Advanced Placement and International Baccalaureate Performance Differences as a Function of Gender and School Size

Abstract

Student participation and student performance on the Advanced Placement and International Baccalaureate exams for the 2008-2009 school year was examined to ascertain the extent to which differences were present as a function of high school size and student gender. Using Academic Excellence Indicator System data for traditionally configured high schools in Texas, statistically significant differences were yielded. In the 2008-2009 school year, girls had higher participation rates than did boys. Moreover, girls outperformed boys regardless of school size. The performance of boys and girls on these exams differed as a function of school size, with students in larger size high schools participating at a higher rate and outperforming students at smaller size high schools. Implications of the findings are discussed and suggestions for further research are made.

Keywords: Advanced Placement, International Baccalaureate, gender, school size, student achievement

Barker and Gump (1964) concluded that smaller schools, though offering fewer activities, allowed a higher percentage of students to become involved in these activities than larger schools. A few years later, Conant’s (1967) examination of over 2,000 high schools suggested that smaller schools (i.e., fewer than 750 students) could not provide as diverse and efficient educational programs as larger schools. Goodlad (1984) entered the school size debate by compiling studies over several years that indicated smaller units should be created from larger schools (i.e., elementary schools should be no larger than 300 students and secondary schools should be 600 or fewer students). In their review of the literature on school size, Slate and Jones (2005) concluded that the differing results that were reported in previous studies might be attributed to the complexity of the issue and the focus of researchers and policymakers on finding a simple answer to a complex issue.

Schools that are at either extreme in terms of size (i.e., very small or very large) might lack the needed resources to serve students efficiently. Moreover, the relationship between school size and school quality may be curvilinear (Slate & Jones, 2005; Stiefel, Schwartz, Iatarola, & Chellman, 2009; Werblow & Duesbery, 2009). More recently, Greeney (2010) noted that students enrolled in larger Texas high schools had higher passing rates on the Texas Assessment of Knowledge and Skills than students enrolled in smaller Texas high schools. Monk and Haller (1993) suggested that an optimal school size that will work best for all students does not exist. Similarly, Howley (2001) concluded that the characteristics of the students served by the school might influence what school size would be best and that a one-size-fits-all answer to the question of optimal school size does not exist. Lee and Smith (1997), in a longitudinal study of over 800 high schools indicated that student performance was strongest in schools of medium size (i.e., between 600 and 900). In their review of school size research, Leithwood and Jantzi (2009) noted that a positive relationship between school size and student achievement was documented in five studies, a curvilinear relationship existed between school size and student achievement in six studies, and a negative relationship between school size and
student achievement was reported in eight studies. Because over 12,500 high schools in the United States participate in the Advanced Placement (AP) or International Baccalaureate (IB) program, examining the impact of school size on the performance of boys and girls taking AP tests and IB exams in Texas may provide useful information.

The AP program was originally designed to provide advanced students the opportunity to earn college credit while enrolled in high school (Hertberg-Davis, Callahan, & Kyburg, 2006). Motivated and academically prepared high school students study a rigorous curriculum and are able to demonstrate their proficiency at college level courses (Dixon, 2006). In recent years, the College Board (2010a) has encouraged schools to recruit and prepare students for enrollment in AP classes. Schools are encouraged to provide PreAP classes that will prepare students for the rigor of AP coursework (College Board, 2010b).

The College Board offered over 35 AP courses in 2009 and 26.5% of the 2009 high school graduates completed at least one AP course during high school. Almost 16% of the 2009 graduating class received a score of 3 or higher and were potentially eligible to receive college credit (College Board, 2010a). Because the same AP exam is administered each year to all students in a subject area, the AP program provides an external standard to evaluate the mastery of content by students (Robinson, 2003). Van Tassel-Baska (2005) stated that the AP program was one way to measure the quality of secondary student learning and collaboration between secondary education and higher education. The AP exams are scored on a scale of 1 (no recommendation) to 5 (extremely well qualified) (College Board, 2010b). A recommendation by the American Council of Education is that students earning a 3 or higher on an AP exam should receive college credit (Ewing, 2006). However, colleges and universities set their own standard and many institutions require a score of 4 or 5 for a student to earn college credit (Ewing, 2006).

