

# **PRINCIPALS' LEADERSHIP AND STUDENT ACHIEVEMENT: AN EXAMINATION OF THE TIMSS 1999**

*Milijana Suskavcevic, University of Texas at El Paso, USA  
Sally Blake, University of Texas at El Paso, USA*

## **Abstract**

Pervasive and sustained student learning is more likely to occur in schools with strong instructional leadership. Some researchers, however, question whether the school leadership style makes a measurable difference in promoting school quality. In this study a nationally representative data set is used to examine the association between U.S.A. middle school principals' leadership styles and student achievement on the TIMSS 1999 mathematics and science testing. Pitner's moderated effect model of the leaders' impact on school outcomes served as a theoretical framework, with the construct of "collaboration and cooperation among teachers" as a moderating variable. Statistically significant correlations among principals' leadership style and student performance on the TIMSS 1999 math and science testing were found, with stronger correlations between the two variables for the sample of schools that have school policies supporting teachers' cooperation and collaboration.

## **INTRODUCTION**

Access to quality science education is not equally dispersed across states, school districts, and individual schools in the United States. The Third International Math and Science Benchmarking Study provides evidence that some schools in the United States are among the best performing schools in the world, but that world-class science education is not available to all children. Providing quality science education requires efforts from multiple stakeholder groups, including teachers, science coordinators, and administrators (National Research Council, 1996). Numerous researchers on school effectiveness have demonstrated some form of association between effective schools and the type of leadership practiced by their principals (Hallinger & Leithwood, 1994). However, a number of researchers who have conducted studies with the sole purpose of finding statistically significant relationships between instructional leadership and student achievement have failed

in that attempt (Leitner, 1994). The failure to some extent can be explained by the methodological obstacles in the research in this area (Fidler, 1997). Introduction of the standards-based school reform additionally contributed to the changes in educational context in the U.S., providing a unique opportunity for researchers to reexamine many issues that may play important roles in school effectiveness. The TIMSS 1999 was conducted after a decade of standards-based instruction in the U. S. schools. Data on school background variables and student achievement from this study provides a solid opportunity for exploring the relationship between leadership style and student achievement at the middle school level. It is our intent to determine the extent to which principal leadership behaviors are related to student math and science achievement.

## **REVIEW OF LITERATURE**

The literature relevant to the topic involves studies that explore the relationship between principals' leadership and school outcomes. Included in this section will be a brief overview of the TIMSS 1999. Through the decades of the twentieth century, the role of school leaders in the United States greatly evolved and could generally be characterized as highly transformative. Metaphorically, the dominant role of school principals in the 1930s was one of a scientific manager. In the 1940s the principal was expected to fulfill primarily the role of a democratic leader. In the 1970s the principal was viewed as a humanistic facilitator, and in the 1980s school principals were expected to serve primarily as instructional leaders (Beck & Murphy, 1993). Even though instructional leadership received great popularity and pervaded leadership literature during the 1980s, this notion was introduced a few decades prior to this period. Mackenzie and Corey in 1954 were among the early writers who referred to the school principal as an instructional leader of a school (Greenfield, 1987). De Bevoise (1984, pp.15) used the term to designate the "actions that school principal takes, or delegates to others, to promote growth in student learning."

A number of researchers have developed theoretical frameworks of instructional leadership roles of school principals, contributing to the clearer conceptualizations of the term. The works of Bossert, Dwyer, Rowan, and Lee (1982) may be considered pioneering efforts directed toward a deeper understanding of instructional leadership roles of a school principal. These researchers emphasized that a school principal, through his or her activities, roles, and behaviors in managing school structures does not affect student achievement directly, in the ways the teachers do. However, classroom teaching may be impacted by principals' actions, such as setting and clearly communicating high expectations for all students, supervising teachers' instructional performance, evaluating student progress, and promoting a positive teaching/learning environment. The works of Bossert et al. were expanded by the studies of Murphy, Hallinger, Weil, and Mitman (1983), Hallinger and Murphy (1985a), Sweeney (1982), and Smith and Andrews (1989).

Over the past decades, several comprehensive reviews have been conducted of the literature on school administrators and their roles in schooling. The findings of these reviews will be presented chronologically, not by their importance.

