DOCUMENT RESUME

ED 299 194 SO 019 336

AUTHOR Ethington, Corinna A.
TITLE Gender Differences in Mathematics: An International Perspective.
PUB DATE 9 Apr 88
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Females; Foreign Countries; Global Approach; Grade 7; Grade 8; Junior High Schools; Males; *Mathematics; Mathematics Achievement; Mathematics Skills; *Sex Differences
IDENTIFIERS Belgium; British Columbia; France; Japan; New Zealand; Ontario; *Second International Mathematics Study; Thailand

ABSTRACT

The purpose of this study was to investigate whether gender accounts for student differences in performance and achievement in mathematics. Data for the study were drawn from the Second International Mathematics Study (SIMS) of seventh and eighth grade students from Japan, Belgium, Canada (British Columbia and Ontario), France, the United States, New Zealand, and Thailand. Students were tested on the same mathematical items and problems. Results indicated no substantial gender effects or differences in any of the mathematics content areas, problem types, or national origins. Other studies suggested that curricula, pedagogies, or cultural factors may interact with gender differences in impacting quantitative performance. Tables and 21 references are included. (JHP)

GENDER DIFFERENCES IN MATHEMATICS: AN INTERNATIONAL PERSPECTIVE

Corinna A. Ethington

University of Illinois at Chicago

This paper was prepared for presentation at the annual meeting of the American Educational Research Association, New Orleans, 1988.

GENDER DIFFERENCES IN MATHEMATICS: AN INTERNATIONAL PERSPECTIVE

Gender related differences on measures of quantitative performance and problem solving abilities consistently appear in national assessments (e.g., Fennema and Carpenter, 1981; National Assessment of Educational Progress, 1975, 1983; Wilson, 1972). Using a variety of performance measures, investigators have examined the nature of these differences and factors associated with them for subjects varying in age from elementary school to undergraduates in college. From these studies, it is generally concluded that no gender differences are evidenced at the elementary school level, but beginning at approximately the seventh grade, any differences that appear favor males (see Fennema (1974, 1980) for a review of this literature.)

However, gender differences in mathematics achievement have not been found to be consistent across countries. Walberg, Harnisch, & Tsai (1986) found differences favoring males after controlling for productivity factors in eight of the twelve countries studied and no gender differences in the remaining four. In the first IEA study, Husen (1967) found differences in achievement to generally favor males, but differences within countries were not always significant. Husen also noted, however, that gender differences were a within-country phenomenon and that across countries, girls may be superior to boys. These between country differences would be attributable to curricular and instructional differences which mirror cultural values. In contrast to these predominant findings of superior performance by males, in a study of Hawaiian students, Brandon, Newton, & Hammond (1987) often found differences favoring females among Japanese-American, Filipino-American, and Hawaiian students but not among Caucasian students. It is such findings that lend support to the suggestion made by Leder (1986) that "a clear recognition of the values, expectations and beliefs of the wider society within which learning takes place is required for a full appreciation of the currently found sex differences in mathematics participation and performance" (p. 6).

Just as higher performance by males on measures of overall mathematics achievement have not been consistent across countries, another body of research notes that the size and direction of gender differences in quantitative performance vary according to problem type. These studies have been conducted with subjects of varying ages and educational backgrounds and have resulted in reasonably consistent conclusions. Several types of problems have been identified in which gender differences appear. Males have been found to excel in problems dealing with measurement and proportionality (Bart, Baxter, & Frey, 1980; Fennema, 1980; Fennema & Carpenter, 1981; Pattison & Grieve, 1984; Wood, 1976) and in problems with a spatial component (Fennema, 1980; Fennema & Carpenter, 1981; Pattison & Grieve, 1984), whereas females were found to perform better on items testing computational skills (Fennema, 1974; Jarvis, 1964; Meece, Parsons, Kaczala, Goff, & Futterman, 1982) and those involving more abstract deductive reasoning such as the algebra of sets (Wood, 1976) and problems involving the construction and analysis of symbolic relationships (Pattison & Grieve, 1984.) Additionally, Maier & Casselman (1971) conclude that while males consistently score higher overall than females on problem solving tests, women's best performance was on idea-getting problems rather than on problems that required making essential distinctions.

