The Alberta Journal of Educational Research *Vol. XLVIII, No. 4, Winter 2002, 275-286*

Gregory A. Pope

Carolyn Wentzel
Alberta Learning, Edmonton

and

Ron Cammaert
Riverside Publishing Company, Itasca, IL

Relationships Between Gender and Alberta Diploma Scores

The purpose of this study was to investigate relationships between gender and Alberta diploma examination and school-awarded scores. Weak relationships between gender and scores were found both on Alberta diploma examination scores and on school-awarded scores. These relationships tended to be relatively greater in magnitude for school-awarded scores than for diploma examination scores, with girls generally attaining higher school-awarded scores than boys. Effect sizes of all correlations between scores and gender were in the small range. R^2 values showed that the overall statistical effect of gender on students' school-awarded scores was greater than the effect on students' diploma examination scores. The greatest statistical effect for gender was on the difference between school-awarded scores and diploma examination scores, with girls having larger differences than boys.

Le but de cette recherche est d'étudier le rapport entre le sexe et les notes aux examens du ministère menant à l'obtention du diplôme du secondaire en Alberta et celles découlant des examens gérés par les écoles. La corrélation entre le sexe et les notes était faible pour les deux types d'examens. Les résultats ont révélé une tendance selon laquelle la corrélation était plus importante pour les notes découlant des examens gérés par l'école, où les notes des filles étaient généralement plus élevées que celles des garçons. L'ampleur de l'effet pour toutes les corrélations entre les notes et le sexe était limitée. Les valeurs R^2 ont indiqué que l'effet statistique global du sexe sur la performance des élèves était plus important pour les notes découlant d'examens administrés par l'école que pour celles des examens du ministère. L'effet statistique le plus important sur le sexe consistait en la différence entre les notes aux examens administrés par l'école et les notes aux examens gérés par le ministre. Les différences les plus importantes se retrouvaient chez les filles.

Many years of research into the area of relationships between gender and academic performance have uncovered some links. To elucidate these links a broad spectrum of theories have been put forth. Biological factors have been proposed to explain the phenomenon of boys outperforming girls in areas such as mathematics (Halpern, 1997). At the school level social factors such as

Gregory A. Pope is a psychcmetrician with the Learner Assessment Branch. He specializes in applied psychometrics and quantitative research methods.
Carolyn Wentzel is a psychometrician with the Learner Assessment Branch. She specializes in applied statistics and quantitative research methods.
Ronald A. Cammaert is Program Director of Contract Management for Riverside Publishing and former Assistant Director of the Analytic Branch of Alberta Learning. He is currently responsible for managing the assessment program for the State of Washington.

self-esteem (Dukes & Martinez, 1994), patterns of classroom interaction (Lindow, Marrett, & Wilkinson, 1985), sex role socialization (Thorne, 1993), and gender role stereotype conformity (Bouchard & St-Amant, 2000) have been explored as variables that could account for relationships between gender and academic performance.

Relationships between gender and student performance on standardized tests have received a great deal of attention in the literature. Findings are inconsistent, however. For example, Han and Hoover (1994), investigated gender differences in performance on three American standardized tests. They reported that average differences in most subject areas were small and similar in magnitude and direction in nearly 30 years of testing. Girls tended to score somewhat higher at the upper grades in verbal tests, whereas boys tended to score higher in mathematics computation. Smith, Young, Bae, Choy, and Alsalam (1997) found that a substantial gender gap in the mathematics and sciences exists among 17-year-old youth in the United States, as measured by the National Assessment of Educational Progress (NAEP). Noble, Davenport, Schiel, and Pommerich (1999), examined high school factors that could explain differences in American College Test (ACT) scores. Academic factors such as high school grades and high school attendance accounted for 34% to 59% of the variation in ACT scores; gender was found to account for only 1% to 2% of the variability in ACT scores. Willingham, Cole, Lewis, and Leung (1997) found that at the grade 12 level, across 74 different standardized tests in 15 categories, gender differences accounted for no more than 1% of the variation in test scores.