Lipham (1964) and Erickson (1967) reviewed the literature on school administration prior to 1967. Whereas Lipham focused on the findings of the research investigations, Erickson was primarily concerned with methodological issues. The latter highlighted methodological weaknesses of the studies which were published in the professional journals in educational administration during the 1964-1966 period. Both authors concluded that the majority of studies reviewed used questionnaires with suspect validity, which served as a predominant mode of data collection.

Bridges (1982) used 322 research reports on school administrators published during the period of 1967-1980 for his review and focused primarily on methodological issues. Three components of Halpin's (1966) classic paradigm for research on administrative behavior were used by Bridges to organize the body of empirical research on school administrators. The three components are the behavior of the administrator, the antecedent variables influencing such behavior, and outcomes, which at least partially could be attributable to the administrator. In classifying studies with respect to outcomes, Bridges (1982) made a distinction between those studies dealing with the impact that school administrators have on school outcomes and ratings of administrator's effectiveness. In the "administrator's impact" studies, researchers attempted to determine whether administrators made measurable differences in schooling. As observed by Bridges (1982, pp. 21), when assessing the impact of school administrators, "researchers are far more likely to focus on organizational maintenance than organizational achievement." In his words: "Organizational maintenance refers to the extent to which the workforce remains intact as a group" (Bridges, 1982, pp.21), and "may be gauged in terms of morale, cooperation among group members working with one another, and other indices of job satisfaction" (Halpin, 1966, pp.37).

Hallinger and Heck (1996) stated that philosophical and methodological shifts (from positivist, to post-positivist, critical theory, and constructivist) occurring in the last decades in the educational research arena did not have a larger impact on the studies of administrative influences on school outcomes. The research on the relationship between the two was largely examined from a positivist paradigm and with a heavy reliance on quantitative methodology.

As the notion of educational leadership style evolved through the past decades, so did the research of the impact of the school leaders' style on the school. At the time when the idea of instructional leadership became dominant, a number of researchers conducted empirical studies in an attempt to determine if the instructional leadership roles, behaviors, and activities practiced by school leaders may be correlated with school outcomes. The large wave of research on instructional leadership occurred in 1980s and 1990s.

Hallinger and Heck (1996) used Pitner's (1982) framework of administrator effects as criteria for classifying 40 studies on instructional leadership and school outcomes published during the period 1980-1995. All studies reviewed were cross-sectional and non-experimental in nature, meaning that researchers had little or no influence on extraneous variables. Five theoretical approaches identified by Pitner represented conceptual models which served as a means for categorizing existing studies of

administrator effects on school outcomes. These five models were direct-effects, moderated-effects, antecedent-effects, mediated-effects, and reciprocal-effects (Pitner, 1982, pp. 105-108).

Current shifts in the area of principals' leadership, from instructional to transformational leadership, have resulted in a substantial decrease in the number of studies focused on examining the instructional leadership style of school principals. The majority of the empirical studies on instructional leadership and school effectiveness have been conducted in the context of the dominating "loosely coupled" educational system of governance in the U. S. Recent changes related to the implementation of the standards-based reform movement emphasize the instructional leadership aspects of school leaders. Among few quantitative studies in the post-1995 period, one conducted by Louis et al. examined behaviors common to principals of schools with high student achievement. Controlling for pertinent principal and school background characteristics such as race/ethnicity, socioeconomic status, and gender, Louis et al. (1996) found that leaders in high achieving schools "worked effectively to stimulate professional discussion and to create the networks of conversation that tied faculty together around common issues of instruction and teaching" (p. 194).

### **Third International Mathematics and Science Study 1999 (TIMSS 1999)**

The International Association for the Evaluation of Educational Achievement (IEA) initiated a set of international studies to measure student achievement in various subject areas, and provide researchers, evaluators, and policy makers with the resultant information about educational performance. One of the most ambitious projects supported by this organization was TIMSS 1999, in which more than half a million students (third and fourth, seventh and eighth, and twelfth grades) from 38 countries around the world were tested in mathematics and science. Besides being interested in measuring trends in students' achievement, the TIMSS 1999 researchers also collected a wealth of data - background information received from students, mathematics and science teachers, and school principals to help the researchers better understand the context in which teaching/learning of math and science is taking place.

