Using Assessment Data to Guide School and Classroom Decision Making: An Examination of the Effects on Student Achievement

Laura Lang

Shana Goldwyn

Christopher Schatschneider

Alysia Roehrig-Bice

Christine Johnson

Florida State University

Paper presented at the annual meeting of the

American Educational Research Association

Chicago, IL.

April 2007

Portions of this research were funded by the Multi-University Reading, Mathematics and Science Initiative (MURMSI) from a grant awarded by the U.S. Department of Education to the Learning Systems Institute, Office of the Provost, Florida State University (FY 04 award number U215K040242). Portions of this research were also funded by NCER grant number R305B04074.

Using Assessment Data to Guide School and Classroom Decision Making: An Examination of the Effects on Student Achievement

While somewhat counterintuitive, recent studies indicate that top down, standards-based accountability policies aimed at increased student achievement have increased the influence of principals, teachers and other local actors on teaching and learning (Marks & Nance, 2007). This nascent influence may be the result of increased requirements of federal and state accountability systems for educators to provide evidence that the decisions they make result in improved student achievement (Institute for Educational Leadership, 2000; Johnson, 2002; Massell, 2000). However, there is a paucity of evidence regarding the extent to which data-driven decision-making effects student performance. The purpose of this study was to examine the impact of using student assessment data to guide school and classroom decision-making, and its effect on student performance in reading and mathematics based on norm- and criterion-referenced outcome measures.

Research related to the relationship between school leadership and improved student performance indicates that principals influence student achievement through the creation of teaching and learning environments that foster achievement and by focusing improvement efforts on the practices most likely to result in student learning gains (Brenninkmeyer & Spillane, 2004; Crow, Hausman, & Scribner, 2002; Hallinger, Bickman, & Davis, 1996; Hallinger & Heck, 1996; Liberman, Falk, & Alexander, 1994; Waters, Marzano, & McNulty, 2003; Witziers, Bosker, & Kruger, 2003). In a meta-analysis examining the effects of leadership on student performance, Waters, Marzano & McNulty (2003) identified 21 leadership responsibilities and practices that were correlated with student achievement. Of these, several were associated with the school leaders role in curriculum, instruction and assessment more generally, and in

monitoring students' progress more specifically. Although methodological issues undermined the causal claims made by Waters, McNulty and Marzano (2003), a number of recent studies corroborate the relationships between monitoring student data and gains in student achievement (Crawford & Torgesen, 2006; Oberman & Symonds, 2005; Reeves, 2006). Despite this evidence, researchers report high levels of variability in the use of assessment results across schools (Stecher & Hamilton, 2006) due, at least in part, to the difficulties educational leaders face in transforming data into tools for improved teaching and learning (Massell, 2000; Chen, Salahuddin, Horsch & Wagner, 2000; Hassler, Buck & Torgesen, 2004; Love, 2004).

Using student performance data effectively involves gathering the results of various assessments, including screening, progress-monitoring, diagnostic and outcome measures, to determine whether a class, school, or district is achieving its various purposes. When student data are disaggregated by groups, domains and/or skills, principals and teachers are better prepared to select appropriate, high quality educational interventions to meet specific goals. With regard to the most central goal of schools—student learning—the call for data-driven decision making is based on the assumption that collecting and analyzing data to inform instructional decisions helps to elucidate the causal link between particular practices and student performance. As a result, the most successful educational interventions can be utilized to produce optimal student achievement (Miller, 2000). Drawing on systemic reform theory and knowledge management theory, Mason (2003) provides the following definition for effective data use: "Learning from data means to transform data into information as it is interpreted in context. Data then becomes knowledge as it is shared, applied, and used to promote change and improvement throughout an organization" (pp. 8-9).

