DOCUMENT RESUME

ED 323 259 TM 015 503

AUTHOR Becker, Douglas F.; Forsyth, Robert A.
TITLE Gender Differences in Academic Achievement in Grades 3 through 12: A Longitudinal Analysis.
PUB DATE Apr 90
NOTE 26p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Ability; *Academic Achievement; *Achievement Tests; *Elementary School Students; Elementary Secondary Education; Females; Individual Development; Information Utilization; Language Usage; Longitudinal Studies; Males; Mathematics Achievement; Problem Solving; Reading Achievement; *Secondary School Students; *Sex Differences; *Standardized Tests; State Programs; Student Development; Testing Programs; Vocabulary
IDENTIFIERS Iowa Tests of Basic Skills

ABSTRACT

Gender differences in achievement were studied in five academic areas: (1) vocabulary; (2) language usage; (3) reading; (4) mathematics problem solving; and (5) using sources of information. Standardized achievement test scores for a sample of 3,002 students (1,642 females and 1,360 males) who were tested in each of 10 consecutive years, grades 3 through 12, were used to assess these differences. Data were from the Iowa Statewide Testing Program, using the Iowa Tests of Basic Skills and the Iowa Tests of Educational Development. Selected percentiles used to estimate both female and male score distributions for each subtest area were: 90th; 75th; 50th; 25th; and 10th. Results show fairly consistent patterns of differences. Males generally performed better at the upper percentile leve_ in vocabulary and mathematics problem solving, and females performed better at lower percentile levels in all content areas, as well as at all percentile levels of the score distribution and across all grades for the language usage area. Developmental patterns of differences between male and female performance appeared when different ability levels were examined. Seven tables and five graphs presen· study data. (SLD)

Gender Differences in Academic Achievement in Grades 3 through 12: A Longitudinal Analysis

Douglas F. Becker
Robert A. Forsyth
The University of Iowa

Paper presented at the Annual Meeting of the
American Educational Research Association
Boston, MA
April 1990

Gender Differences in Academic Achievement in Grades 3 through 12: A Longitudinal Analysis

Douglas F. Becker
Robert A. Forsyth
The University of Iowa

ABSTRACT: This study describes gender differences in achievement in five academic areas: vocabulary, language usage, reading, mathematics problem solving, and using sources of information. Standardized achievement test scores for a sample of 3002 students who were tested in each of ten consecutive years, grades 3 through 12, were used to assess these differences. Selected percentiles (90th, 75th, 50th, 25th, 10th) were estimated for both female and male score distributions for each subtest area. The results indicate fairly consistent patterns of differences. Males generally performed better at the upper percentile levels in vocabulary and mathematics problem solving, and females performed better at the lower percentile levels in all content areas, as well as at all percentile levels of the score distribution and across all grades for the language usage area.

Introduction

Research on human performance, including research on gender differences in educational achievement, has been of considerable interest to educators for many years (Baker, 1987). In an early study, Stroud and Lindquist (1942) described gender differences in test scores on the *Iowa Every-Pupil Tests of Basic Skills* shortly after the battery was developed. However, as Halpern (1986) notes, the literature related to gender differences in academic performance often contains inconsistent findings, contradictory theories, and claims that are unsupported by the research.

Since the publication of Maccoby and Jacklin's 1974 book, *The Psychology of Sex Differences*, it has been generally accepted that differences between [illegible] and females do exist with respect to measured verbal and quantitative abilities. Maccoby and Jacklin's conclusions, though, are stated in terms of developmental trends, reflecting a synthesis of short-term studies that may not have been comparable across age levels. Further, most studies of gender differences in achievement have used a cross-sectional design. A cross-sectional design has serious limitations, however, because age and cohort differences can be confounded with other differences. A longitudinal design that enables comparisons to be made for a fixed group of subjects would be more desirable for the tracking and study of gender differences in academic achievement. Longitudinal studies of gender differences, though, are few in number, and yet, these are the studies needed to explore the emergence of gender differences and to ensure that previously reported differences were not due to sample differences.

One of the most extensive longitudinal studies of gender differences in achievement was undertaken by Martin and Hoover (1987). They studied the relative achievement of more than 9000 males and females over a six year period, grades 3 through 8, using scores from the *Iowa Tests of Basic Skills* (ITBS). The purpose of the present study was to extend the Martin and Hoover study through the high school grades. Specifically, this study sought to describe the differential achievement performance of males and females from grades 3 through 12 using the same basic population as Martin and Hoover.

The size of a high school's enrollment is frequently cited as a variable of interest in research on academic performance (e.g., Kleinfeld et al., 1985; Edington, 1981). Studies that have examined the relationship between school size and academic achievement, however, have yielded varied results. Kleinfeld et al. (1985) concluded that high school size was not an important factor with respect to achievement in standardized tests. Edington (1981), on the other hand, noted that students from small schools tended to achieve the highest scores on the ACT. The present study also sought to investigate the influence of high school size on the achievement of males and females in grades 9 through 12.

Related Research

Most previous studies investigating gender differences in academic achievement have concentrated on reading and mathematics. As Martin and Hoover note (1987), other content areas such as language and work-study skills have received relatively little attention even though these skills are an important part of the school curriculum. Also, conclusions regarding differential verbal and quantitative performance are often reported in terms of broad generalizations that imply that these differences occur in all verbal and quantitative areas and at all ability levels. Friedman (1989), for example, concluded from a meta-analysis of studies on gender differences in mathematical tasks that the average difference was very small and that gender differences in performance had decreased over the years. Linn and Hyde (1988, 1989) concluded from their meta-analyses that cognitive gender differences had declined and no longer existed for verbal ability, spatial visualization, and mathematics computation and concepts.

Based on the results of Martin and Hoover (1987), it appears that there may be an interaction between the ability level of students and score differences between males and females. For instance, in those areas where males had higher scores than females, Martin and Hoover found that the largest differences in favor of males occurred above the medians of the score distributions. When females showed superior performance, the greatest differences occurred below the medians of the distributions. Such interactions are not captured or reflected in analyses of mean differences.

In a recent longitudinal study of gender differences, Wentzel (1988) investigated the performance of 30 males and 30 females in math and English as measured by classroom grades and standardized test scores from grade six through grade twelve. Results from that study suggested "that female and male classroom performance remains relatively stable throughout high school, whereas female mean standardized test scores decline in relation to male mean scores over this same period of time" (p. 697).

