Gender Differences in Large Scale, Quantitative Assessments of Mathematics and Science Achievement

Darren Lauzon

Family and Labour Studies Division

Statistics Canada

Paper Prepared for the Statistics Canada-John Deutsch Institute--WRNET Conference on Empirical Issues in Canadian Education, Ottawa, Nov. 23-24, 2001

Introduction[1]

A vast literature concerning gender differences favouring males in mathematics and science has arisen over the past several decades in academic, practitioner and public policy fields. The enquiry covers differences in cognitive ability, school achievement, achievement in standardized tests, attitudes, motivation, participation and course-taking. For a time, small but persistent achievement gaps in standardized tests, primarily in the U.S., highlighted the need to address teaching practices, attitudes, learning behaviours and other factors associated with the under-performance of girls. In contrast to this, recent evidence reveals that gaps in math and science have narrowed substantially, or perhaps disappeared altogether, suggesting that these efforts have paid off. When performance in math and science is compared to that in language, the difference, if any, between boys and girls seems to lie in the consistently better performance of girls in tests of reading and writing skills.

When research on gender gaps in math and science is disseminated and discussed in the popular press, subtle and not-so-subtle variations in measurement, design and reporting are lost and the result is a bewildering constellation of facts and hypotheses through which non-

[1] The term "gender" difference will be used throughout and its meaning will be synonymous with "sex" difference. That is, gender here refers simply to "boy" and "girl", "male" and "female" in the sense these are generally understood.

experts must navigate. Are there in fact differences? Differences in what, and what are the consequences of such differences?

This chapter focuses on observed gaps, if any, in large-scale, quantitative assessments of math and science. It briefly discusses the broad literature on gender gaps in math and science and the features of assessments that affect the ability to make general conclusions about gender differences based on just a few assessments. While the presence (or not) of achievement gaps has been extensively explored, the consequences of such gaps, particularly those observed in the 70s, 80s and early 90s, for differences among men and women in adult life could remain fertile ground for investigation. The last section of this chapter briefly discusses some of the possible impacts of differences in science and math learning on adult life.

What we know about differences

In the broadest terms, some basic facts have been more or less established in the research of the last two decades on gender differences in math and science. Gender differences favouring boys in mathematics tend not to appear until high school—in earlier grades, differences are either non-existent or favour girls (Hyde et. al. 1990, Randhawa, 1991; Han & Hoover, 1994; Ma, 1999a). They tend to appear in standardized tests rather than school grades[2] and tend to be highest among the best performing students (Becker & Forsyth, 1990; Han & Hoover, 1994; Hedges & Nowell, 1995; Fan, 1995; Lawson et. al., 1999;Lauzon, 1999). The gaps tend also to be declining with time (Friedman, 1989; Lawson et. al., 1999, Ma, 1996a, Stanley & Stumpf, 1997 McClure 1998) though this may be less true of advanced mathematics achievement and upper-level course-taking patterns (Benbow, 1988, Johnson, 2000). Observed gaps tend to be small relative to differences between other population groups—i.e., gaps are larger within genders than between genders (Lawson et. al.. 1999).

With respect to science achievement there has beeen relatively less attention paid to gender differences than in mathematics (Zhang, 1999). Recent analyses suggest that observed gaps are more consistent over time, tend to be strongest in physics and earth sciences than in biology and "life science" or general science (Steinkamp & Maehr, 1983, 1984; Becker, 1989; Lee & Burkam, 1996). The gap has been more persistent than in mathematics. Bruschi and Anderson (1994) found that the early male advantage in physical sciences and earth and space

[2] Bridgeman and Lewis (1996) indicate that when SAT math scores and high school grades are combined, gender differences in calculus and pre-calculus are miniscule.

science became more substantial with age. Females were favoured in the science of nature in all age groups.

