School Performance Indicators, Accountability Ratings, and Student Achievement

Authors

Shirley A. Jackson, Ed.D., is an Adjunct Assistant Professor in the Department of Curriculum & Instruction at Sam Houston State University in Huntsville, Texas

Fred C. Lunenburg, Ph.D., is the Jimmy N. Merchant Professor of Education in the Department of Educational Leadership & Counseling at Sam Houston State University in Huntsville, Texas

Abstract

Researchers have been challenged to find school-level characteristics that make a difference in student achievement. This study focused on a diverse sample of 24 middle schools to examine differences between schools rated Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable on four performance indicator dimensions: academic excellence, developmental responsiveness, social equity, and organizational structures. The study also examined relationships between performance indicators and student achievement. A random sample of teachers from each school and the principal provided data on the school’s performance indicator dimensions. School accountability ratings, student achievement scores, and demographic characteristics were obtained from the state department of education. Significant differences were found on all four performance indicator dimensions based on school accountability ratings, as well as relationships between performance indicators and student achievement.

Keywords: academic emphasis, developmental responsiveness, organizational structures, school accountability ratings, social equity, student achievement
The No Child Left Behind Act of 2001 (Public Law 107-110) sets demanding accountability standards for schools, school districts, and states, including new state testing requirements designed to improve education. For example, the law requires that states develop both content standards in reading and mathematics and tests that are linked to the standards for grades 3 through 8, with science standards and assessments to follow. States must identify adequate yearly progress (AYP) objectives and disaggregate test results for all students and subgroups of students based on socioeconomic status, race/ethnicity, English language proficiency, and disability. Moreover, the law mandates that 100 percent of students must score at the proficient level on state tests by 2014.

Will schools, school districts, and states be able to respond to the demand? The standards demanded by federal and state legislation are particularly problematic for middle schools. Not only are middle schools dealing with documented declines in student achievement (Gottfried, 2010; Kinney, 2008; Linn, Lewis, Tsuchida, & Songer, 2000; Ritter, Barnett, Denny, & Albin, 2009; Usher, 2009), but they are also facing the additional developmental issues associated with the well being of young adolescents (Cook, MacCoun, Muschkin, & Vigdor, 2008; National Middle School Association [NMSA], 2004; Whitley, Lupart, & Beran, 2007; Woolley & Bowen, 2007).

The interest generated by the Carnegie report, *Turning Points: Preparing American Youth of the 21st Century* (Carnegie Council on Adolescent Development, 1989) resulted in the *Middle Schools Initiative* and the establishment of several alliances among educators, researchers, national organizations, and private foundations. One of these alliances is the National Forum to Accelerate Middle-Grades Reform which was established in 1997 because of declining test scores, increasing school violence, and the debate over the nature and purpose of middle-grades education. The Forum is funded by the Edna McConnell Clark Foundation, Bill and Melinda Gates Foundation, Ewing Marion Kauffman Foundation, John S. and James L. Knight Foundation, and the W.K. Kellogg Foundation (The National Forum to Accelerate Middle-Grades Reform [The National Forum], 2004). The Forum has articulated a variety of ways to identify high-performing middle schools, provide best practices, promote effective policies and leadership, and encourage public participation. It has successfully influenced the national discourse on middle school education, and it has received support from the U.S. Department of Education (The National Forum, 2004).

The National Forum has generated several initiatives including *Schools to Watch* (STW), a program designed to impact schools at the classroom level. A second initiative is *Comprehensive School Reform* (CSR), the col-
laboration of seven comprehensive school reform model providers to discuss common educational issues faced by middle school administrators. The third initiative is *The Next Generation of Middle Grade Leaders*, a collaborative effort of 60 leaders from the states working most closely with the National Forum. Together these initiatives form the basis of the Forum's commitment to informing and engaging the public with topics that include school-level performance indicators, school accountability, and student achievement.

Researchers have been challenged to go beyond socioeconomic status in search for school-level performance indicators that make a difference in student achievement. In its vision statement, the National Forum identifies what it means for a middle school to be academically excellent, developmentally responsive, socially equitable, and have the organizational structures in place to fulfill its mission. After seven years of research, The National Forum (2004) developed 61 performance indicators, forming four dimensions: academic excellence, developmental responsiveness, social equity, and organizational structures.

