

Mathematically Talented Males and Females and Achievement in the High School Sciences

CAMILLA PERSSON BENBOW

Iowa State University

and

LOLA L. MINOR

Johns Hopkins University

Mathematically talented youth, whether male or female, tend to have favorable attitudes toward science and to participate in the sciences at a level much higher than average. There were no overall sex differences in course-taking or course-grades in the sciences. Indications of sex differences favoring males, however, were found in participation in high school physics, the taking of and performance on high school and college level science achievement tests, and intention to major in the more quantitatively oriented fields of physics and engineering. No substantial sex differences in attitudes toward the sciences, except possibly physics, were detected. Overall attitude toward science did relate somewhat to participation in science. Moreover, sex differences in mathematical reasoning ability may explain some of the sex difference in science participation and achievement. These results may bear on why women are underrepresented in the sciences.

From 1972 through 1979 the Study of Mathematically Precocious Youth (SMPY) conducted six talent searches for mathematically apt students.

We thank Linda E. Brody and Julian C. Stanley for comments and suggestions and Lois S. Sandhofer for preparation of the manuscript. Financial support was provided by the Spencer and Donner Foundations, and the U.S. Department of Education. This material is based upon work supported by the National Science Foundation Grant MDR-8470387. Part of this work was completed while the first author was affiliated with Johns Hopkins University. Some preliminary results of this work were presented in Benbow and Stanley (1984).

The data from these searches revealed large sex differences in mathematical reasoning ability by grade seven (Benbow & Stanley, 1980, 1983). Since SMPY is a longitudinal study of its participants, the consequences of this sex difference can be assessed at different junctures. Utilizing results from SMPY's first follow-up of its students (i.e., after high school), a relationship between sex differences in achievement in mathematics in high school and the earlier sex difference in mathematical reasoning ability was detected (Benbow & Stanley, 1982).

The possible effects of sex differences in mathematical reasoning ability on science achievement have not, however, been investigated much by SMPY (but see Benbow & Stanley, 1984). Sex differences are pervasive in the pursuit of careers in science and in the numbers majoring in science in college (Benbow & Stanley, 1984; NSF, 1984). A better understanding of why this occurs is critical.

Social and environmental factors may be important determiners of this sex difference. Benbow and Stanley (1984) surveyed some of the literature and concluded that "the typical personality traits associated with scientists are ones more frequently held by males than females. Females and males in the same field, however, tend to be quite similar. There may be some differences in the way parents treat males and females, and parental evaluation may be more important to females than to males" (p. 169). Thus, socialization is implicated in the disproportionate representation of men and women in the sciences. Sex differences in mathematical reasoning ability may also play a role, as it did for mathematics achievement (Benbow & Stanley, 1982). The purpose of this present investigation was to determine if possible sex differences in science achievement are related, at least in part, to sex differences in mathematical reasoning ability.

METHODS

Subjects

Students were drawn from SMPY's first three talent searches (i.e., 1972, 1973, and 1974). In those, seventh- and eighth-graders in Maryland were eligible to participate if they had scored in the upper 5% (1972) or upper 2% (1973, 1974) nationally on any mathematics achievement subtest. Qualified students took the College Board's Scholastic Aptitude Test-Mathematics (SAT-M) and, in 1973, the Scholastic Aptitude Test-Verbal (SAT-V). These tests measure mathematical and verbal reasoning ability, respectively, especially among these young students (Benbow & Stanley, 1983). Additionally, talent search scores of at least 390 on SAT-M or 370 on SAT-V were required for inclusion in this follow-up study. These SAT criteria selected students who as seventh or eighth graders scored as well as the average high school female. Of the total sample ($N = 2,188$) meeting

these criteria, approximately 61% were male. When contacted, over 90% were freshmen in college.