The TIMSS 1999 was a repetition and continuation of the TIMSS 1995. It was a repetition in the sense that one half of the items on which students had been tested were kept the same and the core technical aspects of the study paralleled the 1995 study. It was a continuation in the sense that new testing items were introduced and technical aspects were improved to some extent based on the experiences from the 1995 study. Keeping half of testing items (not released in public) unchanged provided a basis for the longitudinal examination of student achievement. For example, to assess the impact of math and science instruction at the middle school level, students' performance could have been measured at two points in time: in 1995 as fourth graders (pre-middle school scores) and in 1999 as eighth graders (post-middle school scores). The fact that nationally representative samples of students were used in both studies allowed generalizations and, consequently, the above-mentioned comparison.

TIMSS researchers used a two-stage sample data design to select student Population 2. The sampling procedure consists of random sampling of 150 schools, and within each school a random selection of 2 classrooms: one eighth-grade and one seventh-grade math classroom. This type of selection criteria led to a sample of at least 7,500 students. Smaller adjustments by countries were allowed as long as they complied with the prescribed TIMSS standards framework.

The TIMSS 1999 Content Advisory Committee used 308 items (162 mathematics and 146 science) to assess student achievement in these areas. The entire pool of items was distributed among eight Test Achievement Booklets. Each tested student was expected to complete one randomly assigned Test Achievement Booklet in a 90-minute period. Approximately two-thirds of the items that the students received had a multiple-choice format, while the remaining third had a constructed-response format. Eleven content areas from the fields of mathematics and science were used to test math and science comprehension of population 2 (seventh and eighth graders). Six content areas were covered by the mathematics test. These areas and the percentage of items dedicated to each area included fractions and number sense (34%); algebra (18%); geometry (15%); data representation, analysis, and probability (14%); measurement (12%); and proportionality (7%). The eighth-grade science test covered five broader content areas: physics (30%); life science (30%); earth science (16%); chemistry (14%); and environmental issues and nature of science (10%) (Martin et al., 2000).

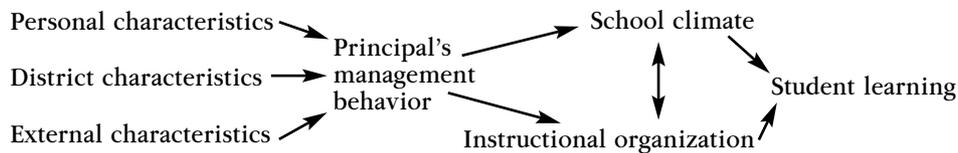
Students' mathematics and science achievement results were summarized using item response theory. This procedure was based on item scoring based on their level of difficulty. In other words, this method provides calculation of test scores by averaging student responses to each item, taking into account the difficulty level of each item (Adams, 1997).

In addition to the Student Achievement Booklets designed to measure students' proficiency in mathematics and science, TIMSS researchers gathered extensive data on student background information, curriculum, instructional practices, school policies, and other aspects of the school environment that directly and indirectly may impact student achievement. The School Background Questionnaire itself contains items about school characteristics, school resources, and school policies related to mathematics and science achievement.

## **THEORETICAL FRAMEWORK**

For the purposes of this research study, a combination of a modified Bossert's (1982) framework and Pitner's moderated effects model (1982) will be utilized. According to Bossert's model, a principal's managerial behavior is shaped by school context (external and district) and the principal's personal characteristics. At the same time, a principal's managerial behavior directly influences school climate and instructional organization, and indirectly school outcomes (student learning and performance). Figure 1 (Bossert, 1982, pp. 38) provides a graphical view of the Bossert's model.

Figure 1. Bossert's model of the principal's influence on student learning



As obvious from the proposed framework, the school principal's managerial behavior (prevalent leadership style) may be at the same time considered a dependent and an independent variable. It plays the dependent variable role in relation to principal's personal characteristics, district characteristics, and external characteristics variables, while it plays an independent variable role when related to school outcomes. It is important to observe the "mono-directional" character of the model, with no shared interaction among variables.

The researchers modified Bossert's framework by leaving out antecedent variables (context and leader's personal characteristics), assuming that these characteristics are already embedded in the leader's dominant leadership style. By leaving out the variables with potential antecedent effects, the nature of the model changes, and researchers look at the indirect effect of leadership style on school outcomes, moderated by the presence of a third variable (school climate). Modification of the Bossert's model leads to the utilization of Pitner's moderated effects model (1982).