In Canada, Large-scale national and provincial assessments have been more limited than that in the U.S. However, recent test evidence seems to confirm these general findings. The 1995 Third International Mathematics and Science Study (TIMSS) showed virtually no difference between boys and girls in grades 3 and 4, only a slight difference in grades 7 and 8 (in science more than mathematics). More substantial differences occurred in the final year of secondary school, particularly in advanced mathematics and physics. The TIMSS Repeat (1999), concentrating on grade 8 students, showed a gap favouring boys in science but parity in mathematics performance.

A number of national assessments in math, science, reading and writing have been administered through the School Achievement Indicators Program (SAIP). The 1997 math assessment showed significantly more 13- and 16-year-old boys performing at higher levels in math problem solving. The 1999 science assessment showed significantly more 13- and 16-year old girls performing at higher levels in practical tasks while there were no significant differences in the written assessment.

In Ontario, the Education Quality and Accountability Office (EQAO) has begun administering assessments as part of that province's move toward greater accountability in educational outcomes. The 1999-2000 assessment showed more girls performing at the upper levels in grade 3 and grade 6, the gap being lower among grade 6 students.

In British Columbia, the Foundation Skills Assessment (FSA) examines whether BC students meet or exceed expectations or are not yet within expected standards in reading, writing and numeracy. The 1999 FSA showed no gap between grade 4, 7 and 10 boys and girls in meeting expectations in numeracy, though the public report shows slightly higher portions of boys exceeding expectations in all grades.

Most recently, the Organization for Economic Cooperation and Development (OECD) through its Programme for International Student Assessment (PISA), released math, science and reading test results for 15-year-olds. No gender differences in average math and science performance were observed in any Canadian provinces.

Table 1 summarizes the results from the recent Canadian assessments. The data suggest that differences favouring boys in mathematics have closed. There is mixed evidence that boys perform better than girls in the sciences.

INSERT TABLE 1 HERE

Explaining the gap

Where they have been observed, a considerable effort has been expended in a variety of studies to explain math and science gender gaps. Biological explanations have been advanced suggesting that there are innate differences between the sexes that imply differential performance on assessments (Benbow & Stanley, 1980, 1988) the recent narrowing of gender gaps suggests that biological differences are not important or can be easily overcome by intervention. Recent meta-analysis supports this view (Luckenbill, 1995).

Socialization or environmental factors receive more attention in the literature. Oakes (1990) provides a useful summary of the individual, school and societal factors thought to explain differences in the way boys and girls participate and achieve in mathematics and science. Individual factors can be loosely grouped into "cognitive" and "affective" classes. With respect to cognitive differences, much has been written about gender differences in spatial ability in particular. Though there is debate about the extent of these cognitive differences and whether they translate into differences on standardized achievement tests, there is evidence that they can be overcome by training.

Affective factors involve the attitude, motivation and self-perception of students. Affective factors often studied include differences between the genders in the relative interest in “people” and “things”, enjoyment of math, perceived utility of math, stereotyping of math as the domain of males, and confidence in one’s own ability. While the existence of cognitive difference is inconclusive, it does seem clear that girls exhibit more negative attitudes toward math and science. However, there is little conclusive evidence that these negative attitudes *cause* lower achievement for girls (Oakes, 1990). The role of attitudes is not fully understood and is likely part of a complex process of decision making that both influences and is influenced by achievement.

School factors might also underlie gender differences and include such things as access to resources, individual guidance and encouragement, the presence of role models, especially proximal ones in the school environment, curriculum tracking, teacher expectations, teaching practices and teacher-student interactions. School factors do represent the most direct policy levers with which to address imbalance in school performance. Karp and Shakesshaft (1997) suggest that gender specific student-teacher interaction could influence observed differences in SAT math scores controlling for differences in course-taking.

Lastly, societal factors that can lead to gender differences in math achievement have been widely studied. The disadvantages associated with low socioeconomic status (SES) lead to lower than average test scores for low SES students. In widely available U.S.

datasets, parents' education, family income and family possessions are usually used as measures of SES. There is some evidence that SES impacts women more than men. Ware et al. (1985) shows that women from more privileged backgrounds are more likely to choose science major and among those in science majors, women are more likely than men to have mothers employed in prestigious occupations (Ware and Lee, 1985). Sax (1996) found that women with science, math or engineering undergraduate degrees whose mothers were research scientists or college teachers were more likely to pursue graduate work.