Procedure

The main sources of data for this study were the seventh- or eighth-grade SAT scores and an after-high-school follow-up questionnaire. The follow-up questionnaire was mailed to each subject at the time of expected high school completion (Benbow & Stanley, 1982). Overall, 91% of the total sample returned the questionnaire for a final number of 1,996 students, of whom 62% were male. Excluding unlocatable persons, the response rate was 93%. There was no appreciable difference in response rates by sex or in SAT scores between respondents and non-respondents (Benbow & Stanley, 1982).

Data Analysis

Statistical analyses were performed with the SPSSX computer program. Since the seventh graders tended to score lower on the SAT than the eighth graders, their SAT scores were equated to avoid confounding because of age. Equating was performed via the method described by Angoff (1971) and separately by sex. Seventh grade scores were converted to be made comparable with the eighth graders.

RESULTS AND DISCUSSION

SAT Scores

A comparison of SAT scores by sex in the talent search is shown in Table I. On the SAT-M boys scored higher than girls ($p < .001$), whereas girls scored higher on the SAT-V ($p < .001$). The effect size values were .53 and .30, respectively. The sex difference in mathematical reasoning ability persisted into high school (Benbow & Stanley, 1982) and was still significant ($p < .001$) with an effect size of .61. This was not the case for the SAT-V.

Science Course-Taking in High School

SMPY males took more semesters of high school science than SMPY females (7.9 vs. 7.1). Although this difference was significant ($p < .001$), its effect size was only .33. Yet there were differences in the types of courses taken. Biology and chemistry were taken most frequently. Fewer students took physics. Of these courses, boys (76%) took physics more often than girls (58%), whereas girls (94%) took biology more often than boys (88%). The sex difference in physics was significant ($p < .01$). In biology, the difference was also significant ($p < .01$) but small.

Of those students beginning the traditional science sequence by taking biology, 90% continued on to chemistry, with 67% completing the se-

TABLE I
Mean SAT Scores of Students in the Talent Search^a and High school

Students	Talent search						High school					
	SAT-M			SAT-V			SAT-M			SAT-V		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
Males	556	73	(1,193)	436	85	(378)	695	70	(1,142)	593	88	(1,140)
Females	519	59	(753)	462	88	(238)	650	71	(700)	599	89	(701)
<i>t</i> value	12.1*			-3.6*			13.3*			n.s.		
College-bound students ^b												
Males	—			—			493			431		
Females	—			—			443			423		

^a Since students could participate in a talent search either as eighth or seventh graders and eighth graders earned slightly higher scores, the talent search SAT scores were equated to be comparable. We used the eighth-grade scores of students participating in a talent search as both a seventh and then as an eighth grader.

^b The sample of 1978-79 high school students taking the SAT was used for comparison purposes. This was the approximate time the SMPY students took the SAT in high school.

* $p < .001$.

quence by taking physics. When broken down by sex, these percentages reflect the greater male participation in physics. Of the boys beginning the sequence, 92% continued to chemistry and 74% completed all three courses. Of the girls, 90% continued to chemistry and 56% completed the sequence. These sex differences were significant, however, only for physics ($p < .001$). Thus, again we find indications of a potential sex difference in physics course-taking.

In addition, there was a small sex difference in the taking of additional science courses beyond the three basic courses. These courses were not enrichment, but advanced courses in biology, chemistry, or physics. The mean for boys was .7 for such courses, whereas for girls it was .5 ($p < .001$, $d = .25$).

A comparison of the mean SAT-M and SAT-V scores of the students participating and not participating in each of the basic sciences courses and in advanced science courses was performed. Four ANOVAs were performed on talent search SAT-M scores and separately on SAT-V scores with sex and participation in each science course or advanced science courses as factors. That is, biology, chemistry, physics, and advanced science were studied individually. Since sex differences in the SAT scores had already been documented, only the possible interaction of sex and participation in a course was of interest.

The association of participation in each science course with a score on the SAT-V was not significant in any analysis, except between students

taking and not taking physics ($F = 5.0, p < .05$). This difference was small (7 points for boys, 26 for girls), however. The interaction of sex and course-taking was not significant for any of the four analyses. Thus, it was concluded that SAT-V scores were not importantly related to the decision to take science.

