V) was computed and then correlated with the number of science and/or mathematics achievement or AP examinations taken in high school. The negative correlations, computed separately by follow-up wave, for the girls were not significant, but the positive *r*'s for the boys were. The *r*'s for the boys were .10 (N = 270, small effect size) and .35 (N = 35, medium effect size).

College Major

Students were asked to report their intended college majors. The percentage of males reporting that they intended to major in the mathematical sciences was 15 percent, while for the females this was 17 percent. The difference favoring the girls was not significant, however. Because some of these intended majors will probably change during the college experience, this may not be a reliable predictor. It does, however, indicate that initially in college the girls were at least as interested in mathematics as the boys and perhaps find it as important for their careers. Thus, this variable may be an indicator of attitudes. Benbow and Stanley (Note 2) investigated the sex difference in college majors in this group and possible causes. No strong explanations for the sex difference were found.

College Mathematics

The students also reported whether they took mathematics during their first semester of college. Of the students in college, 81 percent of the males and 68 percent of the females reported that they took at least one mathematics course during their first semester. This difference was significant beyond the $p < .01$ level, with a small effect size ($h = .30$). Because the difference was not considered important, perhaps it is not worthwhile to ponder on reasons for the discrepancy that, even though slightly more girls were planning to major in the mathematical sciences, the boys took more mathematics during the first semester of college; however, it is of interest.

Mathematics Contests

Several students in the follow-up reported that they had participated in mathematics contests in high school. Approximately 23 percent of the boys and 12 percent of the girls had participated in at least one, significantly different beyond the $p < .01$ level. The effect size was small ($h = .29$), even though the ratio was almost two boys to every one girl.

Attitudes Toward Mathematics

Students in the follow-up were asked to rate their liking for mathematics on a 5-point scale, ranging from strong disliking (1) to strong liking (5). As a group, both girls and boys expressed moderate liking. The males' means in the three waves ranged from 4.28 to 4.44. The females' means ranged from 4.12 to 4.36. Clearly, the difference between the sexes was extremely small. For the combined third and fourth waves of the follow-up, however, the sex difference was significant ($p < .05$) because of the large N. Yet the effect size equaled only .13, which is not even considered small by Cohen (1977). Thus, this statistically significant difference was ignored.

Students also were asked to rank their preference for mathematics relative to biology, chemistry, and physics. Mathematics was most highly ranked by both SMPY males and females. No significant difference occurred.

The favorite courses of SMPY students in high school were mathematics and science. Their favorite courses were grouped into five categories, by sex and total (see Table IV). The favorite subject for both the boys and girls was mathematics. The difference between the SMPY males and females was significant ($p < .01$) because of the large N. When an effect size was calculated, however, it did not reach the criterion for even being considered small ($h = .11$). Thus, this significant difference was ignored. SMPY boys and girls prefer mathematics the most and to an essentially equal degree in high school in comparison with their other high school subjects.

Hence, it cannot be that these girls like mathematics less than boys, causing them to participate at a lower level in mathematics than boys. Perhaps instead girls prefer verbal areas more than boys, which would then tend to lower their participation in mathematics.[6] A chi-square was computed between rated favorite course being verbal or quantitative and sex. The "other" category in Table IV was omitted. The chi-square equaled 26.8 and was significant. The effect size was small ($w = .14$), indicating a weak association.

This possibility was investigated in another way and supported. Students were classified into two groups on the basis of whether they had indicated a

[6] Robert A. Gordon suggested this idea to us.

TABLE IV

The Reported Favorite Courses of SMPY Students in High School by Sex and for 17-Year-Olds

Favorite Course	Percent		
	Males (1228)	Females (768)	17-Year-Olds (NAEP, 1979)
Mathematics	36	31	18
Science	34	25	12
Social Studies	11	9	13
English	7	17	16
Other	13	17	41

verbal subject in high school as their favorite course or whether a quantitative one (i.e., science or mathematics) had been reported. Students not fitting into either category were excluded from the analysis. A discriminant analysis was performed, separately by follow-up wave, using these variables as indicators of participation in mathematics: number of semesters of mathematics taken in high school, whether a mathematics course was taken during the first semester of college, the quantitative index of the student's intended college major on a scale of 1 to 5 (see Benbow & Stanley, Note 2), number of science or mathematics achievement or AP examinations taken, and number of mathematics contests participated in. For the first wave of the follow-up, two variables, mathematics course-taking in college and the quantitative index of intended college major, were not available. Thus, the analysis for the first wave is slightly different from the other two. All analyses presented here were performed on a combined group of boys and girls, because no differences were apparent between the groups that would warrant separate analyses. The results of the discriminant analysis are shown in Table V.