Wood (1976) noted that in some of the schools involved in his study, gender differences within the schools were greater than those for the whole sample, while within others the differences vanished or reversed. He subsequently suggested that a fundamental factor in the presence or absence of gender differences may be the style of instruction. If differences occur within countries, it is also possible that the performance by males and females on specific item types could vary across country just as performance on overall achievement measures do. Should this be the case, it would suggest that the curricula, pedagogy, and culture interact with gender in affecting quantitative performance.

The recently completed Second International Mathematics Study (SIMS) provides an opportunity to determine if the cultural differences suggested by Leder (1986) are manifested in item type differences as well as in overall performance. Twenty-four countries around the world participated in this comprehensive study of school mathematics in which students of approximately the same age and grade level were administered the same core items. Individual item performance was recorded for each student, and thus, performance on item type as well as overall performance can be compared across country as well as gender. The purpose of this study was to investigate whether the patterns of gender differences on specific problem types evidenced in previous studies were consistent across the countries involved in the SIMS longitudinal study.

DATA

Data for this study were drawn from the Population A longitudinal data file of the Second International Mathematics Study (SIMS), a comprehensive survey of the teaching and learning of mathematics in countries around the world. The longitudinal data file contained data for eight of the twenty-four countries involved in the study. They were Japan, Flemish Belgium, British Columbia, France, Ontario, United States, New Zealand, and Thailand. Population A was defined to be the eighth grade in the United States and other countries and seventh grade in Japan. Population A represents a grade level where approximately all of the students in most of the participating countries are still studying mathematics in a common program (Crosswhite, Dossey, Swafford, McKnight, Cooney, Downs, Grouws, & Weinzweig, 1986).

Students were tested at the beginning and end of the 1981-82 academic school year using internationally developed mathematics achievement tests. The items on the achievement tests used in the SIMS study were developed such that the mathematics curriculum of each participating country was adequately sampled. All of the students in Population A took a core test and one of four rotated forms constructed using item sampling procedures. Post-test performance on the core items common to each country was analyzed in the present study. Using the longitudinal form construction strata, the items were clustered according to content areas. The content areas were fractions, ratio/proportion/percent, algebra, geometry, and measurement. The percent of items correct within each cluster was then computed for each student and averaged across country by gender. Table 1 presents these averages.

Insert Table 1 About Here

METHODOLOGY

Analyses were conducted using the exploratory data analysis method of median polishing (see: Tukey, 1977; Velleman and Hoaglin, 1981). The exploratory approach of this method does not test hypotheses but involves a decomposition of the data, producing patterns of effects that are not necessarily apparent in the summary data.[1] The median polish decomposes the data into a common effect, an effect associated with gender, an effect associated with country, and a residual. The common effect is interpreted as a typical score for the entire sample of students. The gender and country effects then indicate performance relative to this typical score that would be expected for a student of specified gender and nationality.

The model used in this study is similar to the additive model of analysis of variance but uses medians rather than means to describe common effects, row effects, and column effects. For the factors involved in this study the model is

$$X_{ij} = M + G_i + C_j + e_{ij}$$

where X_{ij} is the mean proportion for gender i in country j; M is the common effect (median across countries); G_i is the effect of gender i; C_j is the effect of country; and e_{ij} is a residual. The residual indicates how well the model describes the data. Extraordinary values in the table will, after fitting the model by median polish, leave residuals that stand out from other residuals. Differences in country effects would not only reflect differences in student performance across countries, but curricular, pedagogical, and cultural differences as well.

RESULTS

The results of the median polishes of the mean percentages of correct responses over all items and within each content area are given in Table 2. The first panel shows the results over all items. The common effect of 55.64 can be interpreted as the typical percent of items answered correctly for this sample of students, and the gender and country effects indicate increments in performance relative to this typical score as a result of membership in particular categories. The cell residual indicates that portion of the mean score not accounted for by the common value, gender, and country. For example, the mean for females in the United States shown in Table 1 can be expressed as:

$$48.58 = 55.64 + (-.16) + (-7.28) + .39.$$

The mean percent correct for females in the United States is composed of the common value of 55.64, from which .16 is subtracted for being female and 7.28 is subtracted for being from the United States, and .39 residual points. The small residual indicates that the model fit for this group is good. Aside from the common effect, the predominant effects are those associated with countries. In fact, the gender effects are smaller than any of the residual values.