Between those taking and not taking biology, the *non*participants had slightly higher SAT-M scores ($F = 4.4, p < .05$), but the difference was small (<15 points). The interaction term was not significant. For chemistry, the association of course participation and the interaction of it and sex with SAT-M scores were not significant. The difference on the SAT-M between boys and girls taking physics and those not taking it significantly favored the course-takers ($F = 27.4, p \leq .001$), although again the difference was not large (≈ 18 points). The interaction term between sex and participation in physics was not significant. Finally, talent search SAT-M scores showed no statistically significant differences between those taking and not taking advanced science courses. Moreover, the interaction of sex and the taking of such courses was not significant. We conclude, therefore, that the taking of science courses is not importantly related to ability on the SAT-M for boys or girls. Yet in physics there was a *trend* in the data indicative of such a relationship.

Both sexes obtained similar grades, mostly As and Bs, in science. The sex difference in science grade point average was slight, but significant ($p < .05$) and favored girls. An analysis of SAT-M and SAT-V scores for each level (A–F) of the marks achieved found significant sex differences on the SAT-M favoring males on every level. Even those boys and girls obtaining marks less than C had mean SAT-M scores in high school above 600. No significant differences were found for SAT-V scores. We conclude that, for this population, marks achieved are affected more by other variables than by ability.

Advanced Placement Examination

In high school the most advanced and difficult courses are those that prepare students for the Advanced Placement (AP) Program examination. A high score (at least 3 out of 5 points) on an AP examination often leads to college credit. Moreover, just taking the AP examinations is by itself considered a high level of achievement by many college admissions officers. Few of our students took these tests, however, which made it unlikely that truly significant sex differences would be detected. Yet some favoring males were found (Table II).

In the taking of the AP biology examination, a ratio of more than 2:1 favoring boys was found (Table II). Although the boys tended to score higher on it than the girls, no significant sex difference was found. For chemistry, 7% of the males and 2% of the females reported taking the AP

TABLE II
*Advanced Placement and College Board Achievement Test Scores of the
 Students in High School*

Test	Advanced Placement		College Board Achievement Test	
	Males	Females	Males	Females
Biology				
<i>N</i> (percent)	89 (7.2%)	23 (3.0%)	96 (7.8%)	68 (8.9%)
Mean	4.1	4.0	661	623
<i>SD</i>	0.8	1.0	75	86
Chemistry				
<i>N</i> (percent)	91 (7.4%)	15 (2.0%)	221 (18.0%)	76 (9.9%)
Mean	3.7	3.4	677	643
<i>SD</i>	0.9	1.2	80	75
Physics ^a				
<i>N</i> (percent)	32 (2.6%)	5 (0.7%)	165 (13.4%)	24 (3.1%)
Mean	3.8	2.6	677	608
<i>SD</i>	1.1	1.1	77	83
Physics C, Part 1 (Mechanics)				
<i>N</i> (percent)	49 (4.0%)	1 (0.1%)		
Mean	3.5	4.0		
<i>SD</i>	1.1	—		
Physics C, Part 2 (Electricity and Magnetism)				
<i>N</i> (percent)	33 (2.7%)	1 (0.1%)		
Mean	3.4	3.0		
<i>SD</i>	1.1	—		

^a For the Advanced Placement scores reported, this was Physics B.

examination ($p < .01$). Again the boys scored higher than the girls but not significantly. Finally, 9% of the boys took the AP physics examination, but only one girl did. As a result, comparisons could not be made. Clearly, many more boys than girls took AP science examinations, but no significant differences in scores were detected (Table II).

Thus, after studying the approximately 200 students who participated in the most difficult science courses available in the country, we conclude that the data reveals a trend for males to show greater achievement than females at the highest level of high school science.

College Entrance Examination Board Achievement Tests

Although grades in each science course slightly favored females, this may not imply that their knowledge at the end of the course was greater. The College Board's science achievement tests measure student knowledge and understanding of high school science. Again, more SMPY males than females took the science achievement tests. Even then, males scored significantly higher than females (Table II).