Related to parent's education and occupation is the notion of parental involvement in education. Ware & Lee (1985) found that science majors had parents who were involved in their high school activities. Recently, Ma (1999b) found that parental involvement was an important factor in math persistence, but Muller (1998) found that parental involvement *in school activities* was similar for boys and girls and did not underlie differences in achievement.

When considering differences in standardized tests, some suggest that differences are related more to test-taking strategy than to actual differences in knowledge (e,g, Gallagher et. al. 2000 and Gallagher and DeLisi, 1994). Strategy differences might fall into the category of legitimate differences if girls approach test items in ways that do not allow them to reveal their knowledge, i.e., they know the right answer but can't show it. The remedy might be simply to balance the test with items that elicit more natural problem solving strategies for girls, or items that are equally natural for both genders to approach. Other recent evidence, however, suggests that strategy differences are rooted in differences in prior knowledge (Byrnes and Takahira (1993, 1994), or cultural differences (Byrnes, Hong and Xing (1997)). Strategy differences, it would seem, simply reflect deeper processes of the sort discussed by Oakes.

Of course there is a great deal of interaction between individual cognitive and affective factors, school experiences and societal factors and a major criticism of past literature is that few studies attempt to integrate these many factors in any theoretical way (Middleton and Spanias, 1999 and Oakes, 1990). This is certainly a desirable goal when good data exist. However, what Oakes calls predominantly "correlative" studies can provide valuable insight if interpreted carefully. Regression-based analysis of cross-sectional achievement test samples is especially useful for providing insight into how sample composition impacts summary statistics such as mean scores on individual tests. Lauzon (2001) showed that about 21 percent of the observed gap in the Science and Math Literacy Test (of the 1995 TIMSS) could be explained by differences between boys and girls in the highest level of math and science courses taken.

Gaps in large scale assessments

Differences in What?

Focusing, now, on large scale, quantitative tests there are several dimensions along which assessments can differ, apart from obvious differences in the target populations of assessments (e.g., jurisdiction, grade level, or academic background). Differences along the following dimensions directly affect the comparability of assessment results.

1. content,
2. item format,
3. quantification of results (e.g., standardizing scores), and
4. statistics used to report achievement results

Content is chief among these and is perhaps the most significant factor determining the comparability of assessment results. Content differs across assessments along a continuum from the general to the specific. General content refers to the broad knowledge and skills the assessment is meant to evaluate, e.g. the core knowledge expected of all graduating secondary school students or expected proficiency in pre-college calculus). Specific content refers to the actual knowledge and skills tested in individual items, e.g., the manipulation of fractions. Along this continuum, content can be more or less curriculum-based, i.e., reflect more or less the content students were actually taught in school.[3] Differences between boys and girls in the summary results of a test can, in part, be driven by differences in particular content areas (Becker, 1990). For example, much has been made of the apparent disadvantage of girls in "spatial rotation" and geometry content relative to their performance in computation, an area where girls often outperform boys.

Closely related to content is the nature of test items used in assessments. For a given content, test items can vary by format e.g. multiple choice and open-ended or so-called structured response. Item choice and format is the subject of specialist psychometric literature, but it is important to note that item format has been shown to underlie gender differences. For example, Hamilton (1994) found evidence that item features could explain, in part, gender differences in the US National Education Longitudinal Study (NELS). Bielinski & Davison (1998) found some evidence that "easy" test items are easier for females and "harder" test items are harder for females. Zhang & Manon (2000) found that boys did better on multiple choice and girls did better on structured

[3] A distinction can also be made between official content schools are required to teach and content actually delivered (what teachers actually taught). See Mitchell (1999) for a discussion of the literature on curriculum alignment. This distinction was recognized in the design of the 1995 TIMSS.

response.[4] Others have not found such evidence (O'Neil & Brown (1998), DeMars (1998)).[5]

Gender differences in particular items can help explain different results across different achievement tests.[6] When examining any given test that reveals a gender gap in performance, the issue remains that girls perform less well on average on at least some if not all items. An understanding of the process that leads to those differences makes it possible to assess whether there really is a gap in learning or ability that should be remedied or whether the differences reflect "legitimate" learning, performance or behavioral differences in the genders that should be accounted for in assessment design.