Two other longitudinal studies dealing with gender differences in educational achievement were undertaken by Lewis and Hoover (1983) and Martin and Hoover (1987). These studies examined differences at selected percentiles as well as at the mean. Lewis and Hoover compared ITBS scores obtained at the fourth, sixth, and eighth grades and *Iowa Tests of Educational Development* (ITED) scores at the eleventh grade for males and females who later entered college. They investigated gender differences in five general content areas: vocabulary, reading, language usage, mathematics problem solving, and using sources of information. For their sample, females consistently scored higher than males in the language usage area. For the mathematics problem solving area, the average score for males was significantly higher than the average for females only at Grade 11. The differences in favor of males in mathematics achievement were greater at the upper part of the score scale than at the lower part.

As noted above, Martin and Hoover (1987) investigated the relationship between gender and achievement using a sample of 9372 males and females who were tested with the ITBS each year from grade 3 through grade 8. In their sample, females consistently had higher average scores in Spelling, Capitalization, Punctuation, Language Usage, Reference Materials, Mathematics Computation, and Reading Comprehension; the differences between males and females in these areas was most pronounced at the lower end of the score scales. Males showed higher achievement on the Visual Materials, Mathematics Concepts, and Mathematics Problem Solving tests; the largest differences in these areas occurred at the upper portion of the score scales.

The results of the two Iowa longitudinal studies appear to be consistent with recent results from the National Assessment of Educational Progress (NAEP). NAEP's "Reading Report Card" (Applebee et al., 1988) found that the average reading proficiency of 9-, 13-, and 17-year-old females (grades 3, 7, and 11) was "somewhat better" than males for the NAEP reading assessments (1971,75,80,84,86). NAEP's "Mathematics Report Card" (Dossey et al., 1988) noted that the mathematics achievement of males was higher than that of females for each assessment (1973,78,82,86). NAEP also reported that by 1986, gender differences in mathematics performance appeared to be minimal at ages 9 and 13, and only slightly greater at age 17.

Although Denno (1982) noted that consistent longitudinal sex differences had not been found in general (or composite) tests of intellectual ability, the Lewis and Hoover, and Martin and Hoover studies present empirical evidence that suggest a degree of consistency in certain academic areas. The present study sought to determine if the differences observed by Martin and Hoover continued throughout the high school years.

Data Source and Method

The data for this investigation were derived from records of students in schools that had participated in the Iowa Statewide Testing Programs each year from 1978-79 through 1987-88. Participating high schools (grades 9-12) were identified as belonging to one of three size categories: large (350 or more students), medium (200-349 students), and small (199 or less students). Only students who were tested in each of the ten years were included in the study. A sample of 3002 students - 1642 females and 1360 males - was identified.

The achievement tests used in the Iowa Statewide Testing Programs are the *Iowa Tests of Basic Skills* (Hieronymus et al., 1978) for grades 3 through 8, and the *Iowa Tests of Educational Development* (Feldt et al., 1979,1986) for grades 9 through 12. Scores for five major content areas were utilized in this study: vocabulary, language usage, reading, mathematics problem solving, and using sources of information. The subtests from the two batteries used to measure achievement in these areas are shown below:

1) ITBS: Vocabulary (V)
   ITED: Vocabulary (V)

2) ITBS: Language Total (LT) - including Spelling, Capitalization, Punctuation, and Usage and Expression
   ITED: Correctness and Appropriateness of Expression (E)

3) ITBS: Reading Comprehension (R)
   ITED: Ability to Interpret Literary Materials (L)

4) ITBS: Mathematics Problem Solving (M2)
   ITED: Ability to do Quantitative Thinking (Q)

5) ITBS: Reference Materials (W2)
   ITED: Using Sources of Information (SI).

In addition, trends in gender differences from grades 9 through 12 were examined using two other ITED subtests: Analysis of Social Studies Materials (SS); and Analysis of Natural Science Materials (NS). (A description of the various subtests used in this study can be found in the ITBS *Manual for School Administrators* (1986) and the ITED *Teacher, Administrator, and Counselor Manual* (1989).)

Summary statistics including means and standard deviations for ITBS grade equivalent (GE) scores and ITED standard scores were calculated separately for males and females, as well as for the total sample, for each subtest area. Selected percentiles (90th, 75th, 50th, 25th, and 10th) were also estimated for both male and female score distributions. The selection of these five commonly used percentile points permitted the study of the interaction between achievement level and gender over time. For each grade and subtest, differences between corresponding statistical indices for females and males were computed. Each difference was standardized by dividing the obtained difference by the total (females and males combined) standard deviation. The total sample standard deviation was used, as opposed to a pooled standard deviation, in the derivation of the effect size in order to remain consistent with the Martin and Hoover analysis. Effect sizes were also computed using the pooled within groups standard deviation. Differences between the two effect sizes were extremely small. All analyses in this study, therefore, are reported using the total sample standard deviation in the computation of effect size. A positive standardized difference reflects a higher female score while a negative standardized difference indicates a higher male score. Given the differences in the variability of the ITBS GE-score scales across grades and the different score scales used with the ITBS and ITED, the standardized differences provide a useful index for examining longitudinal trends.

Although this study sought to investigate gender differences in achievement, the possibility of a school size effect on these differences was also of interest for the grade 9 through grade 12 data. Therefore, summary statistics and standardized differences were computed by school size using the previously noted size categories.

Results and Discussion

The means and standard deviations of the ITBS grade equivalent scores and the ITED standard scores are reported in Tables 1 and 2, respectively. It should be noted that the combined male and female sample was not representative of all Iowa students. The composite score mean for the combined sample was about 0.38SD units above the mean for all participants in the Iowa program for Grade 3 in 1978-79. The composite score for females in the sample was about 0.31SD units above the mean for all females tested in Grade 3 while the composite score for males in the sample was about 0.41SD units above the mean for all males tested in Grade 3 in 1978-79.

The mean differences between the sample and the population suggest that students who remain in the same school system for their primary and secondary education may have a higher achievement level than those that do not remain in the same system. Regardless, the differences between sample means and population means for Grade 3 were similar to those reported by Martin and Hoover, and would not appear to be critical to this investigation. Because changes in gender differences across grades for an intact group of males and females are the primary focus of this study, the differences should not have a marked effect on the longitudinal trends that are of interest here.