The means for measuring school accountability is stipulated in the No Child Left Behind (NCLB) Act of 2001 (Public Law 107-110). According to the NCLB law, school accountability is ensured by measuring student achievement on state mandated tests. The academic measurements for this federal law vary widely from state to state. Federal auditors recently found that the number of schools facing federal sanctions is growing. Nationwide, 4,509 schools serving more than 2 million children - or about 8 percent of all federally funded schools - have failed to bring enough students to grade level for four or more consecutive years (Ramirez, 2007). Most of these schools are in low-income districts with high numbers of students from racial and ethnic minority groups in California, Illinois, Michigan, New York, Pennsylvania, and Texas.

**Overview of the Study**

In Texas, student achievement is measured by the Texas Assessment of Knowledge and Skills (TAKS) test. This is the latest in a series of state achievement tests administered since 1984. The Texas Education Agency (TEA) (2008) recently reported that one in ten schools was identified as academically unacceptable, the lowest of four ratings, which include “Exemplary,” “Recognized,” “Academically Acceptable,” and “Academically Unacceptable.” In addition, more than 200 Texas schools fell below the federal standards specified by the NCLB law (TEA, 2008). The following research questions were addressed in this study:
1. What are the differences between schools rated Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable on four performance indicator dimensions: academic excellence, developmental responsiveness, social equity, and organizational structures?

2. What are the relationships between performance indicators and student achievement, disaggregated for all students and subgroups of students based on socioeconomic status, race/ethnicity, English language proficiency, and disability?

**Participants**

The study focused on a diverse sample of 24 middle schools. For the purpose of this study, only self-contained middle schools that taught grades 6, 7, and 8, with their own administrative staff and held accountable by the Texas Education Agency were considered. Using these criteria, a population of 797 middle schools was identified.

As a result of their students’ achievement on the Texas Assessment of Academic Skills (TAKS) test in reading, writing, and mathematics, ten of the 797 schools (1%) were designated by the Texas Educational Agency (TEA) as Exemplary; 285 of the 797 schools (36%) were designated Recognized; 494 of the 797 schools (62%) were designated Academically Acceptable; and eight of 797 schools (1%) were designated Academically Unacceptable (TEA, 2009). With this as a foundation, a power analysis (Cohen, 1988, 1992; Olejnik, 1984) was conducted to determine an adequate sample size. According to Cohen, when the population is 797, a sample size of 23 would be required to detect a large effect size at the < .05 level.

The selection of participants began with the 10 Exemplary schools and 8 Academically Unacceptable schools. To guarantee a nonproportional stratified sample of at least 23 schools, an additional 8 Recognized schools and 10 Acceptable schools were randomly selected from the Texas Education Agency’s (TEA) method of creating campus groups, totaling 36 middle schools. These campus groups consisted of public schools that closely matched on six demographic characteristics, including percentages of African American, Hispanic, White, Special Education, Economically Disadvantaged, and Limited English Proficient (LEP) students. From the 36 schools whose principals agreed to participate, 24 (66.7%) of the schools returned usable instruments, which included 6 Exemplary schools, 7 Recognized schools, 6 Academically Acceptable schools, and 5 Academically Unacceptable schools.
INSTRUMENTATION

Two instruments were used in this study. The first instrument, the Middle School Performance Indicator Questionnaire (MSPIQ) was used to measure school performance indicators (National Forum to Accelerate Middle-Graders Reform, 2004). The Likert-type survey consists of 61 performance indicator items forming four dimensions: academic excellence, developmental responsiveness, social equity, and organizational structures. Academic excellence consists of 20 performance indicator items, developmental responsiveness includes 12 performance indicator items, social equity consists of 12 performance indicator items, and organizational structures includes 17 performance indicator items. Respondents select their responses on a four point scale from very frequently to never as follows: Very Frequently (4), Frequently (3), Occasionally (2), or Never (1). Previous research has demonstrated the reliability, construct, and predictive validities of the instrument (National Forum to Accelerate Middle-Grades Reform, 2004; National Middle School Association, 2004). The overall reliability of the instrument in this study was supported with an alpha coefficient of .97.