Pitner speculated that the presence of a third, moderating variable may influence the relationship between the independent and dependent variable (effects of administrator's leadership style on school outcomes). Normally, researchers theorize that administrator effects would occur under one set of conditions and not under another. Figure 2 depicts Pitner's model of moderated effects of administrator's behavior (Pitner, 1982, pp. 108). It is obvious from the figure that  $a_1$  represents a generic variable indicating leadership style, while  $y_1$  represents a generic variable indicating school outcomes.

Figure 2. Modified graphical representation of Pitner's Moderated Effects Model



The researchers were interested in exploring the relationship between the principal's leadership style and student scores in math and science testing, when the third moderating variable is introduced in the model. The "cooperation and collaboration among teachers in school" is the moderating variable in this model. We speculated that the presence/absence of this variable influences the nature of the relationship between the independent and dependent variable. Only in conjunction with the

presence of the third variable, would a statistically significant relationship be found between predictor and criterion.

## **METHODOLOGY**

This section includes a detailed description of the research design and proposed methodology for examining the relationship between principals' leadership style and student achievement in mathematics and science testing. Additionally, information on the participants in the study, instruments used for answering research questions, procedures, and data analysis will be provided.

The major research question addressed in this study was examination of the relationship between students' scores on math and science testing in the TIMSS 1999 and type of school principals' leadership style for the USA sample. Collaboration and cooperation among school teachers was introduced as a grouping variable. A paragraph on procedures will explain in more detail the way in which all variables involved in the study were operationalized.

The primary researchers of the TIMSS-R database used a two-stage stratified sample design to select schools and student samples for the study. First, a random sample of schools was selected, followed by a random sample of two classes from each school. Two hundred and forty schools were randomly selected and within each school two classrooms of eighth graders were randomly selected, totaling 9, 072 eighth graders and comprising a nationally representative sample. The student subjects tested consisted of 50.9% males and 49.1% females. The gender of the 240 middle school principals selected was not included in the School Background Questionnaire (Gonzales & Miles, 2001).

Instruments used are the School Background Questionnaire, completed by principals from the U.S. schools participating in the TIMSS 1999 study, and students' Achievement Tests in Science and Mathematics (8 booklets containing mixed contents in mathematics and science, one of which is taken by each participating student).

Data from the School Background Questionnaire for this particular study involve items BCBGACO1 through BCBGAC13 that address the issue of the number of hours per month spent by principal on instructional issues (teaching, hiring teachers, giving demonstration lessons, discussing educational objectives with teachers, initiating curriculum revision and/or planning, training teachers, and providing professional development activities), and the number of hours per month spent on non-instructional issues (representing the school in the community, representing the school in official meetings, internal administrative tasks, talking with parents, counseling and disciplining students, and responding to education officials' requests). Another three items, BCBGCOL1 through BCBGCOL3 from the principal's instrument, address the issue of collaboration and cooperation among teachers in the school, and will be used for the purposes of categorizing schools for statistical analyses.

The researchers were interested in students' overall achievement in mathematics

and science. Therefore, no particular items from the students' Achievement Test Booklets were selected for the analysis. The total score, representing students' overall performance in these subjects was used. The TIMSS Subject Matter Advisory Committee ensured that the content of the tests reflected contemporary thinking and priorities in the areas of math and science. Students were tested on the concepts that were covered in their classes prior to testing. For Population 2 (eighth graders) the math tests included six areas: fractions and number sense, proportionality, measurement, data representation, analysis and probability, and geometry and algebra. Science pieces in the test booklets covered topics from earth science, life science, chemistry, physics, and environmental science.

The Achievement Test items that measured students' proficiency in the aforementioned content areas in math and science consisted of both multiple choice questions that were easily quantified and free response questions (at about one-third of the total set), for which a strict scoring rubric was developed and consistently applied. For the free-response test items, a random sample of students' responses was reviewed and graded by two experts, and their grading results were compared for the purposes of establishing trustworthy inter-rater reliability. It is important to stress that scientists and mathematicians developed all test items. Also, for the purposes of maximizing students' exposure to content, but minimizing the workload for each participant, a multiple matrix sample design (Adams & Gonzales, 1996) was used to optimize the testing process. This particular design implies the use of a process where each student responds only to a subset of items randomly selected from the total pool of items. This procedure leads to the use of plausible values for student scores, or simply stated, a prediction can be made of how well a student would perform on the entire pool of items, based on the student performance on the subset of items. Item-response theory was used to determine the final set of math and science questions on which students would be tested.