Before the first function was removed the Wilks' Lambda values for the three waves equaled .76, .88, and .80, respectively. This indicated that there existed some discriminating power between the groups in the variables being used (see Table V). The one and only discriminant function for each analysis was, therefore, significant. Its associated canonical correlation values equaled .49, .35, and .45 (see Table V). The contributions of the discriminating variables to the function can be seen in the lower half of Table V. No one variable appeared to be consistently best in discriminating between the groups.

Classification based on the discriminating function was correct for approximately 73 percent of the students, compared to 50 percent expected by chance (see Table V). Thus, there is considerable overlap between the groups. Yet the function does aid in classifying students. Therefore, it was accepted that measures of participation in mathematics can discriminate between students preferring verbal areas in high school and those preferring

TABLE V

Discriminant Analyses Between SMPY Students Preferring Verbal Areas in High School and Students Preferring Quantitative Areas on Several Measures of Mathematics Participation

Follow-up Wave (N)	Eigen value	Canonical Correlation	Wilks' Lambda	Chi-Square	df	Sig.
1 (100)[a]	.32	.49	.76	26.7	3	.001
2 (433)	.14	.35	.88	56.5	5	.001
3 and 4 (1048)	.26	.45	.80	237.7	5	.001

Standardized Discriminant Function Coefficients:	First Wave	Second Wave	Third and Fourth Waves
College Math[b]		.67	.15
Semesters of Math in High School	.69	.43	.20
Choice[c]		.12	.87
Number of Science and Math Tests	.45	.31	.16
Number of Math Contests	.32	.05	.09
Percent Classified Correctly	75	67	75

[a] The analysis for the first wave of the follow-up was performed differently, using only semesters of mathematics in high school, number of science and mathematics tests taken, and number of mathematics contests.

[b] College Math is a measure of whether or not a student took mathematics in his or her first semester of college.

[c] Choice is an indicator of quantitative orientation of a student's college major on a scale of 1 to 5 (see Benbow & Stanley, Note 2).

the quantitative areas. Because slightly more girls than boys prefer the verbal areas, although no differences were seen in the quantitative areas, these findings support the hypothesis that the lower female participation in mathematics in high school can be accounted for to some extent by the greater female than male preference for the verbal areas.

DISCUSSION

In all six SMPY talent searches a large sex difference favoring males was found on the SAT-M, a test of mathematical reasoning ability (Benbow & Stanley, 1980b, 1981). This study followed up at high school graduation the students studied by Benbow and Stanley (1980b) who, as 7th or 8th graders, had participated in any of the first three talent searches and had scored as well as a national sample of 11th and 12th grade females do on the SAT. The objective was to investigate the development and consequences of this initial sex difference. Many consequences were found.

This study and the Benbow and Stanley (1980b, 1981) paper demonstrate that (1) sex differences in mathematical reasoning ability are found at an early age among mathematically talented students; (2) they persist over several years and are related to subsequent differences in mathematics achievement; and (3) the differences in mathematical reasoning ability and achievement do not reflect differential mathematics course taking, at least not among mathematically talented adolescents.

One of the more interesting findings was that the mathematical abilities of SMPY males appeared to develop significantly more rapidly or to improve more during high school than the abilities of SMPY females. Males improved significantly more than females on both SAT-M and SAT-V by about 10 points. This finding partly contradicts a previous study, which showed that the sex with the initial superior ability improved more in that ability during high school (Shaycoft, 1967).