The results of large-scale assessments administered to Canadian students at both the national and provincial levels have also shown some relationships between gender and standardized test performance. The 1997 School Achievement Indicators Program (SAIP) mathematics assessment (Council of Ministers of Education, Canada, 1997) showed several statistically significant gender differences. For the 13-year-old Canadian students, no statistically significant differences were found at any of the five levels on mathematics content. However, the 16-year-old students showed statistically significant gender differences on mathematics content in favor of girls at levels 1 and 2, and significant differences in favor of boys at levels four and five. For the mathematics problem-solving section, 13-year-olds showed significant gender differences at levels 2 and 4, with girls outperforming boys at level 2, and boys outperforming girls at level 4. The 16-year-olds in the mathematics problem-solving section showed significant differences at levels 2, 3, and 5, with girls outperforming boys at levels 2 and 3, and boys outperforming girls at level 5. It would appear that girls tend to outperform boys at the lower levels on the SAIP 1997 mathematics assessment, but boys outperform girls at the higher levels. The 1998 SAIP Reading and Writing assessments (Council of Ministers of Education, Canada, 1999) showed statistically significant gender differences for both 13- and 16-year-old Canadian students for both reading and writing assessments at all levels. The 1999 SAIP for science found few statistically significant gender differences for 13- and 16-year-old students on both the written assessment and the practical tasks portion of the examination for Canadian students (Council of Ministers of Education, Canada, 2000). When gender differences occurred they were mainly in the direction of girls outperforming boys, with the exception of boys outperforming girls at the level 4 science written assessment. Among 13- and 16-year-old Alberta students tested at all levels on the written assessment portion, the only statistically significant gender difference was for 16-year-old students at level 4; more boys than girls performed at that level. Not enough students in Alberta were tested on the practical tasks portion to allow for gender analyses. The Third International Mathematics and Science Study (TIMSS, Robitaille, Taylor, Orpwood, & Donn, 1998) population 3 (grade 12) found statistically significant mean differences between boys and girls for students in both Alberta and Canada on all mathematics and science areas tested, with all differences in the direction of higher mean scores for boys. The Programme for International Student Assessment (PISA, Bussiere et al., 2001) found statistically significant gender differences for Canadian students in favor of girls on the reading literacy component. Canadian boys outperformed girls on mathematics literacy, and no statistically significant difference was found on science literacy. The general pattern that emerges from these results show that boys tend to outperform girls on mathematics tests, especially among the high ability students, and girls tend to outperform boys on reading and writing. The patterns for science assessments seem less clear. Although these large-scale assessments show many statistically significant gender differences, due to the large sample sizes used in these studies even small differences between groups can be expected to yield statistically significant results. Unfortunately, effect sizes were not discussed in these studies. Inclusion would have put the results in a greater practical context.

Generally, results in the literature have found that test score mean differences exist between boys and girls. Those studies that do go beyond examination of mean differences and evaluate the statistical and practical importance of gender relationships find that for the most part differences tend to be small.

Increasing interest has been shown in Alberta regarding the relationships between gender and student performance on provincial tests. The purpose of this study was to investigate empirically gender relationships on both Alberta diploma examination and school-awarded scores at the grade 12 level. Specifically, this study endeavors to determine whether relationships exist between gender and scores both at the school level and the standardized test level. In addition, if relationships exist, this study attempts to determine the statistical and practical importance of these relationships.

Method

The Canadian province of Alberta maintains a high-school level testing program, established in 1984, known as the diploma examinations program. This program was designed for three main purposes: (a) individual student certification, (b) ensuring province-wide achievement standards, and (c) to report results at the individual and group levels. Eleven courses at the grade 12 level have diploma examinations, seven of which offer French translations. Five examination administrations occur per year in November, January, April, June, and August with the two largest administrations being January and June.