The second instrument used in this study was the Texas Assessment of Knowledge and Skills (TAKS) test, a state test administered annually to all students in grade 3-11. Student data resulting from the TAKS tests are made available to school districts through the Texas Education Agency’s annual reports. These reports published in the spring of each year include disaggregated test data as well as other demographic information. Student performance, provided for all required TAKS tests at appropriate grade levels, is disaggregated by gender, ethnicity, disadvantaged economic status, special education, and Limited English Proficient (LEP) students.

PROCEDURES

The first instrument, MSPIQ (National Forum, 2004) was mailed together with a self-addressed, stamped envelope to the principals of the 24 participating schools. The principals in each of these schools completed the survey and also distributed them to their faculties. Teachers completed and returned 451 usable surveys for a 79.6% response rate, and their principals completed 24 surveys for a total sample of 475.

The second instrument used in the data collection procedures was the Texas Assessment of Knowledge and Skills (TAKS) test. Student achievement data, resulting from the administration of the state test (TAKS) annually to all students, grades 3-11 in reading, writing, and mathematics is available through the Texas Education Agency’s Academic Excellence Indicator System (AEIS) reports. These reports, containing disaggregated student achievement data, are
used to rate schools and school districts into four categories: Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable.

**Analysis**

To examine differences between Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable schools on each school’s performances indicator dimensions (academic excellence, developmental responsiveness, social equity, and organizational structure), analysis of variance (ANOVA) was the statistical technique used (Research Question One). To examine relationships between the performance indicators and student achievement, multiple regression analysis was employed (Research Question Two). The independent variables considered in the regression equation were the performance indicators. The dependent variables were the disaggregated achievement scores.

**Results**

Separate analyses of variance (ANOVAs) were conducted to examine differences in each of the performance indicator dimensions (academic excellence [ae], developmental responsiveness [dr], social equity [se], and organizational structure [os]) for Exemplary (E), Recognized (R), Academically Acceptable (AA), and Academically Unacceptable (AU) schools. Means and standard deviations for all variables used in the ANOVA are reported in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>ae M</th>
<th>ae SD</th>
<th>dr M</th>
<th>dr SD</th>
<th>se M</th>
<th>se SD</th>
<th>os M</th>
<th>os SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>58.77</td>
<td>1.50</td>
<td>40.31</td>
<td>2.53</td>
<td>32.30</td>
<td>1.11</td>
<td>61.32</td>
<td>1.84</td>
</tr>
<tr>
<td>R</td>
<td>55.77</td>
<td>2.52</td>
<td>37.43</td>
<td>2.03</td>
<td>30.52</td>
<td>1.07</td>
<td>57.31</td>
<td>1.92</td>
</tr>
<tr>
<td>AA</td>
<td>52.56</td>
<td>3.86</td>
<td>36.15</td>
<td>3.36</td>
<td>29.41</td>
<td>2.22</td>
<td>54.48</td>
<td>4.04</td>
</tr>
<tr>
<td>AU</td>
<td>49.34</td>
<td>2.90</td>
<td>33.55</td>
<td>3.34</td>
<td>28.62</td>
<td>1.82</td>
<td>52.85</td>
<td>3.54</td>
</tr>
</tbody>
</table>

E = Exemplary
R = Recognized
AA = Academically Acceptable
AU = Academically Unacceptable
ae = academic excellence
dr = developmental responsiveness
se = social equity
os = organizational structure
The ANOVA results for each of the performance indicator dimensions are reported in Tables 2-5. The first ANOVA indicated a significant difference among Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable middle schools on the performance indicator academic excellence with a large effect size using Cohen's (1988) criteria, F (3, 20) = 11.86, p < .05 n² = .64. The Tukey post hoc analysis revealed a significant difference between Exemplary and Academically Unacceptable middle schools, p < .000 and between Exemplary and Academically Acceptable middle schools, p < .005. The ANOVA data for the performance indicator academic excellence is shown in Table 2.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>n²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>3</td>
<td>280/46</td>
<td>93.49</td>
<td>11.86*</td>
<td>.64</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>157.61</td>
<td>7.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>438.07</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.

The second ANOVA revealed a significant difference among Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable middle schools for the performance indicator developmental responsiveness with a moderate effect size using Cohen’s (1988) criteria, F (3, 20) = 5.52, p < .05, n² = .45. The Tukey post hoc analysis indicated a significant difference between Exemplary and Academically Unacceptable middle schools, p < .004. The ANOVA data for the performance indicator dimension developmental responsiveness is shown in Table 3.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>n²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>3</td>
<td>130.60</td>
<td>43.54</td>
<td>5.52*</td>
<td>.45</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>157.83</td>
<td>7.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>288.43</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.