Whereas there were no differences in the taking of the biology achievement test, boys scored higher than girls (Table II). The sex difference was significant at the .01 level with an overall effect size of .46. In chemistry, almost twice as many boys than girls reported scores on the achievement test ($p < .01$). Yet boys scored higher than girls ($p < .001$, $d = .42$).

The greatest difference was found in physics, however, greater even than the sex difference in mathematics (Benbow & Stanley, 1982). About 13% of the males versus 3% of the females took the physics achievement test, significant beyond the .01 level. The mean score for males was 677 and for females 608, which was significant at the .001 level. The effect size was .85.

To test the hypothesis that the above sex differences in science achievement could be accounted for partly by the sex difference in mathematical reasoning ability, an analysis of covariance (ANCOVA), controlling for talent search SAT-M scores, was computed separately for each achievement test. Since there was the possibility that differences in advanced course-taking in science might also contribute to achievement scores obtained, this variable was also tested. For all analyses, the SAT-M score at talent search was strongly related to achievement test scores received in science. The sex difference was no longer significant for the biology and chemistry achievement test scores when, controlling for ability on the SAT-M, the ANCOVA was performed by sex and advanced course-taking. The sex difference was still significant for the analysis on physics achievement test scores ($F = 9.3$, $p < .01$), however. The taking of advanced courses in science did not have a significant relationship with science achievement test scores. Moreover, the interaction of sex and the taking of advanced science was also not significant. Since the sex differences on the science achievement tests were greatly reduced when controlling for the initial sex difference on the SAT-M, the differences in understanding and knowledge of science may be related to the fact that, several years earlier, SMPY females tended to have less well developed mathematical reasoning ability.

Moreover, in a series of analyses not presented here, it appeared that the most successful boys and girls in a science class later take the relevant achievement tests and that course grades do not relate to the sex difference in science achievement test scores.

Stepwise multiple regression analyses were then performed to see if talent search SAT scores, sex, and advanced science course-taking could predict the number of science achievement or AP tests taken. There was a significant sex difference ($p < .001$) in the total number taken (.63 for boys vs. .28 for girls), with the effect size equalling .40. The above three variables could account for 22% of the variance in the taking of these examinations ($R = .47$). By itself, talent search SAT-M, the overall best predictor, could account for 12% of the variance. Advanced course-taking in science was

the second best predictor and could account for an additional 9% of variance. Although sex was a significant predictor in the equation, it could account for less than 1% additional variance. Clearly, sex was not an important predictor of taking science examinations after controlling for the other two variables.

Fairs and Contests

A final area of achievement to be investigated was science fair participation. No significant differences were found. About 18% of both males and females participated in at least one fair. Moreover, among those participating no significant sex differences were found in the number completed. In a previous study, Michael (1983) found a modest negative relationship between SAT-M scores and extent of participation in science fairs for girls, but not for boys. He hypothesized that "science fairs may be social occasions for less able girls." We postulated that perhaps those students participating in a science fair would be more likely to major in a quantitatively oriented area in college [see Benbow & Stanley (1984) for coding]. This contention was not supported by the correlation between the two (i.e., $r = .03$).

Attitudes Toward Science

Students were also asked to rate their liking for the various sciences on a five-point scale, ranging from a strong dislike (1) to a strong like (5). On the average, SMPY students at the end of high school appeared to have moderately strong likings for biology, chemistry, physics, and for science in general. Sex differences were found, however. Girls tended to prefer biology; boys showed no preference. In biology, the females' and males' means were 4.0 and 3.8, respectively ($p < .001$). Boys rated chemistry significantly higher ($p < .001$) than did girls (3.8 vs. 3.6). The physics mean for boys was 3.9, and for girls 3.3, which was also significantly different ($p < .001$). An overall rating for the liking for science was computed by averaging the three ratings. Boys had significantly higher ratings for science in general than girls (3.8 vs. 3.7; $p < .001$). Effect sizes for each difference above, however, were small, except in physics (.50).