While a number of variables influence the extent to which student performance data are used effectively in schools and classrooms (Brookhart, 2004; Ingram, Louis & Schroeder, 2004), the results of several studies indicate that principals and teachers lack essential knowledge and skills (Brookhart, 2001; Hassler, Ogawa, Ossont, Nahmias, Roehrig, Johnson & Diefenbach, 2006; Impara, Flake & Pager,1993; Stecher & Hamilton, 2006). A longitudinal three-phase randomized field trial, Applied Data Analysis for Principals and Teachers (ADAPT), was employed to explore this assertion systematically (Hassler et al., 2006). The ADAPT study was designed to (1) determine educators' baseline knowledge, skills and attitudes regarding the use of student assessment data to guide decision making; (2) ascertain the extent to which an online 3 credit hour semester-long graduate course would result in differences in educators' knowledge, skills and attitudes regarding the use of student data; and (3) to examine the extent to which a causal link exists between the use of student performance data and student achievement. Hassler et al. (2006) found that middle school principals, reading coaches, reading teachers and language arts teachers across the state of Florida lacked essential knowledge and skills required to use data effectively (ADAPT Phase I). Further, pretest results indicated there were no significant differences in baseline knowledge and skills among the three groups participating in the study: principals, reading coaches and teachers. Following the successful completion of an online course in applied data analysis during the fall semester in 2005, principals, reading coaches and teachers in the treatment group scored significantly higher than the control group on the posttest of knowledge and skills related to using student assessment data to guide decision making (ADAPT Phase II).

The focus of this paper is on the third phase of this project (ADAPT Phase III). ADAPT Phase III is a follow up analysis of the differences in student outcomes for treatment versus

control principals and teachers following the completion of the intervention (the aforementioned online course in data analysis). The course was a 16-week course, beginning in the fall of 2004, with expected completion in late fall of 2004. However, due to several hurricanes that impacted schools around the state of Florida, principals and teachers in three of the twelve treatment schools began the course several weeks late, and were, therefore, delayed in completing the course. The analyses focused on gains in reading and math performance on Florida Comprehensive Assessment Test-CRT (FCAT-CRT) and in reading and math performance on the Florida Comprehensive Assessment Test-NRT (FCAT-NRT), both administered statewide in grades 3-10. The full extent of ADAPT Phase III includes analysis of student performance gains for the 2005-2006 school year, this paper focuses on preliminary analyses of student gains during the 2004-2005 school year, the year of the intervention implementation.

Methodology

Research Questions

While ADAPT Phase I and ADAPT Phase II focus on the acquisition and application of knowledge, skills and attitudes, ADAPT Phase III takes the analysis one step further, making the connection between educational knowledge and practice regarding the use of assessment results and student performance. The following questions were addressed:

1. What are the effects of the treatment on student performance in reading?
2. What are the effects of the treatment on student performance in mathematics?

Participants

The participants in this study were from 25 Florida middle schools. Schools responded to an open invitation for participation, making participation of schools voluntary. As schools volunteered, they were required to meet specific criteria, including the participation of at least one administrator, one reading coach, and three teachers from any grade who work with students on a regular basis. The schools varied in demographics, with ethnic diversity of students ranging from 3% Black and/or Hispanic, to 80% Hispanic, to 84% Black, and a range of 20%-78% of students receiving free and reduced price lunch. In general, the student performance of schools was 30% or more performing below state standards, with the exception of one school (Lang et. al, 2006).

To assign schools to treatment and control groups, the schools were matched using a combined z-score of the percentage of students receiving free and reduced price lunch and the percentage of students performing below state standards, forming 13 pairs. Within each pair, one school was randomly assigned to the treatment group and one to the control group. One school drooped out before the study began, resulting in 12 treatment schools and 13 control schools (Lang et. al, 2006).

Data Sources and Analyses

Students' reading and mathematics scores on the Florida Comprehensive Assessment Test-CRT (FCAT-CRT), a criterion measure of student performance on the state standards, and on the Florida Comprehensive Assessment Test-NRT (FCAT-NRT), a norm-referenced measure of achievement, were used in the analyses for grades 6-8. Gain scores were obtained by comparing a two-year average (2003 and 2004) to the 2005 scores in reading and mathematics

on both measures. While ADAPT Phase III involves the analysis of student performance gains for 2005 and 2006, the following describes the results for 2005 only.

These analyses were performed through a series of hierarchical linear models (HLM; Bryk & Raudenbush, 2002) in which we estimated mean differences among treatment and control participants on student performance in both reading and mathematics on both measures. HLM was the appropriate statistical technique to use, as it is likely that variability in student performance would be influenced by teacher level variables. Additionally, all of the teachers at the treatment schools did not participate in the online course. Using HLM accounts for the effects of these teacher level differences in knowledge.