Quantification of results is another substantive psychometric issue, details of which are outside the scope of this chapter.[7] However, when assessments employ different methods of producing standardized

---

[4] This observation is particularly important since there has been a trend in assessment design toward the more qualitatively richer open-ended questions on which girls often do better.

[5] Differential reaction to standardized test times has been observed with Canadian students as well (Erickson & Farkas, 1991).

[6] For example, Duffy, Gunther and Walters (1997) examined a set of high achieving Canadian 12 year olds and found gender differences on the Canadian Test of Basic Skills but not on the GAUSS assessment.

[7] For a valuable technical discussion of different assessment types see Bartley & Lawson. 1999. Wolfe, et. al. (1999) provide an excellent discussion of issues of quantification in assessments.

or scale scores to summarize the test results, differences in estimated gender differences can result. This was apparent in the Canadian results from the 1995 Third International Mathematics and Science Study (TIMSS) where gender gaps were highest when measured with plausible value scores and lower using Rausch scores. (Fig. 1.). Such differences may arise when items on which boys and girls perform differently are weighted differently in the summary scale. Standardized scores can also *mask* differences between genders that are linked to specific item types (Kupermintz et. al., 1994).

INSERT FIGRUE 1 HERE

Lastly, the means by which results are presented in public reports and used in research studies varies across assessments. Differences in means of standardized scores are the typical method, particularly in U.S. Studies. Recent Canadian assessments report the percentage of students performing at particular levels of achievement.

Such summary statistics are useful but do not give a complete picture of relative performance. An insignificant difference in means might mask considerable differences at the upper and lower quantiles of the distribution. This is especially the case if boys have greater variability in scores as has been documented in many studies. A related issue is the relative inequality in tests cores between boys and girls, again, some research has suggested that boys test scores are more varied—and hence there is greater inequality in boys' achievement than girls.

Zhang & Manon, 2000 looked at two standardized tests over two years for grades 3, 5, 8 and 10. They do not observe differences in mean performance on the whole, but do observe gaps among the highest and lower 10 percent of students.[8] Han & Hoover (1994) note that male performance on standardized tests in the U.S. between 1963 and 1992 tended to be more variable and that differences favoured females at the low end of test results and favoured males at the high end. Fan et. al. (1997) found similar patterns for the National Education Longitudinal Study (NELS) as did Lauzon (1999) in the 1995 Canadian TIMSS. Feingold (1992) argues this point directly—differences in means are not enough to give the complete picture of differences in ability distributions.

Differences in the proportions of boys and girls achieving at certain levels might also be misleading. While a slightly greater percentage of girls might perform at or above a given level than boys, considering all students who perform at or above that level, the average performance may favour one or the other gender. This is particularly true

[8] They examine the Delaware Student Testing Program and Stanford Achievement Test Series 9. Unlike other studies, they report "no clear patterns of test performance between females and males have been found in the content and cognitive categories of mathematics".

when the levels encompass a wide range of potential performance, such as “meeting expectations” in the BC FSA.

Given the myriad of factors that can affect observed differences between the genders in standardized tests, it is not possible to get a general picture of the relative performance of boys and girls from just one assessment. Instead, a large number of assessments in similar content areas and of similar design are needed for a detailed picture to emerge.