The National Center for Education Statistics team of researchers for the Third International Math and Science Study 1999 was responsible for administering the School Background Questionnaires to the principals from the participating schools. Participation in this study was entirely voluntary. Schools that rejected participation were excluded and replacement schools were provided through random selection. The history of the selection process was carefully documented. The National Center for Education Statistics (NCES - U.S. Department of Education) provided the raw data for these researchers' use via the TIMSS 1999 database stored on the CD-ROM. The CD-ROM contains all the information necessary for conducting secondary data analysis.

## **DATA ANALYSIS**

The researchers used the statistical program, Statistical Package for the Social Sciences PC (SPSS-PC) version 10.7, to conduct statistical analyses required to answer proposed research questions. The SPSS does not provide syntax for correct calculation of standard errors for various quantitative methods for the complex sample data design - the one employed by the primary TIMSS 1999 researchers. The

exception is calculation of the product moment in the correlation data analysis, which represents the ratio of covariance (between predictor and criterion) and the product of their standard deviations. Design effect is a function of both physical quantities, and its effect is canceled for the calculation of the correlation coefficient. Therefore, the correlation coefficient does not require additional adjustments, and has the same value as if it were calculated for the sample random sampling data design. For correct computation of means, percentages, and their standard errors, the use of macros is required (Gonzales & Miles, 2001). The use of currently available software, developed by the NCES related research groups (WesVar and AM), also allows correct computation of standard errors for numerous methods of quantitative methodology.

Students' scores on math and science testing and principals' instructional leadership for the USA sample operationalized through the number of hours dedicated to instructional leadership activities were examined and reported together with the measures of skewness and kurtosis. The normality checks of data proved that coefficients of skewness and kurtosis were outside normal limits, violating the normality assumption, therefore non-parametric correlation (Spearman's rho) was used to answer the research question.

The instructional leadership construct was formed as a composite score of all School Background Questionnaire items that address principals' instructional activities (teaching, giving demonstration lessons, discussing educational objectives with teachers, initiating curriculum revision and/or planning, hiring teachers, training teachers, and providing professional development activities). Similarly, a non-instructional leadership construct represented a composite score of all the questionnaire items that are non-instructional in nature (representing the school in the community, representing the school in official meetings, internal administrative tasks, talking with parents, counseling and disciplining students, and responding to education officials' requests).

As a moderating variable in this analysis the researchers used the construct "collaboration and cooperation among teachers." Three School Background Questionnaire items (school policies that promote collaboration and cooperation, discussing and sharing instructional materials and objectives among teachers, and teachers' meetings on a regular basis to promote instructional issues) represented as dichotomous variables were recoded into the "collaboration and cooperation among teachers" construct by using composite scores. Because all three variables were categorical, with values "1" indicating the answer "no" and "2" indicating the answer "yes", the total possible score for the construct will range from 3 to 6. The researchers imposed the following condition: if the total score is  $\leq 4$ , the school will be considered as one that does not promote collaboration and cooperation among the teachers, otherwise (if the total score is  $> 4$ ), the school will be considered as one that does promote collaboration and cooperation among the teachers.

Even though the primary researchers of the TIMSS 1999 database proposed five math and science achievement plausible values, with no particular preference toward the

use of any of these values, numerous analyses have been conducted using only the first plausible value. Justification for this can be found in the researchers' explanation that "the imputation error can be ignored" (Gonzales & Smith, 1997, ch. 6, p. 3), and the conclusions were reached upon conducting inter-correlations among the five plausible scores. Although any of the five plausible values equally well represent student scores in mathematics and science, the researchers used the mean of the five plausible scores in mathematics and the mean of the five plausible scores in science, as measures of student achievement in these areas. Additional averaging of students' scores in math and in science per school was required to appropriately match these scores with school level data (leadership style). Statistical significance for all statistical analyses was set at .01. Checks for effect sizes were made (Cohen, 1988).

## FINDINGS

Prior to describing the findings of the correlation analyses, a descriptive summary of the measures involved in the analysis will be provided. Descriptive information about the number of hours per month dedicated to instructional and non-instructional activities by the sample of middle school principals from the United States and descriptive information about the eighth grade students' average scores on the TIMSS 1999 mathematics and science tests averaged by school are displayed in Table 1.