Insert Table 2 About Here

The total effects resulting from the median polishes of the five content areas ranged from 51.04 for algebra to 59.61 for measurement, indicating that a typical score for a student, without regard to gender or nationality would be in the 50% to 60% range on any of the types of items. Again, after the common effect was removed the largest effects were associated with the countries, but the order of magnitude differed across subject matter areas. These effects reflect not only differences in the curricula and opportunity to learn across the eight countries, but differences in pedagogies and culture as well. For example, Belgium and France focus on fractions, geometry, and algebra at the Population A level (McKnight, Crosswhite, Dossey, Kifer, Swafford, Travers, & Cooney, 1987), thus having similar curricula. Both countries have large positive effects in the median polish of the algebra items. A substantial positive effect is also seen for Belgium on fractions but the effect for France while positive, is small, and on geometry, the effect for Belgium is negative while France's remains small and positive. Similarly, the United States and British Columbia report approximately the same opportunities to learn across the content areas (McKnight et al., 1987), but the effects for British Columbia are consistently positive and those of the United States negative. These results suggest that opportunity to learn is not the predominant contributor to student achievement.

Of primary interest in this study, however, was the effect of gender. The gender effects were very small in each of the median polishes. The largest gender effects were seen on fractions, but these indicated that on the basis of gender alone, one would expect only about 1.5 percentage points difference favoring females. The largest effect favoring males was found with the geometry items, again approximately 1.5 percentage points. With few exceptions, the gender effects were smaller than the residuals.

Examination of the patterns of residuals suggested that the absence of substantial gender effects was the result of interactive effects. That is, the presence of gender differences was a function of not only content but country. For example, the residuals for females in France were consistently negative, indicating slightly lower scores overall and on each of the separate content areas than would be expected after removing the effects of gender and country, whereas the residuals for females in Thailand were consistently positive, indicating higher scores than would be expected given the model. Residuals for both males and females in the United States were all close to zero. In the absence of a substantial gender effect, this indicates that the average performance for both males and females in the United States is approximately the same and is a function of only the common effect and country effect.

While none of the residuals were extremely large in an absolute sense and the majority were close to zero, those greater than |2| tend to stand out. These residuals indicate greater differences in performance between males and females. Only in France and New Zealand are the larger residuals seen to favor males. In the area of fractions, females in New Zealand scored lower than would be anticipated, yet the females in Belgium, British Columbia, and France scored higher in this area. Females in France scored lower in algebra and geometry, but females in Belgium were higher in algebra and those in Thailand were higher in both algebra and geometry. In fact, females in Thailand had residuals greater than 2 in all areas but measurement, and measurement was the only content area with no residuals greater than |2|.

DISCUSSION

The Second International Mathematics Study presents a unique opportunity to investigate gender differences in measures of mathematics achievement. Students of approximately the same ages from countries throughout the world were tested on the same items, allowing comparisons across countries on identical items. From the analysis of these data it appears as if the cultural differences suggested by Leder (1986) are apparent in item type performance as well as on composite measures of quantitative performance.

There were no substantial gender effects in any of the content areas, and the slight effects shown favored girls more often than boys. These findings differ from those cited previously wherein consistencies were found in gender differences across problem type. For example, previous studies found males to perform better than females on problems dealing with proportionality, yet these results show females in Thailand scoring almost five percentage points higher than males on the ratio/proportion/percent items. Furthermore, within no content area were males found to persistently outperform females across countries or vice versa. The absence of these types of effects supports the suppositions of Wood (1976) and Leder (1986) that perhaps pedagogical and cultural factors lend to the presence or absence of gender differences.