Another program that has been utilized by secondary schools for students is the IB Program (Hertberg-Davis et al., 2006). The IB program began in 1968 at the International School of Geneva to meet the needs of internationally mobile students preparing to enter college (International Baccalaureate Organization, 2005-2010a). The growth of the IB program has been slower than the growth of the AP program. During the 2004 academic year, 1,335 schools were registered IB Diploma Programme schools. In the 2009 academic year, 2,025 schools were registered IB Diploma Programme schools. In 2009, students took a total of 29,962 IB examinations (International Baccalaureate, 2010). Similar to the AP program, students who successfully pass IB examinations may be eligible to earn college credit (International Baccalaureate Organization, 2005-2010a). The scores for IB exams range from 1 (lowest) to 7 (highest) and a score of 4 or higher may allow a student to earn college credit. The goals of the IB program are clarified in the IB mission statement: (a) to develop young people who are caring, knowledgeable, and will work to improve the world; (b) to develop challenging curriculum and assessments; and (c) to develop lifelong learners (International Baccalaureate Organization, 2005-2010b).

Over 3,600 colleges and universities receive an annual AP Exam Report and about 90% of 4-year colleges and universities offer either credit or advancement for students who pass an AP exam (College Board, 2010b). Van Tassel-Baska (2005) contended that students without some evidence of advanced work (e.g., AP, IB, or dual-enrollment) in high school are less likely to be accepted to the top 300 colleges in the United States than are students who completed advanced course work. Traditionally, more men have pursued postsecondary degrees than women. However, in the past 30 years, that trend has reversed (Wells, Seifert, Padgett, Park, & Umbach, 2011). In their study of the gender gap between men and women earning college
degrees, Wells et al. (2011) documented that more women than men completed postsecondary degrees. Hussar and Bailey (2007) predicted that by 2016, 60% of all college students would be women. Moore and Slate (2008) documented that a higher percentage of girls (17%) enrolled in AP courses than the percentage of boys (13%) who enrolled in AP courses, but that the proportion of boys who scored a 3 or higher was slightly higher than the proportion of girls who scored 3 or higher.

In her examination of existing research, Ewing (2006) cited several researchers (e.g., Burnham & Hewitt, 1971; Dodd, Fitzpatrick, De Ayala, & Jennings, 2002; Morgan & Crone, 1993; Morgan & Ramist, 1998) who indicated a positive outcome when students enrolled in a more advanced college course instead of an introductory college course based on their successful performance on an AP exam. Hargrove, Godin, and Dodd (2008) demonstrated strong evidence on a greater number of college credits earned, higher college GPA, and increased graduation rate for students who participated in the AP Program. Dougherty, Mellor, and Jian (2006) documented that students who passed at least one AP exam were more likely to graduate from college in five or fewer years than students who completed an AP course but either did not take the AP exam or did not pass the AP exam. Colangelo, Assouline, and Gross (2004) determined that 59% of students who completed at least one AP course earned a bachelor’s degree compared to 33% for students who did not complete any AP courses in high school.

Theoretical Framework

In this study, economy of scale was used to understand the manner in which student performance on AP and IB exams might be influenced by school size. According to Jewell (1989), economy of scale is an economic term suggesting that larger schools have greater efficiency and lower costs per student. Also, schools may gain efficiency through specialization because teachers are able to teach courses in their field of expertise. As such, larger schools may be able to offer students a more diverse curriculum and employ teachers with greater expertise in the courses being taught with lower costs per pupil (Jewell, 1989). Thus, students in larger schools may benefit from the expanded course offerings and teacher expertise, resulting in increased academic success on AP and IB exams.