Beyond the test results

Though the study of achievement gaps in math and science has been extensive, much less attention has been paid to their longer-term implications, particularly in adult years. Most widely studied are the implications of gender differences in elementary and high school achievement for post-secondary aspirations (Marion and Coladarci, 1993, Sax, 1994 are some of many examples). Turner and Bowen (1999) found that there is “a widening divide between the life sciences and the math/physical science fields in their relative attractiveness to men and women.” Differences in SAT scores account for only a small portion of that gap. Baker (1998) found that small gaps in post-secondary science and engineering completion rates remain after controlling for ability (measured in many ways including Graduate Record Exam (GRE) scores) and that these have been declining with time.

Others point to the fact that the degree to which standardized admissions tests are good predictors of performance in post-secondary programs, differences in performance on these assessments translates into differences in post-secondary performance and graduation rates. The evidence on this seems incomplete. Ramsbottom-Lucier et. al. (1995) found dramatic differences in preadmission Grade Point Averages and standardized entrance tests but smaller difference in program performance in medical school. Spencer (1996) found that while performance on the SAT math component predicted college chemistry grades, there were no obvious gender differences in post-secondary general chemistry GPAs. Odell & Shumacher (1998) found business school grades were better predicted by attitude differences than SAT scores.

From a broader social and economic policy perspective, differential performance of boys and girls in science and math, at least in the past 20 years, could be expected to translate into differences in attitudes and social engagement among adults. Do women participate less in public discussion of important scientific issues such as genetics, the environment and public health? It is not clear how a relative lack of science learning or interest should translate into differences in attitudes and participation. On the one had, less understanding could lead to excessive caution and risk aversion. On the other, lack of awareness could lead to overly optimistic assessments of personal or environmental risk. Bimber (2000) has observed a gender gap (favouring men) in Internet access and participation. Bord & O'Connor (1997) find that the observation that women are more cautions regarding environmental risk is dependent upon the perceived vulnerability to the risk, i.e., risk to personal health.

Documented differences in enrolment patterns could be expected to translate into differential choices among careers for women. There are already numerous studies that documented the under-representation of women in science and engineering fields, though this too has decreased. Brown & Corcoran (1996) find little consistent evidence that the male-female wage gap is due to differences in the pursuit of "traditionally male courses". Abbot et. al. (1999) suggest that field of study differences in Canada may still underlie gender differences in earnings. Baker and Fortin (2000) indicate that occupational differences among men and women do not contribute as much to the observed gap in wages as comparable differences in the U.S..

Lastly, if differences in math and science learning appear in secondary school, are they sustained in adult learning? In particular, are there differences between men and women in access to and the success in on-the-job training programs? Early U.S. evidence from the 1980s suggests some ambiguity as to whether women receive less employer sponsored training then men. Knoke & Ishio (1998) find a persistent gap in employer-provided training due to differences in gender in occupation and industry. There has been much less work on the relative performance of men and women in on-the-job and employer-provided training, much less the impact of differential course-taking or science and math abilities on training incidence and success

Concluding Remarks

This chapter addressed gender differences in mathematics and science achievement tests. Though such gaps were regularly observed in the 1980s and early 90s, particularly in U.S. data, recent evidence in both the U.S. and Canada suggests that boys and girls perform about the same. Gaps in reading and writing skills favouring girls, however, continue to show up in large-scale assessments, most recently the OECD's PISA assessment, now overshadowing gender differences in math and science as the predominant gender disparity in educational achievement.

The conclusion one is tempted to draw from this is that gender differences in math and science achievement ar e no longer important. The gaps have been consistently small when observed, and have disappeared in recent assessments. When they are observed, the factors associated with them seem well understood, and interventions based on this knowledge appear to have helped reduce or eliminate the difference.

This conclusion is subject to a few important caveats. First, unlike the male-female wage gap, achievement gaps in large scale, quantitative assessments vary in part according to variation in the assessments on which they are based. It is difficult to get a complete picture of the magnitude of achievement gaps and the content areas in which they occur because the few available assessments in Canada are not directly comparable. Second, while the present generation of students may no longer experience differences in science and math achievement, a previous generation has and the implications of those differences for adult life may not be well understood. It is safe to conclude, however, that subject to these caveats, gender differences in math and science may now present a "monitoring" issue—something for educators and researchers to "keep an eye on, and perhaps serve as a useful knowledge base for addressing gender gaps in reading and writing.