The diploma examination program offers two streams of courses. As stated in the Alberta Learning (2000) *Curriculum Handbook for Parents*, the 30-level stream courses "are designed primarily for students planning on entering a university or particular programs in colleges and technical schools." Courses in the 33-level stream "are designed primarily for students planning on entering some programs in colleges, technical and trade schools or entering the work force." Students in grade 12 may choose what courses they would like to take, but must complete a certain number of core courses to obtain a high school diploma. All diploma examinations are composed of two sections, a machine-scored section consisting of multiple-choice and numeric-response items (i.e., a type of item where numeric responses are written in cells and then number bubbles are filled in), and an open-ended constructed-response section. The weighting of the examinations is typically 70% machine-scored and 30% constructed-response, resulting in an examination score out of 100%. English 30 and 33 weight the open-ended section 50% each. School-awarded scores are determined by teachers who assess the performance of each student in their classes over the course of the school year. At the end of the year teachers assign each student a school-awarded score out of 100%. Each student's final course mark is arrived at by taking 50% of his or her diploma examination mark and 50% of his or her school-awarded mark. Postsecondary institutions use the final course mark to determine admission.

The development of diploma examination items by the Learner Assessment Branch follows the guidelines outlined in *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999). These standards include creating items and tests with gender-neutral language, content, symbols, words, and phrases. Although the diploma examination achievement standards are designed to be equivalent between administrations, no statistical methods are currently used to ensure that the standards or individual student scores remain constant from examination to examination. As a result, comparisons between sessions are not appropriate and therefore not pursued in this study.

All analyses were performed on both the January and June 2000 diploma sessions because students writing in these sessions are generally considered to be the most typical populations. All English and French students with valid examination and school-awarded scores were included in the analyses (i.e., the entire writing population was used, so sampling was not conducted). Correlation coefficients were used to investigate the gender relationships, with R^2 values derived from the correlations relating the percent of variation in scores accounted for by gender. Pearson correlation coefficients were calculated between gender and diploma examination scores, school-awarded scores, and differences between school-awarded scores and diploma examination scores. Cohen's (1992) effect size guidelines are used throughout this study to describe the magnitude of the correlation coefficients; 0.10, 0.30, and 0.50 represent small, medium, and large effect sizes respectively. Although the magnitude of these effect sizes were designed as general guidelines, there is every reason to believe that in the context of Alberta test scores, these effect size ranges are appropriate for interpreting the practical importance of the results (i.e., correlations in the small effect size range are of little practical importance because the percent of variation accounted for is small). In addition to the Alberta context, Cohen's (1992) effect size ranges are consistent with the interpretation of relationships between gender and test scores examined in the research literature (Willingham & Cole, 1997).

Boys were coded as 1 and girls were coded as 2; therefore, a negative correlation indicates a relationship in the direction of boys, a positive correlation is a relationship in the direction of girls. For example, a statistically significant correlation of 0.11 between gender and diploma examination scores indicates that girls scored higher than boys (i.e., a statistically significant but small mean difference between boys and girls in the direction of girls is present). Means were not listed in this study because of the redundancy in doing so (i.e., reporting correlation coefficients provides the same information).

Throughout this study the word *gender* is used to describe the physiological groups of male and female (i.e., *gender* is used in place of *sex*).

The diploma course Français 30 was not included in any of the analyses because of small numbers of students writing this examination.

A significance of less than or equal to 0.05 was used as the statistical significance criterion for all tests of significance.

Results

Tables 1 through 6 display the Pearson correlation coefficients and regression R^2 values for each relationship. The tables also present the *N* ratio information, which is the number of boys divided by the number of girls.

Tables 1 and 2 present the relationships between gender and diploma examination scores for the January and June 2000 sessions. Table 1 shows that Biology 30, Chemistry 30, English 30, Mathematics 30, Science 30, Social Studies 30, and Social Studies 33 have statistically significant correlation coefficients and R^2 values on the January 2000 examination. Table 2 shows that Chemistry 30, English 33, Mathematics 30, Social Studies 30, and Social Studies 33 have statistically significant correlation coefficients and R^2 values on the June 2000 examination. The correlations for both the January and June sessions are all in the small effect size range and show that for any subject area, gender

Table 1
Relationships Between Diploma Examination Scores and Gender, January 2000 Session