The third ANOVA indicated a significant difference among Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable mid-
The fourth ANOVA revealed a significant difference among Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable middle schools for the performance indicator dimension organizational structures with a large effect size using Cohen’s criteria, $F(3,20) = 9.12$, $p < .05$, $n^2 = .58$. The Tukey post hoc analysis revealed a significant difference between Exemplary and Academically Unacceptable middle schools, $p < .001$ and between Exemplary and Academically Acceptable middle schools, $p < .003$. The ANOVA data for the performance indicator dimension organizational structures is shown in Table 5.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>$n^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratings</td>
<td>3</td>
<td>234.27</td>
<td>78.09</td>
<td>9.12*</td>
<td>.58</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>171.22</td>
<td>8.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.

The results of the ANOVAs and Tukey post hoc analyses indicated not only significant differences between Exemplary and Academically Unacceptable middle schools for one performance indicator dimension (developmental responsiveness) with a moderate effect size but also significant differences...
between Exemplary and Academically Unacceptable as well as between Exemplary and Academically Acceptable middle schools for three other performance indicator dimensions (academic excellence, social equity, and organizational structures) with moderate to large effect sizes. Table 6 provides a summary of the middle school performance indicators identified in the ANOVA analyses as having the greatest significant differences between those middle schools with Exemplary accountability ratings and those middle schools with Academically Unacceptable accountability ratings.

### Table 6. Performance Indicators with the Greatest Significant Differences between Exemplary (N=6) and Academically Unacceptable Middle Schools (N=5)

<table>
<thead>
<tr>
<th>Dimension Item</th>
<th>Descriptor</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ae 6</td>
<td>The curriculum is rigorous and non-repetitious</td>
<td>.001</td>
</tr>
<tr>
<td>ae 14</td>
<td>Most class time is devoted to learning rather than discipline</td>
<td>.001</td>
</tr>
<tr>
<td>ae 15</td>
<td>Students have the supports they need to meet rigorous standards</td>
<td>.001</td>
</tr>
<tr>
<td>ae 19</td>
<td>Adults collaborate on curriculum and instruction</td>
<td>.001</td>
</tr>
<tr>
<td>dr 24</td>
<td>Teachers use a variety of instructional strategies</td>
<td>.000</td>
</tr>
<tr>
<td>se 33</td>
<td>Staff expect high-quality work from all students and commit to helping all students learn</td>
<td>.000</td>
</tr>
<tr>
<td>se 37</td>
<td>All students have equal access to valued knowledge in all classes</td>
<td>.000</td>
</tr>
<tr>
<td>se 44</td>
<td>The suspension rate is low and in proportion to student demographics</td>
<td>.004</td>
</tr>
<tr>
<td>se 45</td>
<td>There is a shared vision of what constitutes a high-performing school</td>
<td>.000</td>
</tr>
<tr>
<td>os 47</td>
<td>Someone in the school is responsible for coordinating this school-improvement process</td>
<td>.000</td>
</tr>
<tr>
<td>os 48</td>
<td>The school is a community of practice in which learning, experimentation, and reflection are the norm</td>
<td>.000</td>
</tr>
<tr>
<td>os 49</td>
<td>At school, everyone’s job is to learn</td>
<td>.000</td>
</tr>
</tbody>
</table>

ae = academic excellence  
dr = developmental responsiveness  
se = social equity  
os = organizational structures

To determine the relationship between school performance indicators and composite academic achievement for the following disaggregated groups (African American, Hispanic, White, economically disadvantaged, students with a disability, and Limited English Proficient students), stepwise multiple
regression analysis was performed using the disaggregated groups as the unit of analysis in order to construct an appropriate equation regressing student composite achievement against the MSPIQ items. Performance indicator items with correlations above .55 were selected from a Pearson correlation table for entry into the regression analysis.