Purpose of the Study

Slate and Jones (2005) suggested that asking open-ended questions might generate results devoid of bias, avoiding the temptation to structure questions to match a predetermined conclusion regarding the best size for a school. However, the debate over school size and student achievement has not yet provided definitive answers concerning the most appropriate size of school (Tanner & West, 2011). The extent to which gender and school size are related to student performance on AP and IB exams is unclear. Therefore, the purpose of this study was to determine the extent to which high school size was related to student performance for boys and girls on the 2009 AP and IB exams in Texas.
Research Question

The following research questions were addressed in this study: (a) What is the difference in the percentage of boys and the percentage of girls taking the 2009 AP and IB exams as a function of high school size? and (b) What is the difference in the passing rate for boys and passing rate for girls on the 2009 AP and IB exams as a function of high school size?

Method

Participants

Data from all Texas high schools with a grade 9-12 configuration for the 2008-2009 school year were used in this study. Charter schools, alternative placement schools, and schools with other grade configurations were excluded from this study because of their differences from traditionally configured high schools. The Texas Academic Excellence Indicator System (AEIS) data, published annually by the Texas Education Agency (TEA), were downloaded from the TEA website. High schools whose data were analyzed in this study were divided into three groups based on size utilizing Green and Barnes’ (1993) definition: small (i.e., between 100 and 799 students), medium (i.e., between 800 and 1,199 students), and large (i.e., more than 1,200 students). Presented in Tables 1 and 2 are the number of schools in each category.

Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>n of schools</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Schools</td>
<td>365</td>
<td>22.99</td>
<td>8.37</td>
</tr>
<tr>
<td>Medium Schools</td>
<td>90</td>
<td>18.20</td>
<td>7.66</td>
</tr>
<tr>
<td>Small Schools</td>
<td>448</td>
<td>17.95</td>
<td>14.52</td>
</tr>
<tr>
<td>Girls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Schools</td>
<td>365</td>
<td>28.04</td>
<td>8.33</td>
</tr>
<tr>
<td>Medium Schools</td>
<td>90</td>
<td>23.44</td>
<td>7.79</td>
</tr>
<tr>
<td>Small Schools</td>
<td>448</td>
<td>22.64</td>
<td>14.86</td>
</tr>
</tbody>
</table>
Table 2

Descriptive Statistics for the Difference in Students Scoring Above the Criterion on AP and IB Exams by Gender and School Size

<table>
<thead>
<tr>
<th>Variable</th>
<th>n of schools</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Schools</td>
<td>365</td>
<td>20.49</td>
<td>9.97</td>
</tr>
<tr>
<td>Medium Schools</td>
<td>90</td>
<td>16.08</td>
<td>11.94</td>
</tr>
<tr>
<td>Small Schools</td>
<td>446</td>
<td>9.78</td>
<td>14.64</td>
</tr>
<tr>
<td>Girls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Schools</td>
<td>365</td>
<td>25.69</td>
<td>11.23</td>
</tr>
<tr>
<td>Medium Schools</td>
<td>90</td>
<td>20.50</td>
<td>14.50</td>
</tr>
<tr>
<td>Small Schools</td>
<td>446</td>
<td>12.37</td>
<td>16.82</td>
</tr>
</tbody>
</table>

Instrumentation and Procedures

Archival AEIS data, obtained from the TEA website for the 2008-2009 school year, were downloaded from the website as .dat files and contained student passing rates for AP and IB exams separated by gender and school size. The .dat files were merged using the Statistical Package for Social Sciences (SPSS). Because the data from the TEA website were reported by school district, errors are assumed to be minimal or few. Readers are referred to the technical manuals on the TEA website for information regarding score reliability and score validity. Included in the AEIS report are three values that are calculated for AP and IB exams: (a) percentage of students in grades 11 and 12 who took at least one AP or IB test; (b) percentage of students who scored at or above criterion score (i.e., a score of 3 on the AP test or a score of 4 on the IB test); and (c) percentage of scores at or above the criterion score (Texas Education Agency, 2011).