Diploma Course	*r*	*N*	*N Ratio Male/Female*	*Significance*	*R^2*	*Percent R^2*
Biology 30	−0.0228	7,577	0.5678	*p*=0.047	0.001	0.1%
Chemistry 30	−0.0315	7,378	0.8505	*p*=0.007	0.001	0.1%
English 30	0.0519	10,807	0.7493	*p*<0.001	0.003	0.3%
English 33	0.0105	6,221	1.4056	*p*=0.407	_	_
Mathematics 30	−0.0680	10,530	0.9738	*p*<0.001	0.005	0.5%
Mathematics 33	0.0017	6,258	0.9167	*p*=0.892	_	_
Physics 30	0.0063	3,636	1.6424	*p*=0.705	_	_
Science 30	−0.0987	1,011	0.9980	*p*=0.002	0.010	1.0%
Social Studies 30	−0.1393	8,883	0.8984	*p*<0.001	0.019	1.9%
Social Studies 33	−0.1290	7,150	1.0949	*p*<0.001	0.017	1.7%

Note. Significance = 2-tailed. Gender coding male=1, female=2.

Table 2
Relationships Between Diploma Examination Scores and Gender, June 2000 Session

Diploma Course	*r*	*N*	*N Ratio Male/Female*	*Significance*	R^2	*Percent* R^2
Biology 30	−0.0124	9,323	0.5714	p=0.231	–	–
Chemistry 30	−0.0642	8,444	0.8970	p<0.001	0.004	0.4%
English 30	0.0152	13,563	0.7557	p=0.077	–	–
English 33	0.0259	7,106	1.3027	p=0.029	<0.001	<0.1%
Mathematics 30	−0.0779	8,972	0.8785	p<0.001	0.006	0.6%
Mathematics 33	0.0053	5,823	1.0265	p=0.684	–	–
Physics 30	0.0162	5,833	1.7003	p=0.216	–	–
Science 30	−0.0330	1,625	0.9624	p=0.184	–	–
Social Studies 30	−0.1626	10,645	0.8534	p<0.001	0.026	2.6%
Social Studies 33	−0.1140	8,237	1.0035	p<0.001	0.013	1.3%

Note. Significance = 2-tailed. Gender coding male=1, female=2.

explains no more than 2.6% of the variation in diploma examination scores (Social Studies 30, June 2000). The results from these tables show that gender plays a small role in the explanation of diploma examination score variation.

The relationships between gender and school-awarded scores for the January and June 2000 sessions are presented in Tables 3 and 4. Table 3 shows that all courses except Chemistry 30 and Social Studies 30 have statistically significant correlations between school-awarded scores and gender for the January 2000 session. Table 4 illustrates that all courses have statistically significant correlations between school-awarded scores and gender in the June 2000 session. Similarly, Biology 30, Chemistry 30, Mathematics 30, Science 30, Social Studies 30, and Social Studies 33 changed correlation direction from negative to positive from diploma examination (Table 1) to school-awarded scores (Table 3). This means that boys outperformed girls on diploma examination scores, but girls outperformed boys on school-awarded scores. The June 2000 session correlations for Biology 30, Chemistry 30, Science 30, and Social Studies 33 showed as similar pattern of direction change, from negative in Table 2 to positive in Table 4. Although all the correlations for the relationships between gender and school-awarded scores are in the small effect size range, the changes in correlation direction and the relative difference in the percent of variation accounted for by diploma examination scores and diploma school-awarded scores is worthy of attention. Girls appear to be at a slight advantage when it comes to school-awarded scores compared with diploma examination scores.

Tables 5 and 6 display the results of correlations between gender and the difference between school-awarded scores and diploma examination scores. These tables show how much variation gender accounts for in the discrepancy between school-awarded scores and diploma examination scores. The results in both Tables 5 and 6 show that all correlations are in the direction of larger discrepancies between school and diploma examination scores being related to

Table 3
Relationships Between School-Awarded Scores and Gender, January 2000 Session

Diploma Course	*r*	*N*	*N Ratio Male/Female*	*Significance*	R^2	*Percent* R^2
Biology 30	0.1011	7,577	0.5678	p<0.001	0.010	1.0%
Chemistry 30	0.0178	7,378	0.8505	p=0.125	_	_
English 30	0.1631	10,807	0.7493	p<0.001	0.027	2.7%
English 33	0.2187	6,221	1.4056	p<0.001	0.048	4.8%
Mathematics 30	0.0212	10,530	0.9738	p=0.030	0.000	<0.1%
Mathematics 33	0.1128	6,258	0.9167	p<0.001	0.013	1.3%
Physics 30	0.1011	3,636	1.6424	p<0.001	0.010	1.0%
Science 30	0.1753	1,011	0.9980	p<0.001	0.031	3.1%
Social Studies 30	0.0006	8,883	0.8984	p=0.958	_	_
Social Studies 33	0.0335	7,150	1.0949	p=0.005	0.001	0.1%