**African American Students’ Predictors of Composite Academic Achievement.** For African American students, three performance indicator items were strongly related to their composite academic achievement: “Teachers use a *variety of methods* to assess student performance” (academic excellence item 9). “Students revise their work based on *feedback* until they meet or exceed performance standards” (academic excellence item 3). “Curriculum, assessment, and instruction are *aligned*” (academic excellence item 4) (italics added for emphasis): $R^2 = .89$, $F (2, 8) = 33.10$, $p = < .001$.

**Hispanic Students’ Predictors of Composite Academic Achievement.** For Hispanic students, three performance indicator items were strongly related to their composite academic achievement: “All students have equal access to valued *knowledge*” (social equity item 37) “The school is a community of practice in which *learning, experimentation, and reflection* are the norm” (organizational structures item 48). “Students have the *supports* they need in school to meet *rigorous academic standards*” (academic excellence item 15) (italics added): $R^2 = .81$, $F (3, 17) = 23.54$, $p = < .001$.

**White Students’ Predictors of Composite Academic Achievement.** For White students, three performance indicator items were strongly related to their composite academic achievement: “The school creates a personalized environment that *supports* each student’s intellectual, ethical, social, and physical development (developmental responsiveness item 21) “The school delineates benchmarks and insists upon evidence and *results*” (organizational structures 55). “Faculty expect *high quality work* from all students and are committed to helping each student produce it” (social equity item 33) (italics added for emphasis): $R^2 = .83$, $F (2, 16) = 38.66$, $p = < .001$.

**Economically Disadvantaged Students’ Predictors of Composite Academic Achievement.** For economically disadvantaged students three performance indicator items were strongly related to their composite academic achievement: “Faculty expect *high-quality work* from all students and are committed to helping each student produce it” (social equity 33). “Most class time is devoted to *learning content* rather than classroom management and discipline” (social equity item 14). “The school insists on having teachers who promote students’ intellectual, social, emotional, physical, and ethical *growth*” (organizational structures item 59) (italics added for emphasis): $R^2 = .83$, $F(3, 19), 31.76$, $p = < .001$. 
Students with Disability Predictors of Composite Academic Achievement. For students with a disability, two performance indicators were strongly related to their composite academic achievement: “Faculty commitment to high-quality work from all students includes tutoring, mentoring, special adaptations, and other supports” (social equity item 34). “All students have equal access to valued knowledge in all classes” (social equity item 37) (italics added for emphasis): $R^2 = .85$, $F(2, 7) = 19.99$, $p = <.01$.

Limited English Proficient Students’ Predictors of Academic Achievement. For limited English proficient (LEP) students, one performance indicator was moderately related to their composite academic achievement: “All students have equal access to valued knowledge in all classes” (social equity 37) (italics added for emphasis): $R^2 = .66$, $F(1, 7) = 13.34$, $p = <.01$.

The multiple regression analyses are shown in Table 7.

Table 7. Predictions of Composite Academic Achievement for Disaggregated Groups from Middle School Performance Indicator Questionnaire Items

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Independent Variables</th>
<th>R²</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American</td>
<td>1. ae 9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. ae 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. ae 4</td>
<td>.89</td>
<td>33.10</td>
<td>.001</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1. se 37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. as 48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. ae 15</td>
<td>.81</td>
<td>23.54</td>
<td>.001</td>
</tr>
<tr>
<td>White</td>
<td>1. dr 21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. os 55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. se 33</td>
<td>.83</td>
<td>38.6</td>
<td>.001</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>1. se 33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. ae 14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. as 59</td>
<td>.83</td>
<td>31.76</td>
<td>.001</td>
</tr>
<tr>
<td>Special Education</td>
<td>1. se 34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. se 37</td>
<td>.85</td>
<td>19.99</td>
<td>.01</td>
</tr>
<tr>
<td>Limited English Proficient</td>
<td>1. se 37</td>
<td>.66</td>
<td>13.34</td>
<td>.01</td>
</tr>
</tbody>
</table>

LIMITATIONS

As with most studies, this one is not free of limitations. First, all of the data collected from principals and teachers are based on self-reports. Although this is a widely used and generally valid approach for obtaining information, educators may be using differing baselines when they are asked to respond to a questionnaire. The moderately large sample size in the current study helps to offset threats to validity inherent in self-report measures by captur-
ing a more accurate over-all population average rather than relying on a few individual educators’ own point of reference.

Second, selection bias is always a major concern for any study. Although we utilized a random sampling approach, several school districts declined our invitation to participate in the study. Acquiring district and school permissions to contact individual principals and teachers and invite them to complete the survey proved to be challenging.