Results

The standardized skewness coefficients (i.e., the skewness value divided by the standard error of skewness) and the standardized kurtosis coefficients (i.e., the kurtosis value divided by the standard error of kurtosis) were calculated to determine normality. The majority (22 out of 24) of the standardized skewness and kurtosis coefficients were outside the range of normality (i.e., +/- 3, Onwuegbuzie & Daniel, 2002). When the statistical analyses were performed, the assumptions for the Box’s Test of Equality of Covariance and Levene’s Test of Equality of Error Variances were violated. However, the robustness of the Multivariate Analysis of Variance (MANOVA) made it the appropriate statistical procedure to be utilized even though the assumptions were violated (Field, 2009). Therefore, a MANOVA was performed to determine if the percentage of boys and the percentage of girls taking AP or IB exams and the percentage of boys and girls above the criterion on AP or IB exams differed as a function of school size.
Focused upon in the first research question was the effect of school size (i.e., small, medium, or large) upon the percentage of boys taking AP or IB exams and the percentage of girls taking AP or IB exams. A statistically significant difference was present for all groups of students, Wilks’ $\Lambda = .95$, $p < .001$, $n^2 = .02$. The effect size for this statistically significant difference was small (Cohen, 1988). Statistically significant differences were present in the univariate follow-up analysis of variance procedures for the percentage of boys taking AP or IB exams, $F(2, 900) = 19.02$, $p < .001$, $n^2 = .04$, small effect size, and for the percentage of girls taking AP or IB exams, $F(2, 900) = 21.19$, $p < .001$, $n^2 = .05$, small effect size (Cohen, 1988). Thus, both the percentage of boys taking AP or IB exams and the percentage of girls taking AP or IB exams differed as a function of school size.

Scheffe post hoc procedures revealed statistically significant differences were present between students enrolled in small and large schools and between students enrolled in medium and larger schools, but no difference was noted between students enrolled in small and medium schools. The percentage of boys taking AP or IB exams was higher in large schools (i.e., 22.92%) than the percentage of boys taking AP or IB exams was in either medium schools (i.e., 18.20%) or small schools (17.95%). Similar results were present for the percentage of girls taking AP or IB tests. Larger schools had a higher percentage of girls taking either AP or IB exams (i.e., 28.04%) than medium schools or small schools (i.e., from 23.44% to 22.64%). Readers are referred to Table 1 for the descriptive statistics for the percentages of boys and girls taking AP or IB tests as a function of school size.

Addressed in the second research question was the effect of school size upon the percentage of boys scoring above the criterion on AP or IB exams and on the percentage of girls scoring above the criterion on AP or IB exams. The MANOVA indicated a statistically significant difference, Wilks’ $\Lambda = .84$, $p < .001$, $n^2 = .08$. The effect size for this statistically significant difference was moderate (Cohen, 1988). Statistically significant differences were present in the univariate follow-up analysis of variance procedures for the percentage of boys scoring above the criterion on AP or IB exams, $F(2, 898) = 72.13$, $p < .001$, $n^2 = .14$, large effect size and for the percentage of girls scoring above the criterion on AP or IB exams, $F(2, 898) = 84.76$, $p < .001$, $n^2 = .16$, large effect size (Cohen, 1988). Thus, both the percentage of boys scoring above the criterion on AP or IB exams and the percentage of girls scoring above the criterion on AP or IB exams differed as a function of school size.

Scheffe post hoc procedures revealed statistically significant differences were present for all groups. The percentage of boys scoring above the criterion increased as the school size increased (i.e., from 9.78% to 20.49%). Similar results were present for the percentage of girls scoring above the criterion. Small schools had the lowest percentage of girls meeting the criterion (i.e., 12.37%), followed by medium schools (i.e., 20.50%), and larger schools (i.e., 25.69%). Readers are referred to Table 2 for the descriptive statistics for the percentages of boys and girls scoring above the criterion on AP or IB tests as a function of school size.
Discussion