Studies examining cross-cultural differences in mathematics performance have identified affective factors that may also contribute to the presence or absence of gender differences within countries. In explaining American kindergarten, first and fifth grade students' low performance relative to Japanese and Chinese students, Stevenson, Lee, & Stigler (1986) cite large differences in the students' lives in school, attitudes and beliefs of their mothers, and students and parental involvement in school work. Also, in a national report on the Second International Mathematics Study (McKnight, et al., 1987) the low achievement evidenced by students in the United States relative to other countries raised concerns about the "nature and quality of the pedagogy demonstrated in the U.S. mathematics classrooms" as well as "the way the content goals are distributed in school mathematics" (p. 9). Future research should investigate how the curricula, pedagogies techniques, and cultural factors interact with gender in impacting quantitative performance.

Footnote

[1] A two-factor (gender by country) ANOVA may appear to be called for to address the research question posed in this study. However, the extremely large sample size ($N > 40{,}000$) resulting from the aggregation of the data by countries produces significant statistics for each effect tested, regardless of how small the effect. While the linear model used in the median polish does not contain a specific interaction component, examination of the residuals can indicate possible interactive effects. If residuals are consistently positive or negative for one gender, the effects of country are the same for each gender; if not, interactive effects are present. Within each country where residuals greater than $|2|$ were observed, independent t-tests were calculated and in each instance were significant with $p < .01$.

REFERENCES

Bart, W.M., Baxter, J., & Frey, S. (1980). The relationship of spatial ability and sex to formal reasoning capabilities. The Journal of Psychology, 104, 191-198.

Brandon, P.R., Newton, B.J., & Hammond, O.W. (1987). Children's mathematics achievement in Hawaii: Sex differences favoring girls. American Educational Research Journal, 24, 437-461.

Crosswhite, F.J., Dossey, J.A., Swafford, J.O., McKnight, C.C., Cooney, T.J., Downs, F.L., Grouws, D.A., & Weinzweig, A.I. (1986). Second international mathematics study detailed report for the United States. Champaign, IL: Stipes Publishing Co.

Fennema, E. (1974). Mathematics learning and the sexes: A review. Journal for Research in Mathematics Education, 5, 126-139.

Fennema, E. (1980). Sex-related differences in mathematics achievement: Where and why. In L.H. Fox, L. Brody, & D. Tobin, (Eds.), Women and the mathematical mystique (pp. 76-93). Baltimore: Johns Hopkins University Press.

Fennema, E., & Carpenter, T.P. (1981). Sex-related differences in mathematics: Results from the National Assessment. Mathematics Teacher, 74, 554-559.

Husen, T. (1967). International study of achievement in mathematics: A comparison of twelve countries, Volume II. Stockholm: Almqvist & Wiksell.

Jarvis, O.T. (1964). Boy-girl ability differences in elementary school arithmetic. School Science and Mathematics, 64, 657-659.

Leder, G.C. (April, 1986). Gender linked differences in mathematics learning: Further explorations. Paper presented at the Research Presession to the National Council of Teachers of Mathematics 64th Annual Meeting, Washington, D.C.

Maier, N.R.F., & Casselman, G.G. (1971). Problem-solving ability as a factor in selection of major in college: Comparison of the processes of idea-getting and making essential distinctions in males and females. Psychological Reports, 28, 503-514.

McKnight, C.C., Crosswhite, F.J., Dossey, J.A., Kifer, E., Swafford, J.O., Travers, K.J., & Cooney, T.J. (1987). The underachieving curriculum: Assessing U.S. school mathematics from an international perspective. Champaign, IL: Stipes Publishing Co.

Meece, J.L., Parsons, J.E., Kaczala, C.M., Goff, S.B., & Futterman, R. (1982). Sex differences in math achievement: Toward a model of academic choice. Psychological Bulletin, 91, 324-348.

National Assessment of Educational Progress (1975). The First National Assessment of Mathematics: An overview (NAEP Report 04-MA-40). Denver: Education Commission of the States.

National Assessment of Educational Progress (1983). The Third National Mathematics Assessment: Results and issues (NAEP Report 13-MA-01). Denver: Education Commission of the States.

Pattison, P., & Grieve, N. (1984). Do spatial skills contribute to sex differences in different types of mathematics problems? Journal of Educational Psychology, 76, 678-689.

Stevenson, H.W., Lee, S., & Stigler, J.W. (1986). Mathematics achievement of Chinese, Japanese, and American children. Science, 14, 693-699.

Tukey, J.W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley.