Lastly, the sample of schools was drawn from a single state. Therefore, results should be generalized with caution, and only to those states that share similar characteristics with the state in this study, including students’ demographics, instructional resources, economic family levels, and community characteristics. Nevertheless, it is believed that a number of intriguing questions have been explored in this inquiry that merit further investigation.

Discussion and Conclusions

The purpose of this study was to examine differences between schools rated Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable on four performance indicator dimensions: academic excellence, developmental responsiveness, social equity, and organizational structures. In addition, relationships between performance indicators and student composite achievement in reading, writing, and mathematics, disaggregated for all students and subgroups of students based on socioeconomic status, race/ethnicity, English language proficiency, and disability were examined.

This work is significant in view of the unprecedented call for secondary school reform (Bill and Melinda Gates Foundation 2008; Center for Collaborative Education, 2004; National Association of Secondary School Principals and The Education Alliance, 2004; National Forum to Accelerate Middle-Grades Reform, 2004; National Middle School Association, 2004; Swaim, 2003; W.K. Kellogg Foundation, 1998) and in light of the need to consider the performance of students from diverse backgrounds (Banks, Gay, Nieto, and Rogoff, 2007; Casteel, 1995; Fry, 2003; Lopez, Gonzalez, and Fierro, 2006; Martin, 2007). The two research questions were tested using data gathered from a sample of 451 middle school teachers and 24 middle school principals, and more than 12,000 middle school students. The statistical methods employed were analysis of variance and multiple regression.

Results of this study suggest that the school performance indicators are meaningful and significantly related to student achievement and school accountability ratings. Differences were found among schools rated Exemplary, Recognized, Academically Acceptable, and Academically Unacceptable on all four dimensions of the MSPIQ: academic excellence, developmental re-
Post hoc analyses revealed significant differences between Exemplary and Academically Unacceptable schools and between Exemplary and Academically Acceptable schools for academic excellence; differences were found between Exemplary and Academically Unacceptable schools for developmental responsiveness; differences were found between Exemplary and Academically Unacceptable schools and between Exemplary and Academically Acceptable schools for social equity; and differences were found between Exemplary and Academically Unacceptable schools as well as differences between Exemplary and Academically Acceptable schools for organizational structures.

The results of our study indicate that teacher behaviors, as well as specific teaching principles and methods, make a difference with regard to student achievement (Ayers, in press; Bulach & Lunenburg, 2008; Emmer & Evertson, 2009; Gage, 2010; Good & Brophy, 2008; Greene, 2008; Nieto, 2009, 2010, in press; Tuckman, 2008). The learning environment in high-performing middle schools challenge all students to use their minds well and meet high academic standards. Curriculum, instruction, and assessment are aligned and provide a coherent vision for learning. The comprehensive services found in high-performing middle schools foster intellectual, ethical, social, and physical development. Teachers in schools that demonstrate these traits employ a variety of instructional strategies with a curriculum that is both socially significant and personally relevant. Norms, structures, and organizational arrangements used to support and sustain excellence are found in high-performing middle schools. In addition, every student is provided with high-quality teachers, resources, learning opportunities, and appropriate supports for learning.

**Implications for Practice**

How can school leaders create high-performing middle schools? We suspect the general way to enhance the academic achievement of middle school students is to focus on the school predictors examined in this study, which were shown to be related to academic achievement. Thus, we briefly consider strategies for enhancing academic excellence, developmental responsiveness, social equity, and organizational structures.

**Academic Excellence.** High-performing schools focus on “authentic” pedagogy (teaching that requires students to think, to develop an in-depth understanding, and to apply academic learning to important realistic problems) and student learning (Newmann & Wehlage, 2010). They achieve this in two ways: greater organizational capacity and greater external support. The most successful schools are those that function as professional learning communities (DuFour, DuFour, & Eaker, 2008; DuFour, DuFour, Eaker, & Karhanek,
That is, they find a way to channel staff and student efforts toward a clear, commonly shared purpose for learning. Moreover, they find that external agencies help schools to focus on student learning and to enhance organizational capacity through three strategies: setting standards for learning of high intellectual quality; providing sustained schoolwide professional development; and using deregulation to increase school autonomy (Blankstein, 2010). In short, dynamic internal learning communities and their relationships with external networks make the difference. Evidence on the critical combination of internal and external learning is mounting (Murphy, 2010).

