The Direct Effects of Principal–Student Discussions on Eighth Grade Students’ Gains in Reading Achievement: An Experimental Study

Jack P. Silva¹, George P. White², and Roland K. Yoshida²

Abstract

Purpose: School effectiveness and instructional leadership research over the past 30 years has largely concluded that principal effects on student achievement are small and indirect. It has been assumed that the principal effect is important but mediated through other school factors. Findings: This experimental study found that one-on-one discussions between a principal and a non-proficient student that focused on the student’s 2008 reading score and a goal for his or her 2009 reading score had a direct and significant effect on the student’s subsequent reading achievement gains on a state reading test. Students in the experimental condition who held discussions with a principal prior to the state reading test showed reading gains significantly larger than students in the control condition who had their discussions after the state reading test. The randomly assigned participants (20 in the experimental condition and 21 in the control condition) were 41 of the 66 eighth grade students who made up the entire non-proficient population of a large

¹Bethlehem Area School District, Bethlehem, PA, USA
²Lehigh University, Bethlehem, PA, USA

Corresponding Author:
Roland K. Yoshida, Lehigh University, 111 Research Drive, Bethlehem, PA 18015, USA
Email: rky2@lehigh.edu
suburban middle school. Student achievement gains were calculated as the difference between the predicted versus actual reading percentile score as reported by the state’s value-added system of assessment. **Conclusions:** The results of this study provide the first steps in investigating the potential benefits to principals who seek the most immediate and direct ways to improve student achievement to meet the leadership imperatives of No Child Left Behind. Future research should investigate the principal–student discussion effects on students of different genders, grade levels, and communities.

**Keywords**
accountability, direct effects, instructional leadership, principals, school effectiveness

**Purpose of the Study**
Increasing the student achievement in the aggregate for a school district, school, and classroom or for individual students has always been one of the primary goals of teachers and school administrators alike. However, national and state learning standards as a result of the No Child Left Behind Act (NCLB), for better or for worse, have increased attention on attaining goals for student performance. The public scrutiny of a school’s performance on the state system of assessment, particularly the scrutiny that accompanies a “warning” or “school improvement” designation, places direct pressure on principals to change school practice and increase student achievement. Although principals once focused on multiyear goals such as increasing communication between the school and the community or improving student reading performance by implementing a new curriculum, they now target annual goals and finite resources to increase the chances that student subgroups make adequate yearly progress (AYP) on the high-stakes state assessment. Failing to make AYP, with its subsequent embarrassment and outcomes-or-else philosophy (Schoen & Fusarelli, 2008), has largely reprioritized the goals of principals and changed their day-to-day notions of effective school leadership.

The first principals who encountered NCLB developed their outlook on school leadership and student achievement from the research conducted on instructional leadership in effective schools from the late 1970s to the early 1990s. The effective schools movement widely held that principals had to assume the role of instructional leader for their school’s students to reach higher levels of achievement. Broad consensus existed around the conclusion that principals indeed do make a difference despite the disagreement over the
degree to which principals affected student achievement as well as the overall lack of theoretical and methodological sophistication of principal-effectiveness studies and the absence of strong empirical verification (Hallinger & Heck, 1996). Specifically, the theoretical model for instructional leadership found most often in the empirical investigations proposed three key principal behaviors: refining the school mission, managing the instructional program, and promoting a positive school learning climate (DeMaeyer, 2007; Hallinger, 2005). However, this theoretical model of instructional leadership found commonly in the literature was developed prior to the full understanding and implementation of NCLB and is now losing its importance and influence on practice. A new focus is needed in the field of principal effects, one that accounts for the nuances of leading schools under NCLB.

Today’s principals, informed by more than three decades of instructional leadership research and challenged by the urgent leadership imperatives of NCLB, seek the most powerful and direct ways to employ leadership theory and improve student achievement. The purpose of this study was to bring greater clarity on principals’ specific behaviors that support the types of immediate student achievement gains required by NCLB.

**Principal Effects in the Effective Schools and Instructional Leadership Eras**

Interest in the study of principals and school effectiveness grew in the beginning of the 1980s, fueled by the urgent tone of *A Nation at Risk* in 1983. Hallinger and Heck (1996) wrote that the summer 1982 issue of *Educational Administration Quarterly* examined the research on principal leadership and contained two landmark reviews, “Research on the School Administrator: The State of the Art, 1967–1980” (Bridges, 1982) and “The Instructional Management Role of the Principal: A Review and Preliminary Conceptualization” (Bossert, Dwyer, Rowan, & Lee, 1982).

These reviews came to two different conclusions. Bridges (1982) examined the methodologies of school effects studies and took researchers to task, concluding that researchers persisted in treating research problems in an ad hoc rather than programmatic fashion and that the research had little or no practical utility. Bridges and subsequent researchers (Creemers, Scheerens, & Reynolds, 2000; Luyten, Visscher, & Witziers, 2005; Pounder, Ogawa, & Adams, 1995; Scheerens, 1997; Scheerens & Bosker, 1997) noted the need for theory building in principal effects. The other perspective was that of Witziers, Bosker, and Kruger (2003), who summarized Bosseret et al.’s (1982) conclusions: “The principal’s routine behaviors create links between characteristics..."
of school organization and instructional climate, which in their turn affect student achievement” (p. 401). Bossert et al.’s conclusions, although recognizing the lack of consistent theory and empirical studies, were more optimistic and supported the conclusion that principals had some indirect effect on the achievement of students. More importantly, Bossert et al.’s work and those of other effective schools researchers (Leithwood & Montgomery, 1982; Sirois & Villanova, 1982) inspired better instrumentation and empirical studies of instructional leadership that investigated specifically the nature and degree of principal effects on student achievement.