Velleman, P.F., & Hoaglin, D.C. (1981). Applications, basics and computing of exploratory data analysis. Boston, MA: Duxbury Press.

Walberg, H.J., Harnisch, D.L., & Tsai, S. (1986). Elementary school mathematics productivity in twelve countries. British Educational Research Journal, 12, 237-248.

Wilson, J.W. (1972). Patterns of mathematics achievement in grade 11: Z population (National Longitudinal Study of Mathematical Abilities, No. 17). Stanford: School Mathematics Study Group.

Wood, R. (1976). Sex differences in mathematics attainment at GCE ordinary level. Educational Studies, 2, 141-160.

Table 1. Mean Proportion of Mathematics Items Correct by Gender and Country

	Country							
	U.S.	Belgium	B.C.	Thailand	France	N.Z.	Ontario	Japan
Gender	**Total**							
Female	48.58	62.24	60.25	48.13	55.30	45.20	52.55	68.09
Male	48.13	59.00	58.32	43.55	58.79	47.08	53.64	69.33
	Fractions							
Female	52.65	68.19	67.85	46.44	56.38	36.14	50.16	82.86
Male	52.18	61.28	62.18	38.64	58.74	38.74	51.52	80.39
	Ratio/proportions/percent							
Female	49.20	59.42	62.12	58.16	40.40	45.67	57.90	37.24
Male	48.92	59.03	60.68	53.54	43.75	48.99	57.85	40.53
	Algebra							
Female	45.16	66.89	57.07	37.14	59.06	41.35	44.06	69.47
Male	44.56	61.58	56.82	33.02	64.14	42.45	46.22	72.37
	Geometry							
Female	45.74	48.75	53.57	48.39	51.14	50.92	52.41	76.13
Male	46.56	49.76	53.61	45.45	56.51	54.18	55.40	77.96
	Measurement							
Female	51.05	66.63	61.55	52.98	64.86	51.65	59.23	74.42
Male	49.23	63.26	59.25	49.51	66.45	51.84	58.43	75.52

Table 2. Median Polish of Proportion of Mathematics Items Correct

	U.S.	Belgium	B.C.	Thailand	France	N.Z.	Ontario	Japan	Gender Effect	Common
					All Items					
Female	.39	1.78	1.13	2.45	-1.59	-.78	-.39	-.46	-.16	
Male	-.39	-1.78	-1.13	-2.45	1.59	.78	.39	.46	.16	
Country Effect	-7.28	4.98	3.65	-9.80	1.41	-9.50	-2.55	13.07		55.64
					Fractions					
Female	-.50	2.72	2.10	3.16	-1.92	-2.04	-1.42	.50	.74	
Male	.50	-2.72	-2.10	-3.16	1.92	2.04	1.42	-.50	-.74	
Country Effect	-3.50	8.82	9.09	-13.38	1.64	-18.48	-5.08	25.70		55.92
					Ratio/proportion/percent					
Female	.06	.11	.64	2.23	-1.76	-1.74	-.06	-1.73	.08	
Male	-.06	-.11	-.64	-2.23	1.76	1.74	.06	1.73	-.08	
Country Effect	-3.40	6.77	8.94	3.40	-10.38	-5.13	5.42	-13.57		52.46
					Algebra					
Female	.51	2.87	.34	2.27	-2.33	-.34	-.87	-1.24	-.21	
Male	-.51	-2.87	-.34	-2.27	2.33	.34	.87	1.24	.21	
Country Effect	-6.18	13.19	5.90	-15.96	10.56	-9.14	-5.90	19.88		51.04
					Geometry					
Female	.30	.21	.69	2.18	-2.03	-.92	-.79	-.21	-.71	
Male	-.30	-.21	-.69	-2.18	2.03	.92	.79	.21	.71	
Country Effect	-6.92	-3.82	.52	-6.15	.81	-.52	.84	23.97		53.07
					Measurement					
Female	.26	1.03	.49	1.08	-1.45	-.75	-.26	-1.20	.65	
Male	-.26	-1.03	-.49	-1.08	1.45	.75	.26	1.20	-.65	
Country Effect	-9.48	5.33	.78	-8.37	6.04	-7.87	-.78	15.35		59.61