**Developmental Responsiveness.** There are instructional strategies that can help teachers increase student learning. In research recently completed at the Mid-continent Research for Education and Learning (McREL) Institute, Marzano (2010) identified classroom practices that generally increase student achievement: identifying similarities and differences; summarizing and note taking; receiving reinforcement for effort and recognition for this achievement; doing homework and practicing; using nonlinguistic representations; learning cooperatively; setting objectives and testing hypotheses; and using cues, questions, and advance organizers. Regardless of whether or not teachers teach to standards, these classroom practices work well.

**Social Equity.** In addition to providing professional development to teachers focused on instructional strategies that can improve learning, all schools need an intervention and support system for students who lag behind in learning the curriculum. Schools need to provide additional help to students who lag behind in core subjects, either in school, after school, on weekends, or during the summer months (Fullan, 2010). School leaders need to supply the financial resources to fulfill this mandate. This involves acquiring materials, information, or technology; manipulating schedules or release time to create opportunities for teachers to learn; facilitating professional networks; or creating an environment that supports school improvement efforts (Smylie, 2010).

**Organizational Structures.** School leaders must develop and sustain school structures and cultures that foster individual and group learning (Bullach & Lunenburg, 2008; Lunenburg & Ornstein, 2008). That is, leaders must stimulate an environment in which new information and practices are eagerly incorporated into the system. Teachers are more likely to pursue their group and individual learning when there are supportive conditions in the school and school district, such as particularly effective leadership (English, 2008; Northouse, 2010). Schools where teachers collaborate in discussing issues related to their school improvement efforts are more likely to be able to take advantage of internally and externally generated information (Senge,
Teachers can become willing recipients of research information if they are embedded in a setting where meaningful and sustained interaction with researchers occurs in an egalitarian context (Kruse & Louis, 2009).

**Future Research**

Admittedly, this study was exploratory; much remains to be accomplished. Nevertheless, our results indicate the potential fruitfulness of some significant linkages within schools that may influence student achievement. Although our data were drawn from middle schools, we believe the findings could be applicable to elementary and high schools.

One might question whether the performance indicator dimensions add any value to the earlier research on effective schools (Levine & Lezotte, 1990; Purkey & Smith, 1983; Scheerens & Bosker, 1997), which identified such factors as clear goals, high expectations, positive home-school relations, frequent monitoring, and opportunity to learn as being related to student achievement. Our findings on performance indicator dimensions are consistent with this earlier research, but they go further to explain how some of these factors influence teachers’ beliefs that lead to student achievement. Positive home-school relations will not support achievement unless this involvement builds trust among students, teachers, and parents. Frequent monitoring and opportunity to learn may be effective because it builds a sense of collective efficacy that promotes teacher motivation and persistence. Students, parents, and teachers will probably be more willing to work toward academically challenging goals if they believe they are capable, and the people around them can be trusted to help them. These are testable propositions in need of further research.

More research in a variety of school settings is needed to build a comprehensive theory of performance indicators in schools. For example, in the tradition of the earlier effective schools research, qualitative researchers could conduct comparative case studies of schools identified as having high or low performance indicator dimensions. What would these schools look like? Are there curricular differences between such schools? What are the experiences of students, teachers, and parents in these schools? How are expectations communicated and monitored? What enables or hinders the development of a culture for learning? What is the role of the principal in developing a culture of high academic performance? On the basis of rich descriptions of life in schools, these relationships and other variables could be identified for further quantitative analysis.

The performance indicator dimensions are fruitful because they emphasize the potential of schools to overcome the power of socioeconomic
factors that may hinder student achievement. It is a social psychological construct that is in part related to the positive psychology of Seligman and Csikszentmihalyi (2000), Bulach and Lunenburg’s (2008) research on school cultures, and the process-product research of Gage (2010), Good and Brophy, (2008), Greene (2008), and Nieto (2009, 2010, in press). The performance indicators attempt to explain and enhance what is best in schools to facilitate student achievement. This conclusion should encourage teachers and principals to move their schools forward with confidence, knowing that possible linkages may exist within schools that predict student achievement.

REFERENCES