Fennema and Sherman (1977) postulated that sex differences in mathematical ability commence in high school because boys take more semesters of mathematics than girls (i.e., differential course taking). Because SMPY boys did take marginally more mathematics in high school than SMPY girls, this might appear to be the reason why the boys improved more on the SAT-M in high school. But this hypothesis cannot explain why the boys improved more on the SAT-V also. Furthermore, the differential course-taking hypothesis cannot explain why there is a sex difference on the SAT-M earlier than senior high school and, for the following reasons which were derived from this and two other studies, cannot even account for the increase in the sex difference in high school: (1) The initial sex difference on SAT-M was found in the 7th or 8th grade, before differential course taking took effect (Benbow & Stanley, 1980b); (2) equal percentages of girls and boys took mathematics in high school up to the 12th grade, when the SATs are normally taken; (3) SMPY boys took only about one semester more of mathematics than SMPY girls, which was mostly accounted for by the larger number of SMPY boys than girls taking calculus (calculus items do not appear on the SAT); and (4) in a separate study it was found that the best predictor of high school SAT-M score was talent search SAT-M, not the number of semesters of mathematics taken in high school, which accounted for little additional variance in high school SAT-M (Benbow, Note 3).

Clearly, the differential course-taking hypothesis does not explain the ability differences found in this population. More research is needed to discover possible reasons why SMPY males improved more than SMPY females in both mathematical and verbal reasoning ability during high school. Perhaps SMPY males are proceeding at a faster developmental rate than SMPY females.

Can ability differences explain differential course taking? When students taking calculus and students not taking calculus in high school were compared, three variables could discriminate somewhat between them: talent search SAT-M and SAT-V, liking for mathematics in talent search, and sex. The most typical student to take calculus in high school was a male who was able and liked mathematics. The significance of the sex difference will be discussed later. It is also of interest to note that the 10-point increase in the sex difference on SAT-M during high school made the mean sex difference for the SMPY group equal the mean difference found for college-bound seniors in high school (ATP, 1979).[7]

Sex differences in mathematical achievement were widely noted. Some of these differences appeared to be related to the sex difference on SAT-M in the talent search. When ability on the SAT-M at talent search was controlled, the significant sex differences in performance on the mathematics achievement tests were either no longer significant or reduced. With regard to taking mathematics and/or science achievement or AP examinations, talent search SAT-M score was the best predictor. Sex accounted for little additional variance. It appears then that the sex difference in mathematical reasoning ability found as early as the seventh or eighth grade may later contribute to sex differences in mathematics achievement.

Slightly (but not significantly) more females than males were planning to major specifically in the mathematical sciences in college, and SMPY females received better grades in their high school mathematics courses than did SMPY males. The mathematics course grade differences can probably be explained by the sex differences favoring girls that have been found in conduct and demeanor in school (Baker, 1981; see Entwisle & Hayduk, 1981, in press). Girls have better conduct and demeanor. This possibility is consistent with the stronger relationship between mathematical reasoning ability and mathematics course grades for boys than girls. Unfortunately, we could not control for conduct or demeanor.

With regard to attitudes toward mathematics, few of the expected sex differences were found. SMPY boys and girls reported that they liked mathematics equally. Furthermore, reported attitudes toward mathematics had little relationship with achievement in mathematics. Attitudes toward mathematics at time of the talent search and in high school could not predict the number of semesters of mathematics taken. In a separate study, attitudes toward mathematics could not predict high school SAT-M score and score on the College Board's Math Level 1 achievement test (Benbow, Note 3). Expressed liking for mathematics, however, could somewhat discriminate between students taking calculus in high school and the ones not doing so. Furthermore, there were some indications that girls participate in mathematics less than boys not because they like it less, but partly because they like verbal areas, especially English, more than boys do.

Therefore, there does not appear to be much relationship between attitudes toward mathematics and achievement in mathematics in a high-aptitude group, unless the variables measured in this study were inadequate indicators of attitudes toward mathematics (i.e., mathematics liking, importance of mathematics for future job, and having rated mathematics a favorite course

[7] ATP (1981) shows the mean score of college-bound 12th-grade males on SAT-M to be 492, versus the analogous female mean of 443. That 49-point difference occurs among much more heterogeneous groups than the SMPY talent search participants constituted.

in high school). For example, Fennema and Sherman (1976) demonstrated that attitude toward mathematics involves several distinct components. Although our measures of attitudes might have been too global, it is hard to imagine that they failed to capture some information that was useful overall.