At the end of high school 49% and 43%, respectively, of the boys and girls had considered possible careers in science. Although this difference was small, it was statistically significant ($p < .05$). Yet girls had considered careers in several areas of science as often as the boys had. We also studied differences by sex in choosing the sciences as one of their favorite courses in high school. Among the boys 57% chose at least one of the sciences as their three favorite courses, whereas for the girls this was 47% ($p < .01$).

Although there were no large differences in each of the above attitudinal variables, there is a possibility that if the measures were combined some

substantial sex differences would emerge. Moreover, a combination of indicators would be a more reliable method of testing attitude toward science than any one variable alone. Thus, a total attitude toward science score was computed. It was defined as the sum of: number of science areas considered as a possible career, number of science courses named as one of the three favorite courses, number of areas of science for which the student held a strong liking, and number of science fairs participated in. The range of resulting scores was from 0 to 11. The mean difference between boys and girls on this variable was significant (2.7 for boys vs. 2.3 for girls, $p < .001$). The effect size was, however, less than .20. Thus, sex differences in attitudes towards science are not noteworthy for this population.

Total science attitude score did significantly relate to the number of advanced science courses taken ($r = .31$ for boys and $.37$ for girls). Rather surprisingly, those who did not take advanced science courses indicated significantly more positive attitudes towards science than those taking advanced science ($p < .001$). The association effect size was .69. In contrast, total science attitude score did relate somewhat but significantly with the quantitative index of college major ($r = .24$, no sex differences), the number of science achievement or AP examinations taken ($r = .28$, no sex differences), and the taking of physics ($r = .23$, no sex differences). For boys only the total science attitude score related somewhat with physics achievement test scores ($r = .20$). No substantial relationship was found for biology or chemistry achievement scores. These relationships indicate, therefore, that attitudes toward science are related to participation in the sciences, but probably not to sex differences in science participation or achievement.

College Majors

Students were asked to report their intended college majors. Approximately 71% of the males and 59% of the females reported a decision to major in the sciences, engineering, or mathematics. Sex differences were found only in engineering and physics, favoring males, and in biology, favoring females. Benbow and Stanley (1984) investigated sex differences in college majors in this group. No strong explanations were found. When we correlated total attitude score with the quantitative index of intended college major, a moderate correlation was found.

CONCLUSIONS

The six initial SMPY talent searches involving almost 10,000 students unexpectedly revealed large sex differences favoring males in scores achieved on the SAT-M, a test of mathematical reasoning ability (Benbow & Stanley, 1980). A subsequent study of approximately 40,000 students confirmed this finding (Benbow & Stanley, 1983). The Benbow and Stanley

studies (1980, 1982, 1983) have shown that "(1) sex differences in mathematical reasoning ability [favoring males] are found at an early age among mathematically talented students [and] (2) they persist over several years and are related to subsequent differences in mathematics achievement" (Benbow & Stanley, 1982, p. 615). The present study found that the differences in mathematical reasoning ability also related to sex differences in measures of science achievement, and possibly to differences in science course participation in the area of physics. Moreover, this study revealed that the greatest sex differences are found in the area of physics.

Generally, physics has the image of being the most demanding, most mathematically rigorous of the high school sciences. This perception of physics may have influenced many of those students (both boys and girls) of lesser ability not to participate, since for physics the differences on the SAT-M between course-takers and noncourse-takers for both boys and girls approached 20 points, even in this select sample. Since boys scored higher on the SAT-M and more boys than girls took physics, it may be that the sex difference on the SAT-M has an important effect on the decision to participate in high school physics. This did not appear to be the case for advanced science courses.

Although course-taking and course-grades in science were similar for boys and girls, knowledge of science as measured by achievement tests was not. SMPY boys achieved at a higher level than SMPY girls, especially in physics. Since, for those taking the tests, participation (generally, only those taking a course in a subject take the test) and grades earned were similar, course-taking cannot account for sex differences in science achievement. Moreover, advanced course-taking in science was not associated with scores.