Note. Significance = 2-tailed. Gender coding male=1, female=2.

being female. This finding is a logical extension of Tables 1 through 4, where it was shown that in general girls obtain higher school-awarded scores than boys. Tables 5 and 6 suggest that in general these slightly higher school-awarded scores, compared with diploma examination scores, are related more to girls than to boys. In general, Tables 5 and 6 show that humanities courses have larger female discrepancies between school-awarded scores and diploma examination scores than math/science courses.

The *N* ratio information in each table demonstrates that Biology 30 had the greatest gender discrepancy among students who chose to take the course; more girls than boys were enrolled in Biology 30. Physics 30 showed the opposite pattern with more boys enrolled in the course than girls. More girls

Table 4
Relationships Between School-Awarded Scores and Gender, June 2000 Session

Diploma Course	*r*	*N*	*N Ratio Male/Female*	*Significance*	R^2	*Percent* R^2
Biology 30	0.0520	9,276	0.5714	p<0.001	0.003	0.3%
Chemistry 30	0.0246	8,476	0.8970	p<0.001	0.001	0.1%
English 30	0.1457	13,301	0.7557	p<0.001	0.021	2.1%
English 33	0.1726	6,885	1.3027	p<0.001	0.030	3.0%
Mathematics 30	–0.0016	9,075	0.8785	p<0.001	<0.001	<0.1%
Mathematics 33	0.0864	5,729	1.0265	p<0.001	0.007	0.7%
Physics 30	0.0930	5,865	1.7003	p<0.001	0.009	0.9%
Science 30	0.1168	1,617	0.9624	p<0.001	0.014	1.4%
Social Studies 30	–0.0087	10,950	0.8534	p<0.001	<0.001	<0.1%
Social Studies 33	0.0536	8,092	1.0035	p<0.001	0.003	0.3%

Note. Significance = 2-tailed. Gender coding male=1, female=2.

Table 5
Relationships Between Gender and the Difference Between School-Awarded Scores and Diploma Exam Scores, January 2000 Session

Diploma Course	*r*	*N*	*N Ratio Male/Female*	*Significance*	*R^2*	*Percent R^2*
Biology 30	0.0932	7,577	0.5678	$p<0.001$	0.009	0.9%
Chemistry 30	0.0766	7,378	0.8505	$p<0.001$	0.006	0.6%
English 30	0.1216	10,807	0.7493	$p<0.001$	0.015	1.5%
English 33	0.1885	6,221	1.4056	$p<0.001$	0.036	3.6%
Mathematics 30	0.1437	10,530	0.9738	$p<0.001$	0.021	2.1%
Mathematics 33	0.1302	6,258	0.9167	$p<0.001$	0.017	1.7%
Physics 30	0.1234	3,636	1.6424	$p<0.001$	0.015	1.5%
Science 30	0.2832	1,011	0.9980	$p<0.001$	0.080	8.0%
Social Studies 30	0.2028	8,883	0.8984	$p<0.001$	0.041	4.1%
Social Studies 33	0.1759	7,150	1.0949	$p<0.001$	0.031	3.1%

Note. Significance = 2-tailed. Gender coding male=1, female=2.

than boys were registered in English 30, whereas more boys than girls were registered in English 33. The *N* ratios for all other courses indicated no major gender discrepancies.