Insert Tables 1 and 2 about here

As noted, the sample standard deviations for males and females are reported in Table 2. The ratios of male to female standard deviations are also reported. These ratios indicate that the scores of males are more variable than those of females in all test areas and at all grade levels. These results are generally consistent with the results reported by Lewis and Hoover (1983) and Martin and Hoover (1987). Maccoby and Jacklin (1974) also noted that the numerical ability scores for males were consistently more variable than female scores. However, they indicated that verbal ability scores did not demonstrate the same consistency. . These differences in variability suggest that examining differences in means alone can not provide a clear picture of the relationship between gender and achievement. A better understanding of gender differences in achievement can be achieved by comparing the entire score distributions of males and females. Such comparisons are presented in Table 3. This table reports standardized differences between females and males at selected percentiles and at the mean. In Table 3, positive differences indicate higher scores for females. These results are discussed by test area below.

Insert Table 3 about here

Vocabulary

The standardized differences between means in vocabulary show a fairly constant but small advantage for males starting at Grade 5 and, except for Grade 10, extending through Grade 12. The differences at the 90th and 75th percentiles indicate an even greater advantage for males. This advantage, however, in effect disappears at the 25th percentile, and in fact, at the 10th percentile the advantage is in favor of females in all but three grades.

Language Usage

The standardized differences between female and male means in the language usage area are somewhat more pronounced than in other content areas, and indicate a difference in favor of females in all grades and at all five selected percentiles. There is also a slight tendency for these differences to increase across grades. Females performed about 0.32SD units higher than males at Grade 3, 0.39SD units higher than males at Grade 8, and 0.44SD units higher than males at Grade 12.

As noted above, the standardized differences at the selected levels of the score distributions reflect a female advantage. These differences, however, are smaller in magnitude and do not exhibit a tendency to increase across grades at the upper percentiles. The standardized difference at the 90th percentile is 0.18 at Grade 3, and decreases to 0.14 at Grade 8. This difference increases to 0.22 at Grade 9, but then decreases to 0.12 at Grade 12. This same general pattern can be noted for the 75th percentile as well. The standardized differences at the lower percentiles do exhibit the characteristics noted for the means. The difference at the 10th percentile, for instance, is 0.36 at Grade

3, and increases to 0.61 by Grade 8. A slight drop is noted at Grade 9, with a difference of 0.45, but this difference is 0.65 at Grade 12.

Reading

In reading, the standardized differences between female and male means again indicate an advantage in favor of females, with the difference at Grade 3 being larger than the differences for grades 4 through 8, which are somewhat constant. The increase in the magnitude of the standardized differences at grade 9 may be attributable to the differing content in the two tests, but the pattern of somewhat constant differences noted for grades 4 through 8 continues for grades 9 through 12.

The standardized differences at the selected percentiles present an interesting pattern. The largest differences across grades 3 through 8 occur, again, in Grade 3 for all percentile levels except the 10th. Further, the differences at the 90th and 75th percentiles indicate relatively similar values for males and females. A female advantage can be observed at the lower percentiles across grades 3 through 8. However, in grades 9 through 12 the female advantage can be observed at all levels, with the largest differences occurring at the 25th and 10th percentiles.

Mathematics Problem Solving

The standardized differences between female and male means in the mathematics problem solving area show a small and fairly constant male advantage for grades 3 through 8. A substantial increase is encountered at grade 9, and again, this may reflect content differences between the two tests. The difference in mean performance between males and females on the ITED Quantitative Thinking test is fairly constant, as noted in grades 3 through 8, but is also somewhat pronounced, with males performing on average (across four grades) abo ' 0.34SD units higher than females.

The standardized differences at the selected percentiles provide a somewhat different picture than one might expect from simply examining the mean values. There is a definite male advantage across all grades at the upper portion of the score scale. Yet, at the lower end of the scale the advantage seems to shift slightly in favor of females for grades 3 through 8, with an average standardized difference of about 0.10SD at the 10th percentile. In grades 9 through 12, all standardized differences still favor males; however, at the 10th percentile these differences are relatively small compared to the differences at other percentiles.

Using Sources of Information

In the sources of information area, the female average mean is about 0.21SD units above the male mean for grades 3 through 8, however, the differences in grades 9 through 12 are not as large. The differences at the selected percentiles exhibit a pattern similar to the other test areas. For all practical purposes, there is no difference between females and males in grades 3 through 8 at the 90th percentile. The average standardized difference at the 90th percentile across grades 9 through 12, however, is about 0.14SD units in favor of males. Yet at the 10th percentile, there is a clear female advantage at all grades, with an average standardized difference across grades of about 0.43SD units.

Briefly, then, the results for these five content areas provide a fairly consistent picture of the trends in gender differences in achievement. In vocabulary, males consistently scored higher than females at the upper portions of the score distributions across all grades. This same trend was noted in the mathematics problem solving area as well. Females generally held a small advantage at the 10th percentile in vocabulary, however, the results at the lower percentiles for mathematics problem solving are basically inconclusive. The results for reading indicate relatively similar performances for males and females at the upper percentiles, with a definite female advantage noted at the lower percentiles across all grades. Definite female advantages could be found at all levels of the score distributions and across all grades for the language usage area, however the advantage was less at the 90th percentile than at other percentiles. The results for using sources of information are somewhat inconclusive; at the 90th and 75th percentiles a slight female advantage could be noted across grades 3 through 8, yet a male advantage was observed across grades 9 through 12 for the same percentiles. A definite female advantage could be noted across all grades, however, at the 25th and 10th percentiles.

Social Studies and Science

As noted earlier, trends in gender differences with respect to the social studies and science subtests of the ITED were also examined in this study. Of course, data for these two subtests were only available for grades 9 through 12. Tables 4, 5, and 6 provide summary data and standardized differences between the female and male distributions.

Insert Tables 4, 5, and 6 about here

The standardized differences between female and male means across grades 9 through 12 in social studies indicate relatively little difference in performance. The differences at the 90th percentile indicate a male advantage that diminishes slightly from Grade 9 to Grade 12. At the 25th and 10th percentiles, a female advantage is observed. with the standardized differences in favor of females increasing slightiy from Grade 9 to Grade 12. The average standardized difference in favor of males at the 90th percentile is about 0.24SD units while the difference in favor of females at the 10th percentile averages about 0.24SD units as well.