In these analyses, we considered the nested structure of a school, performing the analyses as students nested within classrooms. Follow-up analyses that inspected the significant effects in specific areas were also conducted using hierarchical linear models, again looking at students nested within classrooms. Analyses were also conducted considering the nested structure of classrooms within schools, however, because all schools assigned to a treatment included principals, reading coaches and teachers assigned to treatment as well, a 2-level nested approach was appropriate to measure and compare student gains. Since there was no way to determine how students matched up with reading coaches, the analyses were limited to teachers nested in schools, and reading coaches were excluded from the analyses. However, since it was a 2-level model, the reading coaches who participated, would be considered in the school level assignment to treatment or control. Follow up analyses were also conducted to look at the contrast estimates between the treatment and control group as they were separated by grade.

Results

The analyses revealed significant interaction effects for the treatment group in student performance on FCAT-CRT in mathematics, F(2, 3067.321)=40.973 , p=<.001, FCAT-NRT in mathematics F(2, 3054.71)=51.756, p=<.01, and the FCAT-NRT on reading, F(2, 2433.72)=135.916 , p=<.001. Additionally, the effects were not significant in FCAT-CRT in reading , F(2, 2575.574)=.870 , p=.419. Although follow-up analyses revealed that the effect of the treatment was not uniform across participants, there were achievement gains at all three grade levels, supporting a finding that the effects could be attributed to the intervention. Table 1.1 highlights the estimates of fixed effects for each FCAT test, with the treatment group for 8th grade as the reference group. Table 1.2 highlights the estimated marginal means by treatment and control groups.

Table 1.1

	FCAT-CRT Reading		FCAT-NRT Reading		FCAT-CRT Math		FCAT-NRT Math	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Intercept (Grade 8 Tx)	77.39*	3.43	102.68*	7.67	81.05*	3.32	115.96*	6.36
2004 Score	.74*	.1	.87*	.011	.77*	.01	.86*	.01
Group	.26	2.5	.28	1.71	1.25	2.76	3.35	1.99
Grade 6	24.1	22.18	17.68	16.15	-13.51	20.39	-3.29	14.14
Grade 7	1.14	1.79	-14.69	1.3	-6.32*	1.67	-4.91*	1.167
Group*Grade 6	-16.42	24.14	-36.78*	17.55	8.02	22.18	-34.14*	15.37
Group*Grade 7	-2.41	2.39	.59	1.74	-7.44*	2.23	-5.62*	1.55

*p<.05

Table 1.2

	Treatment		Control	
FCAT Test	EM Means	Std. Error	EM Means	Std. Error
FCAT-CRT Reading	309.816	3.496	303.805	7.543
FCAT-NRT Reading	691.829	5.476	680.052	2.511
FCAT-CRT Math	309.645	7.019	311.090	3.384
FCAT-NRT Math*	690.958	4.882	681.058	2.378

These results can be interpreted as the treatment, an online course in applied data analysis, having an effect on student performance in mathematics, but not as much effect on student performance in reading. In that the analyses revealed a significant interaction in both the FCAT-NRT in mathematics and FCAT-CRT in mathematics, post-hoc comparisons were made by each of the grade levels to obtain contrast estimates between the treatment group and the control group (Table 2.1).

Table 2.1

	FCAT-CRT Mathematics		FCAT-NRT Mathematics	
	Estimate	Std. Error	Estimate	Std. Error
Grade 6	819.758	22.27	30.79*	15.45
Grade 7	6.186*	2.94	2.26	2.11
Grade 8	-1.25	2.76	-3.35	1.99

*p<.05

These results can be interpreted as the significant differences in grade occurring during grade 7 for the FCAT-CRT in mathematics and during grade 6 for the FCAT-NRT in mathematics. Therefore, we can assume that the intervention, an on-line course in applied data analysis for principals and reading teachers, was most effective for sixth and seventh grade

students in the area of mathematics. Table 2.2 highlights the estimated marginal means of students to look at these difference.