Although the amount of variation in school-awarded scores and diploma examination scores explained by gender is generally small, the relative difference in the amount of variation explained is worthy of discussion. In some cases, changes in the direction of the correlations and large relative increases in the percent of variation explained by gender are observed when examining school-awarded scores (Tables 3 and 4) compared with diploma examination scores (Tables 1 and 2). For example, English 30 January 2000 has approximately nine times as much variation explained by gender in school-awarded

Table 6
Relationships Between Gender and the Difference Between School-Awarded Scores and Diploma Exam Scores, June 2000 Session

Diploma Course	*r*	*N*	*N Ratio Male/Female*	*Significance*	*R^2*	*Percent R^2*
Biology 30	0.0950	9,276	0.5714	$p<0.001$	0.009	0.9%
Chemistry 30	0.1406	8,476	0.8970	$p<0.001$	0.020	2.0%
English 30	0.1446	13,301	0.7557	$p<0.001$	0.021	2.1%
English 33	0.1365	6,885	1.3027	$p<0.001$	0.019	1.9%
Mathematics 30	0.1215	9,075	0.8785	$p<0.001$	0.015	1.5%
Mathematics 33	0.1005	5,729	1.0265	$p<0.001$	0.010	1.0%
Physics 30	0.0905	5,865	1.7003	$p<0.001$	0.008	0.8%
Science 30	0.1604	1,617	0.9624	$p<0.001$	0.026	2.6%
Social Studies 30	0.2189	10,950	0.8534	$p<0.001$	0.048	4.8%
Social Studies 33	0.1771	8,092	1.0035	$p<0.001$	0.031	3.1%

Note. Significance = 2-tailed. Gender coding male=1, female=2.

scores versus diploma examination scores. The trend seems to be that of gender playing a more important role in explaining school-awarded score variation than explaining diploma examination score variation.

Discussion

The results presented in this article show that gender has a small statistical effect on diploma examination scores. This observation holds true for both the January and June 2000 sessions. These findings are consistent with the literature on standardized test gender differences at the grade 12 level (Willingham et al., 1997). The effect of gender on school-awarded scores was also small, although larger than the effect of gender on diploma examination scores. Girls tended to have larger differences between their school-awarded scores and their diploma examination scores than boys. The individual effects of these differences were as small as the gender differences in school-awarded scores. Although direct comparisons between sessions cannot be made in this study, in general the gender relationship results between the January and June 2000 sessions appear to be similar.

Across groups, school-awarded scores tended to be higher for students than corresponding diploma examination scores. The results presented in this article show that although the statistical effects are small, girls received higher school-awarded scores than boys except in Social Studies 30, Mathematics 30, and possibly Chemistry 30. These results are consistent with Halpern (1997) and Dwyer and Johnson (1997); both studies found that girls tend to earn higher school grades than boys in most subject areas.

The gender patterns for the mathematics, science, and social studies courses are noteworthy because these are subject areas where boys traditionally are thought to outperform girls. The SAIP 1997 mathematics assessment (Council of Ministers of Education, Canada, 1997) found that boys outperformed girls at the higher ability levels, and the TIMSS mathematics and science assessments (Robitaille et al., 1998) found that boys consistently outperformed girls. The PISA (Bussiere et al., 2001) assessment found boys outperforming girls in mathematics proficiency. In the context of Alberta, it has been discussed that the curricula for social studies courses favor boys over girls. Willingham et al. (1997) suggest that performance differences in social studies could be due to differential gender interest in "geopolitical" areas of study. With regard to the mathematics and science areas, a differential gender performance phenomenon has been well documented in the literature (Cleary, 1992; Entwistle, Alexander, & Olsen, 1994; Friedman, 1989; Hyde, Fennema, & Lamon, 1990; Loewen, Rosser, & Katzman, 1988). Although the effect is small, the results in this article show that boys either score slightly higher or the same as girls on the mathematics, science (with the exception of Biology 30, January 2000), and social studies diploma examinations. The school assessment shows the opposite pattern, with girls generally outperforming boys. Among the science courses (i.e., Physics 30), no significant differences between male and female diploma examination performance are observed; however, a significant relationship for school-awarded scores in favor of girls is noted. This finding is counter to research finding that boys tend to outperform girls in areas of physics (Bae & Smith, 1997). It is unlikely that the differences we see between boys and girls on school versus diploma examination results are due to natural differences in performance. If this were the case, theoretically both diploma examination scores and school-awarded scores would show the same pattern. For example, if the Social Studies 30 curriculum is really biased toward boys, one would expect to see a strong gender relationship (in favor of boys) between both school-awarded and diploma examination scores.