References

Baker, J. (1998) “Gender, Race and Ph.D. Completion in Natural Science and Engineering” *Economics of Education Review* 17 (2) 179-88

Bartley, A. W. & Lawson, A (1999) “Varieties of Assessment: Issues of Validity and Reliability” Research Series, Number 2, (Toronto, EQAO)

Becker, B. J. 1989. “Gender and science achievement: A reanalysis of studies from two metaanalyses.” *Journal of Research in Science Teaching* 26, 141-169.

Becker, B. J. (1990), “Item Characteristics and Gender Differences on the SAT-M for Mathematically Able Youths”, *American Educational Research Journal* 27, 65-87

Becker, D. F. & Forsyth, R. A. (1990), “Gender Differences in Academic Achievement in Grades 3 through 12: A Longitudinal Analysis” Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990)..

Benbow, C. P. (1988), “Sex Differences in Mathematical Reasoning Ability in Intellectually Talented Pre-Adolescents: Their Nature, Effects and Possible Causes”, *Behavioral and Brain Sciences* 11, 169-183

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. (1983). “Sex differences in mathematical reasoning ability: More facts.” *Science* (222) 1029-1031.

Bielinski, J & Davison, M. L. (1998) “Gender Differences by Item Difficulty Interactions in Multiple Choice Mathematics Items”, *American Educational Research Journal* 33, 455-76.

Bimber, B. (2000) “Measuring the Gender Gap on the Internet” *Social Science Quarterly* 81 (3) 868-76

Bord, R. J. & O'Connor, R. E. (1997) “The Gender Gap in Environmental Attitudes: The Case of Perceived Vulnerability to Risk” *Social Science Quarterly* 78 (4) 830-40

Bridgeman, B, & Lewis, C. (1996) “Gender Differences in College Mathematics Grades and SAT-M Scores: A Reanalysis of Wainer and Steinberg”, *Journal of Educational Measurement* 33 (3) 257-70

Brown, C. & Corcoran, M. (1997) “Sex-Based Differences in School Content and the Male-Female Wage Gap” *Journal of Labor Economics* 15 (3), Part 1, 431-65..

Bruschi, B. A. & Anderson, B. T. “Gender and Ethnic Differences in Science Achievement of Nine-, Thirteen-, and Seventeen-Year-Old Students”. Paper presented at the Annual Meeting of the Eastern Educational Research Association (Sarasota, FL, February)

Byrnes, J. P. & Takahira, S. (1994) “Why Some Students Perform Well and Others Perform Poorly on SAT Math Items”, *Contemporary Educational Psychology* 19 (11, 63-78

Byrnes, J. P. & Takahira, S. (1993) “Explaining Gender Differences on SAT Math Items”, *Developmental Psychology* 29, 805-810

Byrnes, J. P.; Hong, L.; & Xing, S. (1997). “Gender Differences on the Math Subtest of the Scholastic Aptitude Test May Be Culture-Specific”, *Educational Studies in Mathematics* 34, 49-66

Catsambis, S. (1994) “The Path to Math: Gender and Racial-Ethnic Differences in Mathematics Participation from Middle School to High School” *Sociology of Education* 67 (3) 199-215

DeMars, C. E. “Gender Differences in Mathematics and Science on a High School Proficiency Exam: The Role of Response Format”, *Applied Measurement In Education* 11 (3) 279-99

Duffy, J,; Gunther, G, and Walters, L, (1997). “Gender and Mathematical Problem Solving”, *Sex Roles: A Journal of Research* 37 (7-8) 477-94

Erickson, G, and Farkas, S. (1991) “Prior Experience and Gender Differences in Science Achievement.” *Alberta Journal of Educational Research* 37 (3) 225-39