The standardized differences between female and male means in science suggest a male advantage that diminishes slightly across grades 9 through 12. This pattern generally holds at the upper three percentiles. At the 10th percentile, however, a female advantage is once again noted, with this advantage increasing slightly from Grade 9 to Grade 12. The average standardized difference at the 90th percentile is about 0.32SD units in favor of males while the difference at the 10th percentile averages about 0.15SD units in favor of females across grades 9 through 12.

Size Effects on Gender Differences

To examine the relationship between high school size and gender differences, high schools were sorted into the size categories noted previously. The results for this phase of the study are, at best, inconclusive. Table 7 shows the average standardized difference across grades 9, 10 , 11, and 12 for each ITED subtest. Differences are provided for the total sample (TS), as well as for the three school size categories (S1, S2, S3). No clear patterns or trends are evident in the data. Two factors that might account for this lack of association are the relative homogeneity of Iowa high schools with respect to ITED scores and the particular size categories used.

Insert Table 7 about here

Summary

The results of this study generally concur with the results of a previous study on gender differences conducted by Martin and Hoover (1987); trends in gender differences noted in grades 3 through 8 generally remain consistent through the high school grades. In addition, these results provide additional empirical evidence to indicate that males generally exhibit greater variability in achievement than females. The results also support the position of Martin and Hoover that differences between the means of males and females often provide little information regarding the relationship between the gender of the students and academic achievement. It seems clear that there is some interaction between achievement level and the magnitude of the differences in achievement between males and females.

The results for the five major content areas suggest some degree of similarity with the Martin and Hoover study. Females, for instance, generally scored higher than males in reading and in language usage. In each of these areas, standardized differences were most pronounced at the lower percentiles of the score distributions across all grades. Males generally scored higher than females in mathematics problem solving. In this area, the greatest differences were observed at the upper percentiles. However, unlike the results reported by Wentzel (1988), no decline in female mean test scores in relation to male mean test scores for the high school grades was observed.

Unlike the Martin and Hoover study, the results of this study showed that males consistently scored higher than females in vocabulary at the upper percentiles of the score distributions across all grades. A small female advantage was noted in all but three grades at the 10th percentile. In addition, the results for using sources of information were generally inconclusive, with females holding a slight advantage over males at the upper percentiles for grades 3 through 8, and males holding an advantage over females at the same percentiles for grades 9 through 12. Females, however, did score higher than males across all grades at the lower percentiles. The phenomenon exhibited at the upper percentiles in the using sources of information area is perhaps the clearest indication of how differing test contents between the ITBS and the ITED might have influenced the results of this study.

The results for social studies and science indicate that male advantages at the upper percentiles of the score distributions tend to diminish slightly over grades 9 through 12. Females, on the other hand, increase their advantage over males at the lower percentiles across the same four grades.

Finally, in this study, the size of a high school was not related to gender differences in academic achievement.

The primary purpose of this study was simply to describe gender differences in academic achievement in several areas, and therefore provides limited information regarding the cause of these differences or the context within which the differences occur. The results indicate, however, that examining male and female performance at different ability levels gives a clearer picture of the relationship between the gender of students and academic achievement than is achieved by merely considering group averages. Further, the results suggest that there appear to be developmental patterns of differences between male and female performance when examined at different ability levels. The causes and/or implications of these patterns of differences should be of interest and concern to researchers and educators, and continued attempts to understand why and how these differences arise would seem to be warranted.

REFERENCES

Applebee, A.N., Langer, J.A., & Mullis, I. (1988). *The Nation's Report Card: Who reads best? Factors related to reading achievement in grades 3, 7, and 11.* Princeton, N.J.: Educational Testing Service.

Baker, M.A. (1987). *Sex Differences in Human Performance.* Chichester: John Wiley & Sons.

Denno, D. (1982). Sex differences in cognition: A review and critique of the longitudinal evidence. *Adolescence, 17*, 779-788.

Dossey, J.A., Mullis, I., Lindquist, M.M., & Chambers, D.L. (1988). *The mathematics report card: Are we measuring up? Trends and achievement based on the 1986 National Assessment.* Princeton, N.J.: Educational Testing Service.

Edington, E.D. (1981). *ACT scores of incoming freshmen to New Mexico State University by high school size.* New Mexico: New Mexico State University. (ERIC Document ED272354.)

Feldt, L.S., Forsyth, R.A., & Lindquist, E.F. (1979). *Iowa Tests of Educational Development, Forms X-7 and Y-7.* Iowa City: The University of Iowa.

Feldt, L.S., Forsyth, R.A., & Alnot, S.D. (1986). *Iowa Tests of Educational Development, Forms X-8 and Y-8.* Iowa City: The University of Iowa.

Feldt, L.S., Forsyth, R.A., & Alnot, S.D. (1989). *Teacher, Administrator, and Counselor Manual: Iowa Tests of Educational Development.* Chicago: The Riverside Publishing Company.

Friedman, L. (1989). Mathematics and the gender gap: A meta- analysis of recent studies on sex differences in mathematical tasks. *Review of Educational Research, 59*, 185-213.

Halpern, D.F. (1986). *Sex Differences in Cognitive Abilities.* Hillsdale: Lawrence Erlbaum Associates.

Hieronymus, A.N., Lindquist, E.F., & Hoover, H.D. (1978). *Iowa Tests of Basic Skills, Forms 7 and 8.* Chicago: The Riversdie Publishing Company.

Hieronymus, A.N. & Hoover, H.D. (1987). *ITBS Manual for School Administrators.* Chicago: The Riverside Publishing Company.

Hyde, J.S. & Linn, M.C. (1988). Gender differences in verbal ability:A Meta-analysis. *Psychological Bulletin, 104*, 53-69.

Kleinfeld, J.S. and others (1985). *Alaska's small rural high schools: Are they working?* Fairbanks: Alaska University, Center for Cross-Cultural Studies.

Lewis, J. & Hoover, H.D. (1983, April). *Sex Differences on Standardized Academic Achievement Tests - A Longitudinal Study.* Paper presented at the annual convention of the American Educational Research Association, Montreal.

Linn, M.C. & Hyde, J.S. (1989). Gender, mathematics, and science. *Educational Researcher,18*, 17-27.

Maccoby, E.E. & Jacklin, C.N. (1974). *The Psychology of Sex Differences.* Palo Alto: Stanford University Press.