The focus of this study was the effect of school size on the performance of girls and boys on AP and IB participation and exam performance in Texas for the 2008-2009 school year. In this study, large schools had the highest percentage of both boys and girls taking AP or IB exams and the highest percentage of boys and girls than did either small schools or medium schools. These results differ from Barker and Gump (1964) and Goodlad (1984) who suggested that smaller schools are better than larger schools. However, Greeney (2010) noted similar results in his study of school size and TAKS performance, with larger schools having a higher passing rate than smaller schools. More recently, Zoda, Combs, and Slate (2011) documented higher passing rates on TAKS measures by Black students in larger schools than by Black students in smaller schools. The higher participation rates and passing rates on AP and IB exams may be explained by the economies of scale theoretical framework (Jewell, 1989) previously mentioned. Larger schools might be more able than smaller schools to provide a variety of AP or IB classes and teachers who have more expertise in these specialized courses. Although the effect size for the overall participation rate was small, the number of students impacted and the pattern that developed indicated the need for further study into the effect of school size on student participation on the AP and IB tests. The moderate effect size for the overall performance on AP and IB tests for boys and girls as a function of school size also suggests the need for further study.

As documented in Figure 1, the percentage of girls taking AP or IB exams was higher than the percentage of boys taking AP or IB exams regardless of the school size. This finding agreed with Moore and Slate (2008), who also reported that a higher percentage of girls than boys participated in AP courses. The percentage of boys and girls scoring above the criterion was the highest in large schools. As shown in Figure 2, regardless of school size, a higher percentage of girls scored above the criterion than the percentage of boys who scored above the criterion. This result differs from the finding of Moore and Slate (2008) and Morris and Slate (2012), in which a higher percentage of boys than girls scored above the criterion.

Important questions remain that need to be addressed in future research studies: (a) Are teacher expectations different for boys and girls in these classes?; (b) What is the impact of school size on the expectations of students and teachers?; and (c) What other influences impact student performance on these tests? Few multiyear studies have been conducted to determine the effect of gender and school size on the participation in and performance on AP and IB tests. Therefore, little information is known about the effect of gender and school size on the performance on AP and IB exams. Results from this study and other studies could be used to inform practice and make curriculum decisions regarding these courses. School leaders may benefit from an analysis of gender differences and school size on the participation in and performance on AP and IB exams as these two programs continue to grow. Given the importance of providing students with a high quality education, more research is needed to determine if programs such as AP and IB are the best method of educating all students.

Several cautions are in order lest readers overgeneralize the findings of this study. First, this study is limited to students in Texas who completed AP or IB exams and may not be generalizable to high school students in other states. Second, because students
enrolled in AP or IB courses are not required to take the AP or IB exams, the sample of students is self-selected. As such, the results may not be generalizable to students who enroll in AP and IB courses. Third, because not all students complete AP and IB courses at the same rate, differential attrition may be a threat to validity. Therefore, readers are urged to exercise caution when generalizing the results of this study.

Figure 1. **Percentage of boys and girls taking AP and IB exams as a function of school size**

Figure 2. **Percent of boys and girls scoring above criterion on AP and IB exams as a function of school size**
In conclusion, our analysis of student participation and performance by school size on the Advanced Placement and International Baccalaureate exams for the 2008-2009 school year revealed that girls had higher participation rates than did boys. Moreover, girls outperformed boys regardless of school size. Similar to recent studies (Greeney & Slate, 2012; Zoda et al., 2011), students enrolled in larger-size schools performed better on state-mandated assessment measures than did students enrolled in smaller-size schools. Researchers are urged to conduct research investigations in other states to determine the extent to which these results, based on Texas students and Texas schools, are generalizable to students and schools in other states.

References


Jill B. Morris is a doctoral student at Sam Houston State University. Also, she is a Master Teacher at Bryan High School, in Bryan, TX.

John R. Slate, Ph.D., is a Professor in the Department of Educational Leadership and Counseling at Sam Houston State University. He teaches academic and professional writing and statistical analysis.