Girls traditionally outperform boys in the area of language arts (Cole, 1997; Willingham et al., 1997). The SAIP 1998 Reading and Writing assessment (Council of Ministers of Education, Canada, 1999) found that girls consistently outperformed boys. Similarly, the PISA assessment (Bussiere et al., 2001) found that girls outperformed boys on reading literacy. In this study only two statistically significant results were found for the diploma examination scores: English 30, January 2000, and English 33, June, both in favor of girls. The statistical effect of these differences was extremely small. In stark contrast, significant relationships were found for both January and June 2000, English 30 and English 33, between gender and school-awarded scores all in favor of girls. The percent of variation accounted for in school-awarded scores by gender was many times greater than the percent of variation accounted for in diploma examination scores by gender. These results present a muddled picture of the relationships between gender and student achievement in the English language arts subject areas.

The gender relationship patterns for diploma examination scores versus school-awarded scores are not consistent for most of the courses examined in this study. For example, the gender relationships for diploma examination scores are mixed, with boys outperforming girls in some courses and girls outperforming boys in others. For the school-awarded score results, every course that showed statistically significant gender relationships, with percent R^2 values greater than 0.1, had results in the direction of girls outperforming boys. As both school-awarded scores and diploma examination scores are considered to be measures of how well students have learned and understood curriculum material, it is expected that similar gender difference patterns (in direction and magnitude) would be found for both school-awarded and diploma examination scores. This study found, however, that this was not the case. It would seem that when it comes to school-awarded scores, girls have a slight advantage over boys in virtually every diploma course. There are a number of possible reasons why female students tend to obtain slightly higher school-awarded scores than boys. Perhaps female students do better on assignments or are better behaved in the classroom than boys and therefore are subject to a halo effect when teachers assign school-awarded marks. Or perhaps girls participate more in class discussions than boys, and therefore receive higher school marks. The literature describes various social factors that may contribute to differences in student achievement including: conformity to gender stereotypes (Bouchard & St-Amant, 2000), classroom interaction patterns (Lindow et al., 1985), or self-esteem (Dukes & Martinez, 1994; Francis & James, 1996). Based on the results of this study it is unlikely that there are systematic gender biases in any of the diploma course curriculums. This conclusion results from the fact that one would expect medium to large effect sizes (i.e., great statistical and practical importance) along with consistency between school and examination results for curriculum bias to be considered a reasonable possibility. There does appear to be some sort of differential favoritism in favor of girls in terms of school-awarded scores. This suggests that school-awarded scores are assigned to students using information that is at least somewhat independent of student curriculum knowledge (e.g., factors such as classroom behavior may influence school-awarded scores).

To help unravel some of the gender patterns between gender and school-awarded scores versus diploma examination scores, a study that examines relationships between teacher gender, student gender, and scores would be useful. In addition, examining the component parts (i.e., multiple-choice, written-response) of the diploma examinations may provide greater understanding about which sections of the examinations show relationships with student gender, teacher gender, and student-teacher interactions. An additional study that would be interesting to undertake in Alberta would be to examine whether relationships exist between teachers' expectations of students' performance in relation to gender. A study looking at whether gender plays a role in classroom social behavior and what influence these social factors have on school-awarded scores would help to unravel the relationships between student gender and school-awarded scores. Clearly additional studies such as those suggested in this article would be useful in addressing questions raised by the current study.

Although not all findings in this study are consistent with the literature, it is clear from the results presented that there are small relationships between gender and diploma examination scores in Alberta. The same holds true for school-awarded scores. It appears that girls have a slight advantage over boys when it comes to school-awarded scores, but the overall statistical importance of these differences is small. The results of this study contribute to the field of education by providing information on the role of gender in the context of Alberta standardized tests and school-awarded scores.

References

Alberta Learning. (2000). *Curriculum handbook for parents*. Edmonton, AB: Crown in Right of Alberta, Alberta Learning. [Online] Available e-mail: curric.contact@edc.gov.ab.ca

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bae, Y., & Smith, T.M. (1997). Women in mathematics and science. In National Center for Educational Statistics, *The condition of education 1997* (pp. 14-21). Washington, DC: US Department of Education, Office of Research and Improvement.