Martin, D.J. & Hoover, H.D. (1987). Sex differences in educational achievement: A longitudinal study. *Journal of Early Adolescence, 7*, 65-83.

Stroud, J.B. & Lindquist, E.F. (1942). Sex differences in achievement in the elementary and secondary schools. *Journal of Educational Psychology, 33*, 657-667.

Wentzel, K.R. (1988). Gender differences in math and English achievement: A longitudinal study. *Sex Roles, 18*, 691-699.

TABLE 1
ITBS/ITED MATCHED LONGITUDINAL SAMPLE
MEAN ACHIEVEMENT LEVELS AND FEMALE-MALE DIFFERENCES

GRADE		V/V	LT/E	R/L	M2/Q	W2/SI
3	FEMALES	42.99	51.22	46.58	43.23	45.43
	MALES	42.42	47.69	44.48	43.48	43.38
	DIFFERENCE	0.57	3.53	2.10	-0.25	2.05
	STD. DIFF.*	0.06	0.32	0.18	-0.02	0.19
4	FEMALES	54.18	62.30	58.29	54.17	58.65
	MALES	54.31	58.30	57.11	54.78	56.14
	DIFFERENCE	-0.13	4.00	1.18	-0.61	2.51
	STD. DIFF.*	-0.01	0.35	0.09	-0.05	0.20
5	FEMALES	65.45	74.67	67.91	65.54	70.23
	MALES	66.64	70.60	67.57	66.81	67.82
	DIFFERENCE	-1.19	4.07	0.34	-1.27	2.41
	STD. DIFF.*	-0.12	0.33	0.03	-0.11	0.19
6	FEMALES	75.16	85.19	79.12	77.06	81.28
	MALES	76.67	80.42	78.07	77.27	78.21
	DIFFERENCE	-1.51	4.77	1.05	-0.21	3.07
	STD. DIFF.*	-0.12	0.36	0.07	-0.02	0.22
7	FEMALES	84.68	96.60	88.58	88.45	91.36
	MALES	86.89	91.00	87.49	89.23	87.96
	DIFFERENCE	-2.21	5.60	1.09	-0.78	3.40
	STD. DIFF.*	-0.15	0.37	0.07	-0.05	0.22
8	FEMALES	95.84	108.34	100.78	99.87	103.62
	MALES	98.19	102.29	99.32	101.53	100.34
	DIFFERENCE	-2.35	6.05	1.46	-1.66	3.28
	STD. DIFF.*	-0.15	0.39	0.09	-0.12	0.21
9	FEMALES	15.90	16.51	15.18	14.59	17.83
	MALES	16.08	14.78	14.24	16.22	17.64
	DIFFERENCE	-0.18	1.73	0.94	-1.63	0.19
	STD. DIFF.*	-0.04	0.40	0.19	-0.31	0.03
10	FEMALES	17.63	18.31	16.98	16.78	20.12
	MALES	17.50	16.23	16.08	18.36	19.82
	DIFFERENCE	0.13	2.08	0.90	-1.58	0.30
	STD. DIFF.*	0.03	0.46	0.17	-0.29	0.05
11	FEMALES	18.64	20.11	19.84	18.13	22.27
	MALES	19.08	18.14	18.33	20.96	21.77
	DIFFERENCE	-0.44	1.97	1.51	-2.83	0.50
	STD. DIFF.*	-0.09	0.40	0.25	-0.42	0.08
12	FEMALES	19.65	21.22	20.42	19.63	23.33
	MALES	19.97	18.87	18.50	21.97	22.34
	DIFFERENCE	-0.32	2.35	1.92	-2.34	0.99
	STD. DIFF.*	-0.06	0.44	0.30	-0.33	0.14

*STD. DIFF. = (FEMALE MEAN - MALE MEAN) / TOTAL SAMPLE SD
FEMALES: N=1642 ; MALES: N=1360

TABLE 2
ITBS/ITED MATCHED LONGITUDINAL SAMPLE
MALE and FEMALE SDs and RATIOS of SDs

GRADE		V/V	LT/E	R/L	M2/Q	W2/SI
3	MALES	10.32	11.17	11.84	10.41	10.96
	FEMALES	9.83	10.76	11.68	9.83	10.22
	RATIO	1.04	1.04	1.01	1.06	1.07
4	MALES	11.65	11.62	13.38	12.04	12.89
	FEMALES	10.60	10.79	12.45	10.80	11.63
	RATIO	1.10	1.08	1.07	1.11	1.11
5	MALES	12.82	13.00	13.78	12.17	13.29
	FEMALES	11.86	11.60	12.46	11.54	11.58
	RATIO	1.08	1.12	1.11	1.05	1.15
6	MALES	13.56	14.14	14.65	13.32	15.12
	FEMALES	12.33	12.08	13.46	12.30	12.69
	RATIO	1.10	1.17	1.09	1.08	1.19
7	MALES	15.30	15.92	16.40	15.45	16.85
	FEMALES	14.19	14.07	14.33	13.29	13.93
	RATIO	1.08	1.13	1.14	1.16	1.21
8	MALES	15.81	16.68	16.07	16.87	16.90
	FEMALES	15.06	14.13	14.81	14.13	13.58
	RATIO	1.05	1.18	1.09	1.19	1.24
9	MALES	4.66	4.36	5.01	5.60	6.12
	FEMALES	4.28	4.13	4.78	4.94	5.28
	RATIO	1.09	1.06	1.05	1.13	1.16
10	MALES	4.78	4.64	5.60	5.75	6.44
	FEMALES	4.39	4.23	5.11	5.21	5.29
	RATIO	1.09	1.10	1.10	1.10	1.22
11	MALES	5.20	5.23	6.31	7.05	7.11
	FEMALES	4.68	4.41	5.52	6.07	6.01
	RATIO	1.11	1.19	1.14	1.16	1.18
12	MALES	5.35	5.78	6.80	7.44	7.81
	FEMALES	5.03	4.79	5.74	6.40	6.09
	RATIO	1.06	1.21	1.18	1.16	1.28