Table 2.2

	FCAT CRT-Math	FCAT NRT-Math
C - Grade 6	312.012	659.615
C - Grade 7	303.747	686.517
C - Grade 8	317.511	697.042
T- Grade 6	302.743	690.403
T- Grade 7	309.934	688.781
T- Grade 8	316.258	693.69

Additionally, these follow-up analyses revealed that the effect of the treatment was not uniform across participants. However, there were trends in the increases across grade levels, supporting a finding that the effects could be attributed to the intervention. Figures 2.1 and 2.2 highlight the estimated marginal means by grade level by treatment and control.

Figure 2.1

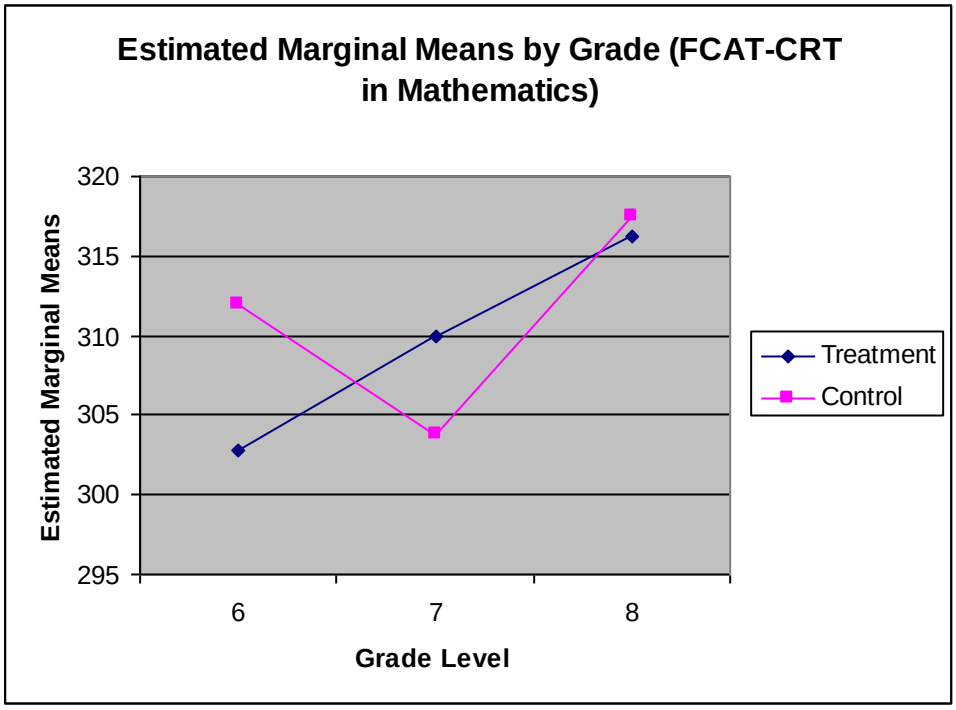
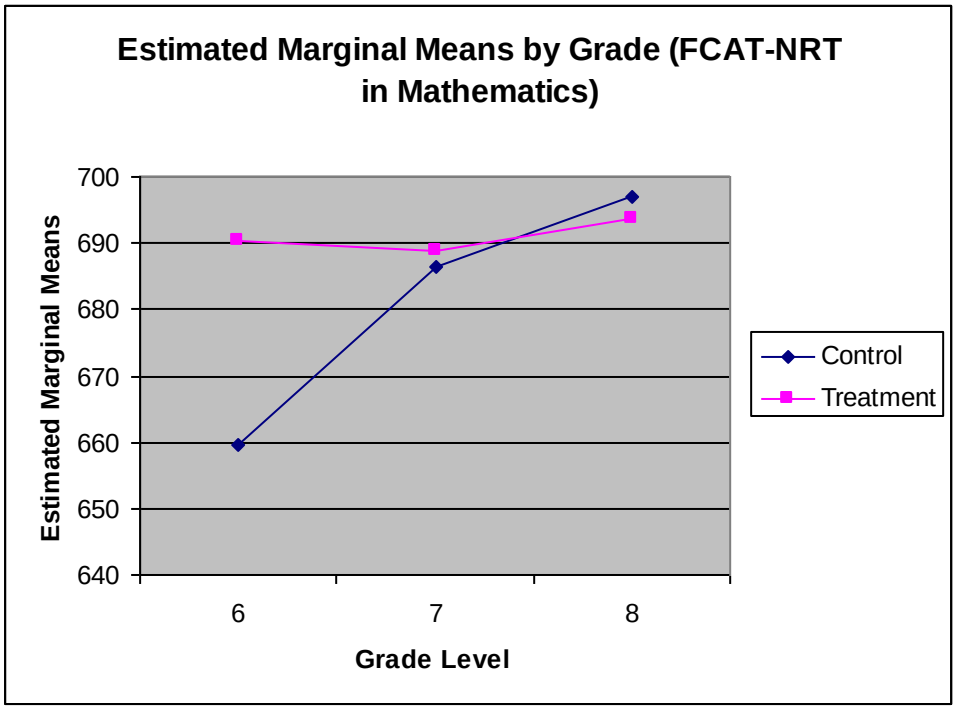