Bouchard, P., & St-Amant, J.-C. (2000). Gender identities and school success. *Alberta Journal of Educational Research, 46*, 281-283.

Bussiere, P., Cartwright, F., Crocker, R., Ma, X., Oderkirk, J., & Zhang, Y. (2001). *Measuring up: The performance of Canada's youth in reading, mathematics and science*. Ottawa: Human Resources Development Canada, Council of Ministers of Education Canada, and Statistics Canada.

Cleary, T.A. (1992). Gender differences in aptitude and achievement test scores. In *Sex equity in educational opportunity, achievement, and testing: Proceedings of the 1991 ETS invitational conference* (pp. 51-90). Princeton NJ: Educational Testing Services.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Cole, N.S. (1997). *The ETS gender study: How females and males perform in educational settings.* Princeton, NJ: Educational Testing Service.

Council of Ministers of Education, Canada (CMEC). (1997). *Report on mathematics assessment School Achievement Indicators Program (SAIP) 1997*. Toronto, ON: Author.

Council of Ministers of Education, Canada (CMEC). (1999). *Report on reading and writing assessment School Achievement Indicators Program (SAIP) 1998*. Toronto, ON: Author.

Council of Ministers of Education, Canada (CMEC). (2000). *Report on science assessment School Achievement Indicators Program (SAIP) 1999*. Toronto, ON: Author.
Dukes, R.L., & Martinez, R. (1994). The impact of ethgender on self-esteem among adolescents. *Adolescence, 29*, 105-115.
Dwyer, C.A., & Johnson, L.M. (1997). Grades, accomplishments and correlates. In W. Willingham & N.S. Cole (Eds.), *Gender and fair assessment* (pp. 127-156). Mahwah, NJ: Erlbaum.
Entwistle, D.R., Alexander, K.L., & Olsen, L.S. (1994). The gender gap in math: Its possible origins in neighborhood effects. *American Sociological Review, 59*, 822-838.
Francis, L.J., & James, D.J. (1996). The relationship between Rosenberg's construct of self-esteem and Eysenck's two-dimensional model of personality. *Personality and Individual Differences, 21*, 483-488.
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. (1997). Sex differences in intelligence: Implications for education. *American Psychologist, 52*, 1091-1102.
Han, L., & Hoover, H.D. (1994, April). *Gender differences in achievement test scores*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans.
Hyde, J., Fennema, E., & Lamon, S. (1990). Gender differences in mathematics performance: A meta analysis. *Psychological Bulletin, 107*, 139-155.
Lindow, J., Marrett, C., & Wilkinson, L.C. (1985). Overview. In L.C. Wilkinson & C.B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 1-15) Orlando, FL: Academic Press.
Loewen, J.W., Rosser, P., & Katzman, J. (1988). *Gender bias in SAT items*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
Noble, J., Davenport, M., Schiel, J., & Pommerich, M. (1999). *High school academic and noncognitive variables related to the ACT scores of racial/ethnic and gender groups. Research Report #ACT-RR-99-6*. Iowa City, IA: American College Testing Program.
Robitaille, D.F., Taylor, A.R., Orpwood, G., & Donn, J.S. (1998). *The third international Mathematics and Science Study TIMSS—Canada Report. Volume 4: Senior secondary*. Vancouver, BC: Department of Curriculum Studies, University of British Columbia.
Smith, T. M., Young, B.A., Bae, Y., Choy, S.P., & Alsalam, N. (1997). The condition of education 1997. *Research Report #NCES 97-388*. Washington, DC: US Department of Education, National Center for Education Statistics.
Thorne, B. (1993). Girls and boys together ... but mostly apart: Gender arrangements in elementary schools. In L. Richardson & V. Taylor (Eds.), *Feminist frontiers III* (pp. 115-125). New York: McGraw Hill.
Willingham, W.W., & Cole, N.S. (1997). *Gender and fair assessment*. Mahwah, NJ: Erlbaum.
Willingham, W.W., Cole, N.S., Lewis, C., & Leung, S.W. (1997). Test performance. In W. Willingham & N.S. Cole (Eds.), *Gender and fair assessment* (pp. 55-126). Mahwah, NJ: Erlbaum.