Fan, X et. al. (1997) “Gender Differences in Mathematics Achievement: Findings from the National Education Longitudinal Study of 1988.”, *Journal of Experimental Education* 65 (3) 229-42

Feingold, Alan. (1992) “The Greater Male Variability Controversy: Science versus Politics”. *Review of Educational Research* 62 (1) 89-90

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 (2) 185-213

Gallagher, A. M. et. al.. (2000) “Gender Differences in Advanced Mathematics Problem Solving”, *Journal of Experimental Child Psychology* 75, 165-90

Gallagher, A. M & Delisi, R. (1994) “Gender Differences in Scholastic Aptitude Test—Mathematics Problem Solving Among High Ability Students”, *Journal of Educational Psychology* 86, 204-11

Hamilton, L. et. al. (1994) “Gender Differences on High School Science Achievement Tests: Do Format and Content Matter”? *Educational Evaluation and Policy Analysis* 20 (3) 179-95

Han, L. & Hoover, H.D. (1994) “Gender Differences in Achievement Test Scores” Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, April 5-7, 1994)

Hedges, L. V. & Nowell, A. “Sex Differences in Mental Test Scores, Variability, and Numbers of High Scoring Individuals”, *Science*, 269, 41-45

Hoffer, T. et. al. “Social Background Differences in High School Mathematics and Science Coursetaking and Achievement.” *Statistics in Brief* Report NCES-95-206 (Washinton DC. National Centre for Education Statistics).

Hyde, J.; Fennema, E. & Lamon, S. (1990). “Gender Differences in Mathematics Performance: A Meta-Analysis”, *Psychological Bulletin* 107 (2), 139-155

Johnson, R. M. (2000), “Gender Differences in Mathematics Performance: Walberg's Educational Productivity Model and the NELS:88 Database”, Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).

Kupermintz, H. et. al.(1994) “Enhancing the Validity and Usefulness of Large-Scale Educational Assessments: I. NELS:88 Mathematics Achievement” Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8)

Lawson, A.; Penfield, R. & Naggy, P. (1999) “Relating Attitudes, Gender and Student Achievement in Grades 3 and 6” Research Series, Number 1, (Toronto, EQAO)

Lauzon, D. (1999) “Gender Differences in Science and Math Achievement in the Final Year of Secondary School: Evidence From the 1995 TIMSS” Unpublished working paper, Centre for Education Statistics, Statistics Canada Ottawa

Lauzon, D (2001) “How Much Do Gender Differences in Math and Science Participation Account for Gender Differences In High School Mathematics Achievement?” Unpublished working paper, Centre for Education Statistics, Statistics Canada Ottawa

Luckenbill, L. M. (1995) “Biological Superiority in Math: Calvin or Susie? Spotlight: Gender Differences”. *Montessori Life* 7 (4) 28-32

Ma, X. (1999a) “Gender Differences in Growth in Mathematics Skills During Secondary Grades: A Growth Model Analysis”, *Alberta Journal of Educational Research* 45, 448-66

Ma, X. (1999b) “Dropping Out of Advanced Mathematics: The Effects of Parental Involvement”, *Teachers College Record* 101, 60-81

Madiggan, T. (1997) “Science Proficiency and Course Taking in High School: The Relationship of Science Course-Taking Patterns to Increases in Science Proficiency between 8th and 12th Grades” Research Paper NCES-97-838 (National Centre Education Statistics: Washington DC)

Marion, S. F. & Coladarci, T. (1993) “Gender Differences in Science Course-Taking Patterns among College Undergraduates”, Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April).