RATIO = MALE SD / FEMALE SD
FEMALES:N=1642;MALES:N=1360

TABLE 3
ITBS/ITED MATCHED LONGITUDINAL SAMPLE

STANDARDIZED DIFFERENCES FOR TESTS V/V

GRADE	3	4	5	6	7	8	9	10	11	12
MEAN	.06	-.01	-.10	-.12	-.15	-.15	-.04	.03	-.09	-.06
90th %ile	-.04	-.11	-.24	-.25	-.27	-.20	-.21	-.11	-.34	-.13
75th %ile	.02	-.07	-.18	-.19	-.25	-.20	-.09	-.07	-.14	-.15
50th %ile	.07	-.02	-.15	-.09	-.20	-.20	-.04	.04	-.10	-.08
25th %ile	.07	.09	-.03	-.03	-.07	-.10	.02	.08	-.01	-.01
10th %ile	.13	.16	.05	-.07	-.05	-.06	.04	.18	.18	.08

STANDARDIZED DIFFERENCES FOR TESTS LT/E

GRADE	3	4	5	6	7	8	9	10	11	12
MEAN	.32	.35	.33	.36	.37	.39	.40	.46	.40	.44
90th %ile	.18	.21	.12	.13	.13	.14	.22	.29	.25	.12
75th %ile	.29	.27	.24	.23	.22	.23	.34	.36	.29	.30
50th %ile	.40	.47	.44	.53	.54	.57	.46	.52	.53	.60
25th %ile	.41	.40	.39	.39	.43	.42	.47	.48	.40	.49
10th %ile	.36	.40	.43	.52	.50	.61	.45	.59	.67	.65

TABLE 3 (continued)
ITBS/ITED MATCHED LONGITUDINAL SAMPLE

STANDARDIZED DIFFERENCES FOR TESTS R/L

GRADE	3	4	5	6	7	8	9	10	11	12
MEAN	.18	.09	.03	.07	.07	.09	.19	.17	.25	.30
90th %ile	.18	.00	-.17	-.06	-.11	.00	.08	.00	.09	.12
75th %ile	.15	.02	.00	-.03	-.04	.03	.17	.07	.08	.10
50th %ile	.24	.09	.06	.07	.04	.08	.25	.22	.23	.27
25th %ile	.23	.16	.12	.23	.23	.15	.30	.28	.48	.55
10th %ile	.22	.18	.11	.18	.27	.23	.15	.24	.39	.59

STANDARDIZED DIFFERENCES FOR TESTS M2/Q

GRADE	3	4	5	6	7	8	9	10	11	12
MEAN	-.02	-.05	-.11	-.02	-.05	-.11	-.31	-.29	-.42	-.33
90th %ile	-.06	-.23	-.10	-.25	-.32	-.42	-.44	-.26	-.60	-.45
75th %ile	-.03	-.11	-.20	-.09	-.15	-.21	-.41	-.29	-.65	-.46
50th %ile	.02	-.03	-.08	-.06	-.07	-.11	-.42	-.43	-.50	-.47
25th %ile	.00	-.01	-.06	.03	.07	.03	-.15	-.22	-.27	-.17
10th %ile	.06	.10	.01	.08	.20	.12	-.09	-.04	-.12	-.03

TABLE 3 (continued)
ITBS/ITED MATCHED LONGITUDINAL SAMPLE

STANDARDIZED DIFFERENCES FOR TESTS W2/SI

GRADE	3	4	5	6	7	8	9	10	11	12
MEAN	.19	.20	.19	.22	.22	.21	.03	.05	.08	.14
90th %ile	.06	.03	.02	.00	.00	-.03	-.17	-.14	-.11	-.12
75th %ile	.13	.10	.07	.07	.09	.03	-.12	-.09	-.07	-.10
50th %ile	.22	.21	.21	.19	.17	.13	.03	-.02	-.04	.09
25th %ile	.28	.29	.30	.33	.41	.36	.15	.16	.17	.25
10th %ile	.29	.32	.46	.47	.45	.51	.27	.36	.43	.66

STANDARDIZED DIFFERENCE = (FEMALE MEAN - MALE MEAN) / TOTAL SAMPLE SD
Positive differences favor females.

TABLE 4
MEAN ACHIEVEMENT LEVELS, FEMALE-MALE DIFFERENCES, AND TOTAL SAMPLE STANDARD DEVIATIONS

GRADE		TEST SS	NS
9	FEMALES	15.01	16.14
	MALES	15.23	17.23
	DIFFERENCE	-.22	-1.09
	TOTAL SD	4.93	5.22
10	FEMALES	16.83	18.48
	MALES	16.99	19.53
	DIFFERENCE	-.16	-1.05
	TOTAL SD	5.13	5.61
11	FEMALES	19.25	19.89
	MALES	19.38	20.41
	DIFFERENCE	-.13	-.52
	TOTAL SD	6.15	6.37
12	FEMALES	20.11	20.24
	MALE	19.86	20.90
	DIFFERENCE	.25	-.66
	TOTAL SD	6.46	6.64

Positive differences favor females.

TABLE 5
MALE AND FEMALE SDs and RATIOS of SDs

GRADE		SS	NS
9	MALES	5.26	5.57
	FEMALES	4.65	4.86
	RATIO	1.13	1.15
10	MALES	5.62	5.94
	FEMALES	4.69	5.27
	RATIO	1.20	1.13
11	MALES	6.61	6.94
	FEMALES	5.74	5.85
	RATIO	1.15	1.19
12	MALES	7.03	7.21
	FEMALES	5.95	6.10
	RATIO	1.18	1.18