We tentatively conclude that mathematical reasoning ability by itself seems to be a better predictor of achievement in mathematics, even in this intellectually rather homogeneous group, than attitudes toward mathematics.

Yet mathematical reasoning ability cannot be the sole reason why girls perform less well than boys in high school, because it could not totally account for the sex difference in the mathematics achievement tests. Other factors must be operating. Performance in the class and having taken more courses in the subject matter were ruled out because: (1) SMPY females received better grades than the SMPY males in their mathematics course work; (2) the same proportion of females and males took mathematics up through the 11th grade, which is right before or when these achievement tests are taken; and (3) most of the sex difference in the taking of mathematics can be accounted for by the higher percentage of males taking calculus (calculus items do not appear on these achievement tests). What factors are involved in the performance differential between the sexes needs to be investigated. This study suggests strongly what these factors are not likely to be.

Along this line, Fox, Brody, and Tobin (1982) investigated the family background of the SMPY talent participants. They found few differences. Especially, no indications of differential training or encouragement of boys and girls were discovered.

A study of sex differences in science achievement among this population has also been conducted (Benbow & Stanley, Note 2). The sex difference in mathematical reasoning ability also related to the sex difference in science achievement, but could not account for the sex difference in intended college majors of the group.

Large sex differences were found in taking higher level mathematics. Among the SMPY group, almost twice as many boys as girls took calculus in high school. In college fewer females than males took mathematics during their first semester. Actually, the time when less SMPY females than males were taking mathematics seemed to begin in the 12th grade. Thus, if one wants to increase the participation of able women in mathematics, earlier than the 12th grade would seem to be the time to use some intervention strategies. As Sells (1980) pointed out, women are closed out of certain career options because they do not take enough mathematics in high school.

In 1973 such a program was implemented at SMPY (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 the girls to study a number of years of mathematics. The girls who successfully completed the program (i.e., those who were placed in Algebra II that fall) did take more advanced mathematics in high school and college (Fox, Benbow, & Perkins, Note 4). That was, however, the only major difference between this group of girls and an equally able group of girls not invited to attend the program. No effects were found for the girls who attended the accelerated algebra program but were not successful. Clearly, an early intervention strategy can improve the participation of girls in higher level mathematics, but the girls have to be successful in such a program. Unfortunately, many of the girls in this study may not have been able enough to benefit sufficiently from such intensive training.

A potential problem with our study is that it is based on self-reported data. Moreover, there is the possibility of differential accuracy (sex-related) in recall. Hamilton (1981) found, however, that females tend to exaggerate (positively) more than males. If this is the case, then our results might be underestimating the magnitude of the sex differences favoring males.

A further limitation of the study is that the students involved are mathematically able and highly motivated. Therefore, we urge caution when generalizing to the total population.[8] If one is interested in the question of why women do not pursue careers in mathematics and science as frequently as men do, however, the students whom we studied are an appropriate nonprobability sample of the population. Because all of them are intellectually (especially mathematically) talented, they are the most likely to enter the sciences at those academically difficult colleges where most top-level scientists received their undergraduate education (Davis, 1965; Werts, 1967).

Another problem was our use of Cohen's (1977) effect sizes to evaluate the importance of a difference found. Values of effect sizes have been classified as being either small, medium, or large. This classification is arbitrary. We choose to accept as important a difference classified as being either medium or large. As Cohen warns, however, in some areas even small differences can be important. Thus, we may have been too rigid in adhering to this standard. Many differences were found that we felt were important despite the small effect size. Readers should evaluate whether a small effect size might be important in some cases.

We conclude that sex differences in mathematical reasoning ability and achievement are widely noted in this highly able group of students, they persist over several years, and they are better accounted for by the sex difference in mathematical reasoning ability than by sex differences in expressed attitudes toward mathematics and mathematics course taking in junior and senior high school. The slightly greater female than male prefer-

[8] The influence of the researchers' expectations on these students' responses was probably minimal, because there was little personal contact and the SMPY study was not conducted primarily to find sex differences in this population.

ence for the verbal areas in high school might also explain some of the sex difference in participation in mathematics. The reason for this difference in preference for verbal areas is not clear and needs to be investigated. Moreover, why boys tend to reason better than girls from at least as early as second grade (Dougherty et al., Note 5) onward is also, of course, not clear. What interactions of factors such as environment, female versus male hormones during prenatal development, physiologically induced differences in activity levels, and different brain-hemisphere lateralization (Goy & McEwen, 1980; Harris, 1979; Levy, 1981; Wittig & Petersen, 1979) might be responsible cannot be ascertained yet. It "seems likely that putting one's faith in boy-versus-girl socialization processes as the only permissible explanation of the sex difference in mathematics is premature" (Benbow & Stanley, 1980b, p. 1264).