Marhs, H. W. & Yeung, A. S. (1998) “Longitudinal Structural Equation Models of Academic Self-Concept and Achievement: Gender Differences in the Development of Math and English Constructs”, *American Educational Research Journal* 35 (4) 705-38

McLure, G. T. (1998) “High School Mathematics Course-Taking and achievement Among College Bound Students: 1987-1996”, *NASSP Bulletin* 82, 110-118

Middleton, J. A. & Spanias, P. A. (1999) “Motivation for Achievement in Mathematics: Findings, Generalizations, and Criticisms of the Research.” *Journal for Research in Mathematics Education*; 30 65-88

Mitchell, F. M. (1999) “All Students Can Learn: Effects of Curriculum Alignment on the Mathematics Achievement of Third-Grade Students”, Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Quebec, Canada, April 19-23,

Muller, C. (1998) “Gender Differences in Parental Involvement and Adolescents’ Mathematics Achievement” *Sociology of Education* 71, 336-56

Navarro, C. (1989) “Why Do Women Have Lower Average SAT Math Scores Than Men?” Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989

Oakes, J. (1990). “Opportunities, Achievement, and Choice: Women and Minority Students in Science and Mathematics.” *Review of Research in Education* 16 153-222.

O’Neil, H. S. jr. & Brown, R. S. (1998), “Differential Effects of Question Formats in Math Assessments on Metacognition and Affect” *Applied Measurement in Education* 11, 331-51

Odell, P. M. & Schumacher, P. (1998), “Attitudes toward Mathematics and Predictors of College Mathematics Grades: Gender Differences in a 4-Year Business College”, *Journal of Education for Business*; 74 (1) 34-38

Pallas A.M. & Alexander, K.L. (1983) ”Sex Differences in Quantitative SAT performance: New Evidence on the Differential Coursework Hypothesis” *American Educational Research Journal* 20, 165-182

Ramsbottom-Lucier, M. et. al. (1995), “Age and Gender Differences in Students' Preadmission Qualifications and Medical School Performances”, . *Academic Medicine* 70 (3) 236-39

Randhawa, B. R. S (1991) “Gender Differences in Academic Achievement: A Closer Look At Mathematics” *Alberta Journal of Educational Research* 38 (3), 241-257

Sax, L. J. (1994), “Mathematical Self-Concept: How College Reinforces the Gender Gap”, *Research in Higher Education* 55 (2) 141-66

Sax, L. J.. (1996) “The Impact of College on Post-College Commitment to Science Careers: Gender Differences in a Nine-Year Follow-Up of College Freshmen”, Paper presented at the Annual Meeting of the Association for the Study of Higher Education (Memphis, TN, November)

Sharp, L. M. (1989) “Gender Gap: Looking at Micro Level Data”, Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989)

Stanley, J. C. and Stumpf, H. (1997), “Gender Differences, Especially on Fifty College Board Achievement Tests” Paper presented at the Annual Meeting of the Eastern Psychological Association (Washington, DC, April).

Steinkamp, M. W., & Maehr, M. L. (1984) "Gender differences in motivational orientations toward achievement in school science: A quantitative synthesis." *American Educational Research Journal* 21 39-59.

Turner, S. E. & Bowen, W. G. (1999) "Choice of Major: The Changing (Unchanging) Gender Gap" *Industrial and Labor Relations Review* 52 (2) 289-313.

Ware, N. C. & Lee, Y. (1985) "Predictors of Science Major Choice in a National Sample of Male and Female College Students" Unpublished Manuscript, Radcliffe College, Cambrige MA.

Ware, N.C., Steckler N.A. & Lesserman, J. (1985) "Undergraduate Women: Who Chooses a Science Major?", *Journal of Higher Education* 56 73-84

Wolfe, R.; Wiley, D. & Traub, R (1999) "Psychometric Perspectives for EQAO: Generalizability Theory and Applications" Research Series, Number 3, (Toronto, EQAO)

Zhang, L. & Manon, J. (2000), "Gender and Achievement--Understanding Gender Differences and Similarities in Mathematics Assessment" Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).

Zhang, Y. (1999) "Gender Differences in Mathematics and Science Achievement: Role of Engagement in Out-of-School Activities and Experimentation in Science Lessons of Eighth Graders" Human Resources Development Canada, Applied Research Branch, Research Paper Series R-99-18Ea