FEMALES:N=1642;MALES:N=1360
RATIO = MALE SD / FEMALE SD

TABLE 6
ITED STANDARDIZED DIFFERENCES AT SELECTED PERCENTILES

GRADE		E	Q	SS	NS	L	V	SI	C	RT
09	MEAN	0.40	-0.31	-0.04	-0.21	0.19	-0.04	0.03	-0.01	0.04
10		0.46	-0.29	-0.03	-0.19	0.17	0.03	0.05	0.01	0.06
11		0.40	-0.42	-0.02	-0.08	0.25	-0.09	0.08	-0.02	0.12
12		0.44	-0.33	0.04	-0.10	0.30	-0.06	0.14	0.03	0.15
09	P90	0.22	-0.44	-0.29	-0.46	0.08	-0.21	-0.17	-0.20	-0.09
10		0.29	-0.26	-0.26	-0.33	0.00	-0.11	-0.14	-0.18	-0.14
11		0.25	-0.60	-0.23	-0.24	0.09	-0.34	-0.11	-0.30	-0.08
12		0.12	-0.45	-0.17	-0.25	0.12	-0.13	-0.12	-0.22	-0.11
09	P75	0.34	-0.41	-0.17	-0.30	0.17	-0.09	-0.12	-0.14	-0.10
10		0.36	-0.29	-0.21	-0.43	0.07	-0.07	-0.09	-0.12	-0.10
11		0.29	-0.65	-0.15	-0.30	0.08	-0.14	-0.07	-0.21	-0.07
12		0.30	-0.46	-0.19	-0.28	0.10	-0.15	-0.10	-0.16	-0.04
09	P50	0.47	-0.42	-0.04	-0.21	0.25	-0.04	0.03	-0.02	0.07
10		0.48	-0.43	-0.04	-0.23	0.22	0.04	-0.02	-0.01	0.07
11		0.40	-0.50	-0.03	-0.13	0.23	-0.10	-0.04	-0.03	0.09
12		0.49	-0.47	0.02	-0.16	0.27	-0.08	0.09	0.02	0.12
09	P25	0.46	-0.15	0.05	-0.10	0.30	0.02	0.15	0.10	0.15
10		0.52	-0.22	0.12	-0.11	0.28	0.08	0.16	0.10	0.19
11		0.53	-0.27	0.07	0.05	0.48	-0.01	0.17	0.13	0.28
12		0.60	-0.17	0.20	0.02	0.55	-0.01	0.25	0.18	0.32
09	P10	0.45	-0.09	0.20	0.05	0.15	0.04	0.27	0.22	0.18
10		0.59	-0.04	0.20	0.02	0.24	0.18	0.36	0.29	0.28
11		0.67	-0.12	0.24	0.22	0.39	0.18	0.43	0.25	0.44
12		0.65	-0.03	0.30	0.19	0.59	0.08	0.66	0.34	0.55

STANDARDIZED DIFFERENCE = (FEMALE MEAN - MALE MEAN) / TOTAL SAMPLE SD
Positive differences favor females.

TABLE 7
ITBS/ITED MATCHED LONGITUDINAL SAMPLE
AVERAGE STANDARDIZED DIFFERENCES ACROSS GRADES
FOR ITED SUBTESTS, AT SELECTED PERCENTILES

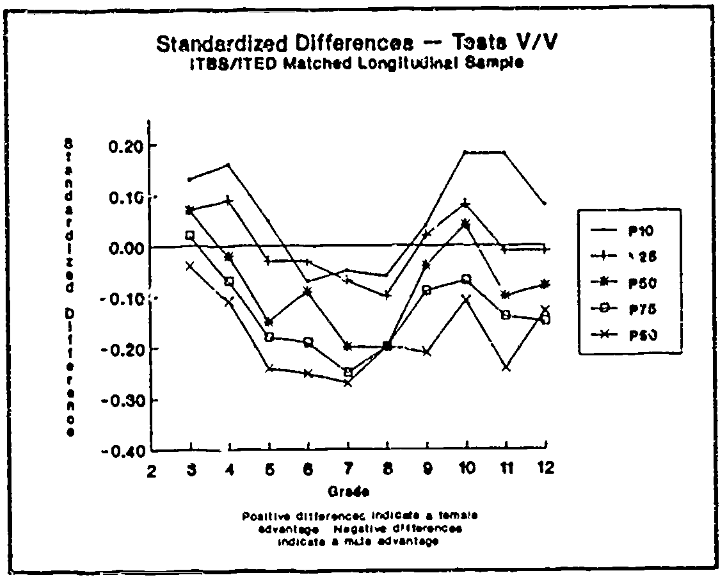
		E	Q	SS	NS	L	V	SI
P90	TS	.22	-.44	-.24	-.32	.07	-.20	-.14
	S1	.24	-.46	-.24	-.31	.07	-.20	-.15
	S2	.18	-.44	-.22	-.39	.10	-.21	-.11
	S3	.20	-.42	-.26	-.30	.07	-.20	-.12
P75	TS	.32	-.45	-.18	-.33	.11	-.11	-.10
	S1	.30	-.44	-.23	-.33	.10	-.13	-.11
	S2	.32	-.43	-.12	-.35	.13	-.07	-.07
	S3	.39	-.50	-.13	-.31	.10	-.10	-.06
P50	TS	.46	-.46	-.02	-.18	.24	-.05	.02
	S1	.43	-.46	-.07	-.20	.22	-.08	-.03
	S2	.49	-.42	-.02	-.18	.25	.02	.06
	S3	.50	-.47	.04	-.15	.26	-.03	.08
P25	TS	.53	-.20	.11	-.04	.40	.02	.18
	S1	.52	-.20	.12	-.03	.43	-.02	.15
	S2	.58	-.12	.09	-.02	.38	.08	.19
	S3	.50	-.26	.13	-.06	.38	.05	.25
P10	TS	.59	-.07	.24	.12	.34	.12	.43
	S1	.59	-.07	.25	.16	.35	.08	.44
	S2	.67	-.01	.23	.12	.36	.17	.44
	S3	.53	-.13	.22	.05	.34	.13	.40

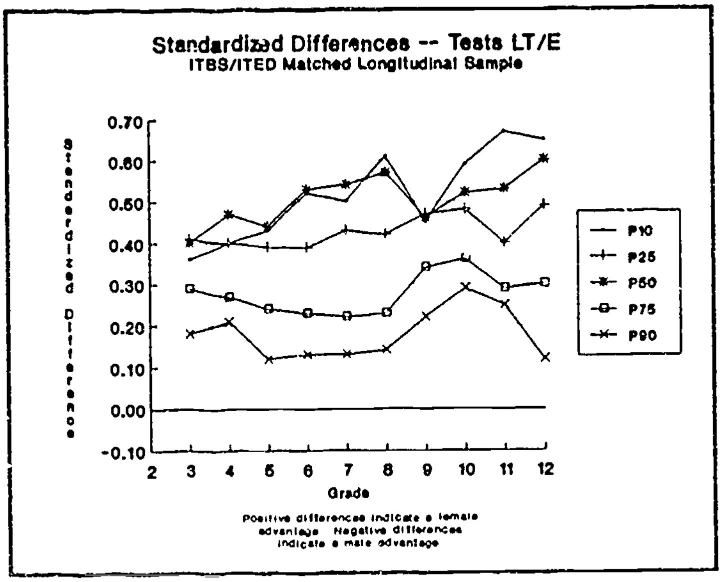
TS - Total Sample — Females:N=1642;Males:N=1360