Figure 2.2



It is important to note that there is an extremely high level of variability within the sixth grade level. This means that at the sixth grade level, students perform very differently, which may influence the effect of the treatment on student performance.

Conclusions

These analyses were conducted to gain a preliminary understanding of the effect of the intervention on student performance in reading and mathematics. They were conducted comparing student assessment data collected during March 2005, the same school year treatment participants completed the online course (intervention), to data collected during March 2004. Given the progressive nature of the intervention, changes in practices resulting from participation in the online course were not fully infused in classrooms nor schools, and, minimal differences, if any, were expected in comparing student performance outcomes for treatment versus control participants. Any differences resulting from the intervention would be more likely to be greater the following year (2006) when principals, reading coaches and teachers began the year equipped with the necessary knowledge and skills to use student performance data to drive school and classroom decision-making.

Earlier phases of the ADAPT research indicated that (1) middle school principals, reading coaches and teachers lacked essential knowledge and skills to use data to drive school and classroom decision making and (2) following successful completion of an online course in applied data analysis, there were significant differences in related knowledge and skills for the treatment group when compared to the control group. Preliminary evidence from ADAPT Phase III suggests a causal link between the use of student performance data and student achievement in mathematics.

Discussion

Despite the preliminary nature of the initial analyses, the findings provide insights about the manner in which accountability policies may be increasing the influence of principals and teachers on teaching and learning. The also have implications for practice, provide insights regarding the nature of using data to inform instruction and indicate areas for further study. The significant results found in the area of mathematics suggest that as middle school leaders knowledge, skills and use of data increases, the most immediate effect is on student performance in mathematics. This was an unexpected preliminary finding given that the content areas of the participating teachers were reading and language arts and that the focus during the on-line applied data analysis course was on reading assessment. However, given that the content area preparation of middle school mathematics teachers includes training in data analysis, their baseline knowledge and skills would be expected to be higher than those of reading teachers and language arts teachers. As school leaders increased their use of student performance data in discussions with teachers, the mathematics teachers may have been better poised to act on the information to improve instruction. Further, while the effects on reading outcomes were not statistically significant, the upward trends in the reading data may have influenced performance in mathematics.

These preliminary findings also suggest that principals and teachers must both use student performance data to drive decision-making in order for its impact on student performance to be significant. The analyses of year two student performance data should examine the relationship between principal and teacher knowledge in terms of its impact on student performance. Additionally, it is important to note that the most significant results occurred at the sixth and seventh grade levels. Longitudinal analyses could answer questions

regarding the extent to which sustained knowledge of principals, teachers and reading coaches is critical to students' success at different grade levels. Over time, it will be important to examine these differences and the extent to which they are maintained. Overall, the results substantiate earlier findings (Massell, 2000; Chen, Salahuddin, Horsch & Wagner, 2000; Hassler, Buck & Torgesen, 2004; Love, 2004) indicating that educators' lack of knowledge and skill related to student assessment and applied data analysis impedes their ability to use the results of both progress monitoring and outcome measures to make school and instructional decisions. They also suggest that an on-line course is effective in providing the required knowledge and skills; and that using data to guide decision-making may improve student performance in some subject areas and grade levels. The final analyses, comparing March 2005 assessment results to March 2006 assessment results should provide further insights regarding the practice and its impact on teaching and learning.