Support was found, however, for the hypothesis that the sex difference in talent search SAT-M scores could account partly for subsequent differences in science achievement. When controlling for the SAT-M score measured in the seventh or eighth grade, sex differences on the College Board achievement tests either were no longer significant or were reduced greatly. Additionally, the sex differences in mathematical reasoning ability could partly account for fewer females taking the science achievement or AP examinations. Taking advanced science courses also related to the completion of such tests.

Attitudes toward sciences, which related somewhat to number of advanced science courses and science achievement tests taken, quantitative index of college major, and the taking of physics, exhibited only slight sex differences at the end of high school. Generally, SMPY students appeared to have moderately strong likings for the sciences. Trends in the data indicated, however, that there were sex differences favoring girls in biology and favoring boys in chemistry and physics. These findings are consistent

with Walberg (1969). He found that boys prefer inanimate aspects of science and were attracted to activities involving physical manipulation, whereas girls preferred the more animate aspects. Such differences, along with the greater mathematical reasoning ability of males, may explain partially the large sex difference in physics achievement. In physics SMPY males took more courses, had higher achievement test scores, had more positive attitudes, and more often intended to major in it than did SMPY females.

Some limitations of this study should be noted. One important caution is that this study is based on self-reported data. The possibility exists of a sex-related difference in accuracy of recall. Hamilton (1981) found that females tend to exaggerate (positively) more often than do males. If this is true for this population, our results may be underestimating the actual magnitude of the sex difference favoring males.

Another limitation is the nature of our respondents. All of the students involved were mathematically talented and most were highly motivated. Therefore, we urge careful consideration before generalizing to the general population of all students. For that population of students considered most likely to enter the sciences at colleges with demanding curricula, however, the SMPY students are an appropriate nonprobability sample. Finally, our results are based on group data and therefore it is not appropriate to apply them to individual cases.

We conclude, therefore, that sex differences in science achievement, especially physics, are found, even in a population that is highly able and quantitatively oriented. These differences did relate to the sex difference in mathematical reasoning ability found approximately five years earlier but not to course-taking or course grades. Sex differences in attitudes toward science were slight, but attitudes did relate to various measures of science achievement. Our results may bear on why women are underrepresented in the sciences.

REFERENCES

- ANGOFF, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement*, 2nd ed. (pp. 508–600). Washington, DC: American Council on Education.
- BENBOW, C. P., & STANLEY, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, 210, 1262–1264.
- BENBOW, C. P., & STANLEY, J. C. (1982). Consequences in high school and college of sex differences in mathematical reasoning ability: A longitudinal perspective. *American Educational Research Journal*, 19, 598–622.
- BENBOW, C. P., & STANLEY, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science*, 222, 1029–1031.
- BENBOW, C. P., & STANLEY, J. C. (1984). Gender and the science major. In M. W. Steinkamp & M. L. Maehr (Eds.), *Advances in motivation and achievement*, Vol. 2 (pp. 139–164). Greenwich, CT: JAI Press.

BENBOW AND MINOR

- HAMILTON, L. C. (1981). Sex differences in self-report errors: A note of caution. *Journal of Educational Measurement, 18*, 221–228.
- MICHAEL, W. B. (1983). The manifestation of creative behaviors by maturing participants in the Study of Mathematically Precocious Youth (SMPY). In C. P. Benbow & J. C. Stanley (Eds.), *Academic precocity: Aspects of its development* (pp. 38–50). Baltimore, MD: Johns Hopkins University Press.
- National Science Foundation. (1984). *Women and minorities in science and engineering*. Washington, D.C: Author.
- WALBERG, H. J. (1969). Physics, femininity, and creativity. *Developmental Psychology, 1*, 47–54.
-

AUTHORS

- CAMILLA PERSSON BENBOW, Associate Professor, Iowa State University, Department of Psychology, Ames, IA 50011. *Specializations: Individual differences, intellectual talent, sex differences.*
- LOLA L. MINOR, formerly at Johns Hopkins University, Department of Psychology, Baltimore, MD. (Deceased).