*Table 1: Means, Standard Deviations, and Ranges for Leadership Activities and Students Scores for the Middle Schools from the U.S. Sample of Schools that Promote Collaboration and Cooperation of Teachers*

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Number of hours per month on instructional activities	2450	38.81	23.98	4	123
Number of hours per month on non-instructional activities	2412	102.67	42.77	23	212
Students' mathematics score per school	2636	500.67	48.12	400.59	666.42
Students' average science score per school	2636	514.89	52.42	383.64	634.49

Descriptive information about the number of hours per month dedicated to instructional and non-instructional activities as reported by the middle school principals and eighth grade students' average scores on the TIMSS 1999 mathematics and science tests from the schools that promote collaboration and cooperation among teachers is provided in Table 2.

*Table 2. Means, Standard Deviations, and Ranges for Leadership Activities and Students Scores for the Middle Schools from the U.S. Sample that Do Not Promote Collaboration and Cooperation of Teachers*

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Number of hours per month on instructional activities	4018	40.79	23.43	0	121
Number of hours per month on non-instructional activities	3907	102.80	45.84	26	199
Students' mathematics score per school	4725	494.41	51.23	347.32	625.08
Students' average science score per school	4725	503.13	57.84	359.80	621.54

Findings of the correlation analysis between students' scores on math and science testing and principals' leadership style - instructional and non-instructional, for both types of schools - those that have written policies that encourage and support collaboration and cooperation among teachers and those that do not have these written policies in place, are summarized in the correlation tables 3 and 4, respectively.

*Table 3: Correlation Matrix of Leadership Styles and Student Scores on Math and Science for the Schools that Do Not Promote Collaboration and Cooperation among Teachers*

	<i>Non-instructional leadership</i>	<i>Mathematics</i>	<i>Science</i>
Instructional Leadership	( $\rho = .299$ ) ( <i>n</i> = 3819)	( $\rho = .083$ ) ( <i>n</i> = 4018)	( $\rho = .103$ ) ( <i>n</i> = 4018)
Non-instructional leadership		( $\rho = .080$ ) ( <i>n</i> = 3907)	( $\rho = .092$ ) ( <i>n</i> = 3907)
Mathematics			( $\rho = .931$ ) ( <i>n</i> = 4725)

*Table 4: Correlation Matrix of Leadership Styles and Student Scores on Math and Science for the Schools that Promote Collaboration and Cooperation among Teachers*

	<i>Non-instructional leadership</i>	<i>Mathematics</i>	<i>Science</i>
Instructional Leadership	( $\rho = .259$ ) ( <i>n</i> = 2412)	( $\rho = .156$ ) ( <i>n</i> = 2450)	( $\rho = .147$ ) ( <i>n</i> = 2450)
Non-instructional leadership		( $\rho = .116$ ) ( <i>n</i> = 2412)	( $\rho = .196$ ) ( <i>n</i> = 2412)
Mathematics			( $\rho = .911$ ) ( <i>n</i> = 2636)

## CONCLUSIONS AND RECOMMENDATIONS

Two important questions discussed in the previous sections emerged after performing the data analysis: 1) How is the school principal's dominant leadership style correlated with student achievement? and 2) What is the role of the school policies that promote collaboration and cooperation among teachers in the examination of principal's leadership style and student achievement?

No significant differences were found in the strength of relationship between either non-instructional leadership and student scores on math and science tests or the instructional leadership and student scores on the same tests. The specific activities performed by the school principal that were associated with both instructional and non-instructional leadership styles have been demonstrated to be weakly but positively correlated with students' scores on the math and science tests. This finding supports the conclusions of previously conducted studies, which indicate that variables associated with the leadership style of school principals account but for a small portion of student performance.

Due to the stronger relationship between both instructional and non-instructional leadership and student scores present in the set of schools that promote collaboration and cooperation among teachers, school principals may consider having a set of written policies that encourage meetings among the teaching staff on a regular basis - by grade level, by subject, and by other grouping criteria. Such an effort may represent another important contributor to a school climate that is conducive and supportive of student learning.