REFERENCE NOTES

1. COHN, S. J. *Two components of the Study of Mathematically Precocious Youth's intervention studies of educational facilitation and longitudinal follow-up.* Unpublished doctoral dissertation, Johns Hopkins University, 1980.
2. BENBOW, C. P., & STANLEY, J. C. Gender and the science major. To appear in M. W. Steinkamp & M. L. Maehr (Eds.), *Women in science.* Greenwich, Conn.: JAI Press.
3. BENBOW, C. P. *Development of superior mathematical ability during adolescence.* Unpublished doctoral dissertation, Johns Hopkins University, 1981.
4. FOX, L. H., BENBOW, C. P., & PERKINS, S. *An accelerated mathematics program for girls: A longitudinal evaluation.* Unpublished manuscript, 1982.
5. DOUGHERTY, K., HERBERT, M., EDENHART-PEPE, M., & SMALL, A. *Sex-related differences in mathematics, grades 2-5.* Unpublished manuscript, 1980.

REFERENCES

Admissions Testing Program. *National report, college bound seniors, 1979.* Princeton, N.J.: Educational Testing Service, 1979.

Admissions Testing Program. *National report on college-bound seniors, 1981.* Princeton, N.J.: Educational Testing Service, 1981.

ANGOFF, W. (ED.). *The College Board admissions testing program.* Princeton, N.J.: College Entrance Examination Board, 1971.

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

BAKER, D. P. Personal communication, September 1981.

BENBOW, C. P., & STANLEY, J. C. Intellectually talented students: Family profiles. *Gifted Child Quarterly,* 1980, *24,* 119–122. (a)

BENBOW, C. P., & STANLEY, J. C. Sex differences in mathematical ability: Fact or artifact? *Science,* 1980, *210,* 1262–1264. (b)

BENBOW, C. P., & STANLEY, J. C. Mathematical ability: Is sex a factor? *Science,* 1981, *212,* pp. 118, 121.

BENBOW, C. P., & STANLEY, J. C. Intellectually talented boys and girls: Educational profiles. *Gifted Child Quarterly,* 1982, *26,* 82–88.

BIERI, J., BRADBURN, W., & GALINSKY, M. Sex differences in perceptual behavior. *Journal of Personality,* 1958, *26,* 1–12.

COHEN, J. *Statistical power analysis for the behavioral sciences.* New York: Academic Press, 1977.

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

ENTWISLE, D. R., & HAYDUK, L. A. Academic expectations and the school attainment of young children. *Sociology of Education,* 1981, *54,* 34–50.

ENTWISLE, D. R., & HAYDUK, L. A. *Schooling of young children.* Baltimore, Md.: Johns Hopkins University Press, in press.

ERNEST, J. Mathematics and sex. *American Mathematical Monthly,* 1976, *83*(8), 595–612.

FENNEMA, E. Mathematics learning and the sexes: A review. *Journal for Research in Mathematics Education,* 1974, *5,* 126–139.

FENNEMA, E., & SHERMAN, J. Fennema-Sherman mathematics attitude scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. JSAS *Catalog of Selected Documents in Psychology,* 1976, *6,* 31–32. (Ms. No. 1225)

FENNEMA, E., & SHERMAN, J. Sex-related differences in mathematics achievement, spatial visualization, and sociocultural factors. *American Educational Research Journal,* 1977, *14,* 51–71.

FINN, J. D., & BOCK, R. D. *Multistat/multivariance manual.* Chicago, Ill.: International Educational Services, 1981.