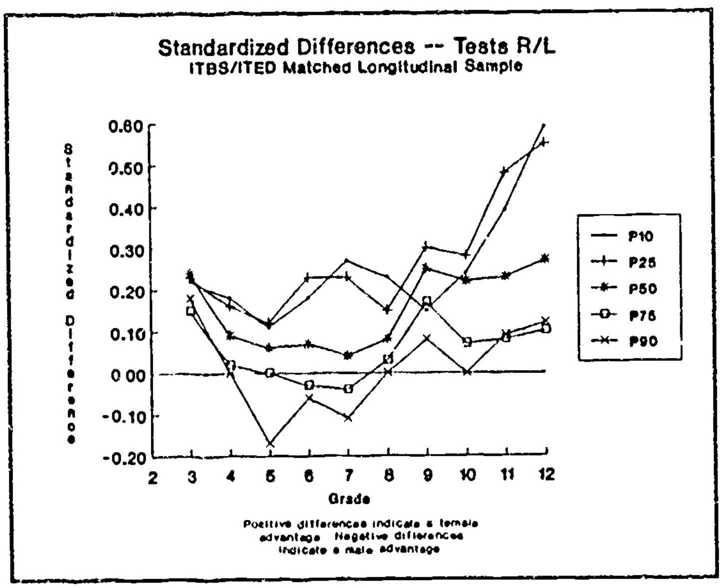
S1 - Size Category 1 (350 or more students) — Females:N=872;Males:N=736

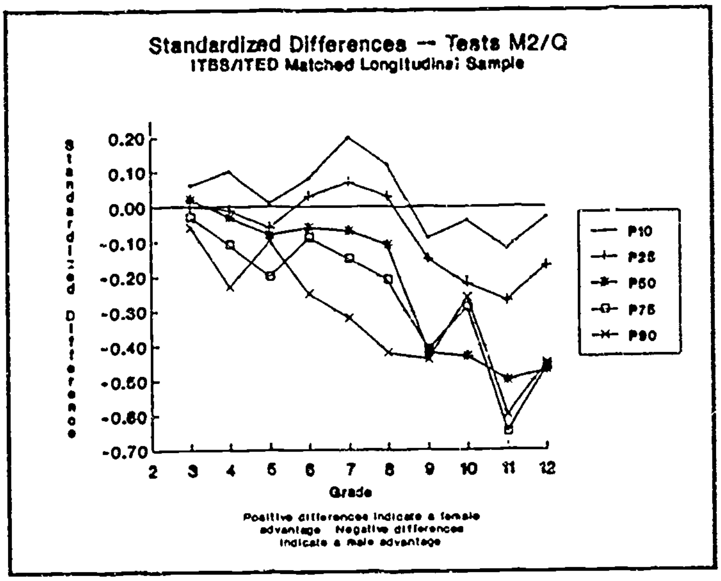
S2 - Size Category 2 (200 - 349 students) — Females:N=342;Males:N=287

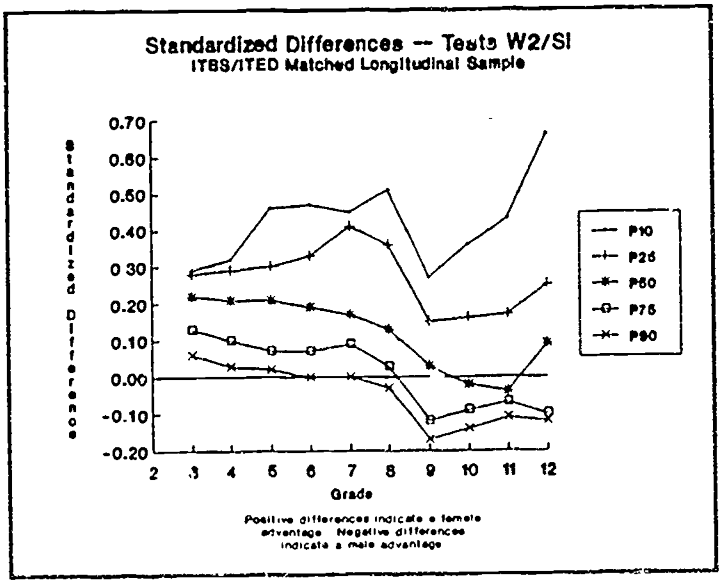
S3 - Size Category 3 (199 or less students) — Females:N=428;Males:N=337


Standardized Differences -- Tests V/V
ITBS/ITED Matched Longitudinal Sample
Standardized Difference
0.20
0.10
0.00
-0.10
-0.20
-0.30
-0.40
2
3
4
5
6
7
8
9
10
11
12
Grade
P10
P25
P50
P75
P90
Positive differences indicate a female advantage. Negative differences indicate a male advantage


Standardized Differences -- Tests LT/E
ITBS/ITED Matched Longitudinal Sample
Standardized Difference
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
-0.10
2
3
4
5
6
7
8
9
10
11
12
Grade
P10
P25
P50
P75
P90
Positive differences indicate a female advantage. Negative differences indicate a male advantage


Standardized Differences -- Tests R/L
ITBS/ITED Matched Longitudinal Sample
Standardized Difference
0.60
0.50
0.40
0.30
0.20
0.10
0.00
-0.10
-0.20
2
3
4
5
6
7
8
9
10
11
12
Grade
P10
P25
P50
P75
P90
Positive differences indicate a female advantage. Negative differences indicate a male advantage


Standardized Differences -- Tests M2/Q
ITBS/ITED Matched Longitudinal Sample
Standardized Difference
0.20
0.10
0.00
-0.10
-0.20
-0.30
-0.40
-0.50
-0.60
-0.70
2 3 4 5 6 7 8 9 10 11 12
Grade
P10
P25
P50
P75
P90
Positive differences indicate a female advantage. Negative differences indicate a male advantage


Standardized Differences -- Tests W2/SI
ITBS/ITED Matched Longitudinal Sample
Standardized Difference
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
-0.10
-0.20
2 3 4 5 6 7 8 9 10 11 12
Grade
P10
P25
P50
P75
P90
Positive differences indicate a female advantage. Negative differences indicate a male advantage