The major purpose of this study was to examine the relationship between principals' leadership style and student achievement in the era of standard-based instruction and the most recent systemic reform in education in the USA. This same issue has been examined in the previous decades through a number of empirical studies, but considering the changing milieu of the field of education, in the researchers' view, it was important to reexamine the relationship between the two constructs in the new context. This new context imposes higher expectations on school principals, who are contemplated to serve primarily as instructional leaders. The findings of this study did not provide evidence of the superiority of instructional leadership over the non-instructional, considering that the strength of the relationships between instructional leadership and student performance and non-instructional leadership and student achievement displayed similar trends.

## References

- Adams, R. J., Wu, M. L., & Macaskill, G. (1997). Scaling methodology and procedures for the mathematics and science scales. In M. O. Martin and D. L. Kelly (Eds.) *Third International Mathematics and Science Study Technical Report*, Volume 2, Chestnut Hill, MA: Boston College.
- Beck, L.G. and Murphy, J. (1993). *Understanding the Principalship: Metaphorical Themes 1920s-1990s*. Teachers College, Columbia University New York and London

- Bossert, S. T. (1988). School effects. In N. J. Boyan (Ed.), *Handbook of research on educational administration* (pp. 341-352). New York: Longman.
- Bossert, S., Dwyer, D., Rowan, B., & Lee G. (1982). The instructional management role of the principal. *Educational Administration Quarterly*, 18 (3), 34-64.
- Bridges, E. M. (1982). Research on the school administrator: The state of art, 1967-1980. *Educational Administration Quarterly*, 18 (3), 12-33.
- Erickson, D. A. (1967). The school administrator. *Review of Educational Research*, 37 (4), 417-432.
- Fidler, B. (1997). School leadership: Some key ideas. *School Leadership and Management*, 17 (1), 23-38.
- Gonzales, E. J., & Miles, J. A. (2001). *TIMSS 1999 User guide for the international database*, International Study Center, Lynch School of Education, Boston College.
- Gonzales, E. J., & Smith, T. A. (1997). *Users guide for the TIMSS international database*, Chestnut Hill, MA: TIMSS International Study Center.
- Hallinger, P., & Heck, R. H. (1996a). The principal's role in school effectiveness: An assessment of methodological progress, 1980-1995. In K. Leithwood, J. Chapman, D. Corson, P. Hallinger, and A. Hart (Eds.), *International Handbook of Educational Leadership and Administration* (pp. 723-783). Dordrecht, Netherlands: Kluwer.
- Hallinger, P., & Heck, R. H. (1996b). Reassessing the principal role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly*, 32 (10), 5-44.
- Hallinger, P., & Leithwood, K. (1994). Introduction: Exploring the impact of principal leadership. *School Effectiveness and School Improvement*, 5, 206-218.
- Hallinger, P., & Murphy, J. (1985a). Assessing the instructional management behavior of principals. *Elementary School Journal*, 86 (2), 217-247.
- Halpin, A. W. (1966). A paradigm for research on administrator behavior. In *Theory and Research in Administration* (Halpin A. W., Ed.) New York: The Macmillan Company, pp. 22-77.
- Leitner, D. (1994). Do principals affect student outcome: An organizational perspective. *School Effectiveness and School Improvement*, 5, 219-238.
- Lipham, J. M. (1964). Organizational character of education: Administrative behavior. *Review of Educational Research*, 34 (4), 435-454.
- Louis, K. S., Kruse S. D., & Marks, H. M. (1996). Schoolwide Professional Community. In Fred M. Newmann and Associates (Eds.) *Authentic Achievement: Restructuring Schools for Intellectual Quality*, San Francisco: Jossey-Bass.
- Martin, M. O., Mullis, I. V. S., Gregory, K. D., Hoyle, C., & Shen C. (2000). *Effective schools in science and mathematics*. TIMSS International Study Center, Boston College, Chestnut Hill, MA.

- Murphy, J., Hallinger, P., Weil, M., & Mitman, A. (1983). Instructional leadership: A conceptual framework. *Planning and Changing*, 14 (3), 137-149.
- Pitner, N. (1982). The study of administrators effect and effectiveness. In N. Boyan (Ed.), *Handbook of research in educational administration* (pp. 99-122). New York: Longman.
- Smith W. F., & Andrews, R. L. (1989). Instructional leadership: How principals make difference. *Association for Supervision and Curriculum Development, Alexandria, Virginia*.
- Sweeney, J. (1982). Research synthesis on effective school leadership. *Educational Leadership*, 39 (5), 346-352.