References

Brenninkmeyer, L. D., & Spillane, J. P. (2004). Instructional leadership: How expertise and subject matter influence problem solving strategy. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Brookhart, S. M. (2001). The "standards" and classroom assessment research. Paper presented at the annual meeting of the American Association of Colleges of Teacher Education, Dallas, TX (ERIC Document Reproduction Service No. ED451189).

Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, *106*(3) 429-458.

Chen, J., Salahuddin, R., Horsch, P., & Wagner, S. L. (2000). Turning standardized test scores into a tool for improving teaching and learning: An assessment-based approach. *Urban Education*, *35*(3) 356-384.

Crawford, E. C., & Torgesen, J. K. (2006, July). Teaching all children to read: Practices from Reading First schools with strong intervention outcomes. Paper presented at Just Read, Florida! 5th Annual K-12 Leadership Conference, Orlando, FL. Retrieved July 20, 2006, from http://www.fcrr.org/Interventions/pdf/teachingAllStudentsToReadComplete.pdf.

Creighton, T. B., & Jones, G. D. (2001). Selection or self-selection: How rigorous are our selection criteria for education administration preparation programs? Paper presented at the 2001 conference of the National Council of Professors of Educational Administration, Houston, Texas. (ERIC Document Reproduction Service No. ED 457557)

Crow, G. M., Hausman, C. S., & Scribner, J. P. (2002). Reshaping the role of the school

principal. In J. Murphy (Ed.), The educational leadership challenge: Redefining leadership for the 21st century. Chicago: National Society for the Study of Education.

Crow, G. M., & Matthews, L. J. (1998). *Finding one's way: How mentoring can lead to dynamic leadership*. Thousand Oaks, CA: Corwin Press

Hallinger, P., Bickman, L., & Davis, K. (1996). School context, principal leadership, and student reading achievement. *The Elementary School Journal*, *96*(5), 527-549.

Hallinger, P., & Heck, R. H. (1996b). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly*, *32*(1), 5-44.

Hassler, L. B., Buck, J. A., & Torgesen, J. K. (2004, April). Preparation for high stakes reading assessments: The scope and nature of test preparation activities. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Hassler, L., Ogawa, G., Ossont, D., Nahmias, C. Goldwyn, S., Diefenbach, B., & Albee, A. (2006, April). Data-driven school reform and instructional decision-making: Barriers and challenges. Paper presented at the annual meeting of American Educational Research Association in San Francisco, CA.

Impara, J. C., & And Others. (1993). Educational administrators' and teachers' knowledge of classroom assessment. *Journal of School Leadership*, *3*(5), 510-521.

Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, *106*(6), 1258087.

Institute for Educational Leadership. (2000). *Leadership for student learning: Reinventing the principalship. School leadership for the 21st century initiative: A report of the task force*

on the principalship. Washington, DC: Author.(ERIC Document Reproduction Service No. ED458685)

Johnson, R. (2002). *Using data to close the achievement gap: How to measure equity in our schools.* Thousand Oaks, CA: Corwin Press.

Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. Center for Applied Research and Educational Improvement, University of Minnesota; Ontario Institute for Studies in Education, The University of Toronto; The Wallace Foundation.

Leithwood, K., & Steinbach, R. (1995). *Expert problem solving: Evidence from school and district leaders*. Albany, NY: State University of New York Press.

Love, N. (2004). Taking data to new depths. *Journal of Staff Development*, *25*(4).

Mason, S. A. (2003, April). Learning from data: The role of professional learning communities. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Massell, D. (2000). The district role in building capacity: Four strategies. (CPRE Policy Brief RB32). Philadelphia, PA: Consortium for Policy Research in Education.

Miller, C. A. (2000, April). School reform in action. Paper presented at the annual meeting of the American Educational Research Association Conference, New Orleans, LA.

Stecher, B., Hamilton, L. S. (2006, April). Using test-score data in the classroom. Santa Monica: CA: Rand Corporation.

Waters, T., Marzano, R. J., & McNulty, B. (2003). Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement. Aurora, CO: Mid-continent Research for Education and Learning.

Witziers, B., Bosker, R. J., & Kruger, M. L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly, 39*(3), 398-425.