FOX, L. H. Sex differences in mathematical precocity: Bridging the gap. In D. P. Keating (Ed.), *Intellectual talent: Research and development.* Baltimore, Md.: Johns Hopkins University Press, 1976.

FOX, L. H., BRODY, L., & TOBIN, D. (EDS.). *Women and the mathematical mystique.* Baltimore, Md.: Johns Hopkins University Press, 1980.

FOX, L. H., BRODY, L., & TOBIN, D. *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 (NIE-G-79-0113), January 1982.

GARAI, J. E., & SCHEINFELD, A. Sex differences in mental and behavioral traits. *Genetic Psychology Monographs,* 1968, *77,* 169–229.

GLENNON, V. J., & CALLAHAN, L. G. *A guide to current research: Elementary school mathematics.* Washington, D.C.: Association for Supervision and Curriculum Development, 1968.

GORDON, R. A. Personal communication, September 1981.

GOY, R. W., & MCEWEN, B. S. *Sexual differentiation of the brain.* Cambridge, Mass.: Massachusetts Institute of Technology Press, 1980.

HAMILTON, L. C. Sex differences in self-report errors: A note of caution. *Journal of Educational Measurement,* 1981, *18*(4), 221–228.

HARRIS, L. J. Sex-related differences in cognitive functioning. *Science,* 1979, *206,* 50–52.

HODGE, R. W., SIEGEL, P. M., & ROSSI, P. H. Occupational prestige in the United States, 1925–1963. *American Journal of Sociology,* 1964, *70,* 286–302.

KEATING, D. P. The Study of Mathematically Precocious Youth. In J. C. Stanley, D. P. Keating, & L. H. Fox (Eds.), *Mathematical talent: Discovery, description, and development.* Baltimore, Md.: Johns Hopkins University Press, 1974.

LEVY, J. Sex and the brain. *The Sciences,* 1981, *21*(3), 20–28.

MACCOBY, E. E., & JACKLIN, C. N. *The psychology of sex differences.* Stanford, Calif.: Stanford University Press, 1974.

MESSICK, S., & JUNGEBLUT, A. Time and method in coaching for the SAT. *Psychological Bulletin,* 1981, *89*(2), 191–216.

National Assessment of Educational Progress. Males dominate in educational success. *National Assessment of Educational Progress Newsletter,* 1975, *8*(5), insert.

National Assessment of Educational Progress. *Attitudes toward science.* Report to the National Institute of Education (#08-5-02), 1979.

NIE, N. H., HULL, C. H., JENKINS, J. G., STEINBRENNER, K., & BENT, D. H. *Statistical package for the social sciences, 2nd edition,* New York: McGraw-Hill, 1975.

REISS, A. J. *Occupations and social status.* New York: Free Press of Glencoe, 1961.

SELLS, L. W. The mathematics filter and the education of women and minorities. In L. H. Fox, L. Brody, & D. Tobin (Eds.), *Women and the mathematical mystique.* Baltimore, Md.: Johns Hopkins University Press, 1980.

SHAYCOFT, M. F. *The high school years: Growth in cognitive skills.* Palo Alto, Calif.: American Institutes for Research, 1967.

SUYDAM, M. N., & WEAVER, J. F. *Individualizing instruction.* Washington, D.C.: U.S. Office of Education, 1970.

VERY, P. S. Differential factor structures in mathematical ability. *Genetic Psychology Monographs,* 1967, *75*(2), 169–207.

WERTS, C. E. Career changes in college. *Sociology of Education,* 1967, *40*(1), 90–95.

WILSON, J. W. *Patterns of mathematics achievement in grade 11: Z population.* Stanford, Calif.: School Mathematics Study Group, 1972.

WITTIG, M. A., & PETERSON, A. (EDS.). *Sex-related differences in cognitive functioning: Developmental issues.* New York: Academic Press, 1979.

AUTHORS

CAMILLA PERSSON BENBOW, Associate Research Scientist in Psychology, Johns Hopkins University, Baltimore, MD 21218. *Specializations:* Individual differences; psychometrics; intellectual precocity; sex differences.

JULIAN C. STANLEY, Professor, Dept. of Psychology, Johns Hopkins University, Baltimore, MD 21218. *Specializations:* Individual differences; psychometrics; intellectual precocity; sex differences.