Despite the call to develop better designed studies, Hallinger and Heck (1996) challenged the theoretical and methodological benchmarks of principal-effect studies in their landmark review of 40 empirical studies conducted between 1980 and 1995. They concluded that many of the studies contained weak methodologies, small sample sizes, and a wide range of unrelated measures. Nevertheless, Hallinger and Heck defended the value of nonexperimental studies, citing a growing number of studies’ “ability to exploit large samples using sophisticated data analysis techniques [that] compensate to some degree for other limitations” (p. 15). These studies completed over those 15 years grew in scientific precision and showed a systematic effort to build on the conceptual and methodological work of others. Each of the 40 empirical studies was analyzed to determine its theoretical model. Hallinger and Heck adapted the Pitner model (Pitner, 1988) to classify the empirical studies into two overarching theoretical models by which principals influence student achievement: the direct-effects model and the mediated-effects model.

Direct-effects studies explored the relationship between principal leadership and student achievement in the absence of other features of school organization. Direct-effects studies revealed, at best, very small principal effects on student achievement. Hallinger and Heck (1996) concluded that the direct-effect studies did not account for the environmental conditions or school-based variables that influenced student achievement. Their skepticism was consistent with previous reviews that, even with various control variables, it was not clear what exactly was found when the results revealed a positive relationship. Clearly as the 1980s progressed, shedding more light on the principal leadership–student achievement relationship required empirical study of relevant antecedent variables such as the organizational context of the school and the unique qualities of the principal.

Hallinger and Heck (1996) wrote that the more theoretical mediated-effects studies assumed “some or all of the impact attained by administrators on desired school outcomes occurs through the manipulation of, or interaction with, features of the school organization” (p. 18) such as school size, the students’
socioeconomic status (SES), and principals’ gender, teaching experiences, and leadership philosophy. Similar to direct-effect studies, the mediated-effects studies often contained weak methodologies and simplistic statistical techniques that yielded very small principal effects. Given that the studies tested antecedent effects on leadership and leadership effects on outcomes separately, the authors concluded that although principals may have exerted some direct effect on students’ learning, the actions of others in schools significantly contributed to the relationship between principal leadership and student learning. The majority of the mediated-effects studies reported a significant effect of principal leadership on organizational processes and, in turn, student achievement.

Mediated-effects studies also contained a variety of definitions of student achievement, ranging from standardized achievement test scores to teacher-conceived outcomes. Yet the trends in the research showed that the effects of principal leadership were stronger on in-school process than student achievement gains. The preference toward investigating indirect effects on student achievement was reflected consistently within the developing body of effective schools and pre-NCLB instructional leadership research—a body of research that tested many different mediating organizational factors that interact with principal effect and student achievement.

One such factor was the principal acting in his or her more traditional role of instructional manager. Leitner (1994) used Hallinger’s (1983) instructional management framework to divide instructional management into three dimensions: defining school mission, managing instructional programs, and promoting positive school climate. Leitner found little or no relationship between any dimension of instructional management and student achievement. However, Sammons, Hillman, and Mortimore’s (1995) literature review concluded that the quality of the principal’s leadership and shared vision and goals with teachers were significantly related to student achievement gains. Similarly, Gullatt and Lofton’s (1996) and Gurr-Mark, Drysdale-George, and Mulford’s (2010) reviews supported the conclusion that principals’ self-perceptions as instructional leaders were correlated with their students’ improved achievement. The growing consensus of effective schools research, developed largely before NCLB, supported the view that a principal’s leadership had a significant yet indirect effect on the success of individual students when the principal provided instructional leadership beyond managerial leadership.

Another key mediating factor was the teacher perception of principal leadership. Short and Spencer (1990) analyzed teacher perceptions of a principal’s instructional leadership and student levels of motivation in the classes of 16 secondary schools. Specifically, significant correlations were found between
higher levels of student motivation and teacher perceptions of principals as effective communicators of school goals and active supervisors of instruction. Andrews and Soder (1987) conducted a more comprehensive study that explored 12 organizational characteristics categorized into four sets hypothesized to increase student achievement. Teachers rated their principals as (a) resource provider, (b) instructional resource, (c) communicator, and (d) a visible presence. Based on these ratings, principals were categorized as strong, average, or weak. Findings showed that typical equivalent gain scores of students in the strong leader schools were significantly higher than those of students in the average or weak leader schools. In short, teachers’ perceptions of their principal as an instructional leader were highly correlated with the reading achievement gains of students, particularly among low-achieving students. However, explorations of the relationship between student perceptions of the principal’s instructional leadership and student achievement gains were left out of most of the other principal self-perception and teacher-perception studies.

It was not surprising that Hallinger’s (2005) review of empirical studies of instructional leadership concluded that principals contributed to school effectiveness and student achievement indirectly through actions they take to influence school and classroom conditions. The greatest principal effect on student achievement occurred when principals acted as instructional leaders in which they focused on defining school mission, managing the instructional program, and promoting a positive school learning climate. These effects were statistically small, mostly accounted for when principals shaped the school’s mission (Bamburg & Andrew, 1990; Goldring & Pasternak, 1994; Hallinger, 2005).

Similar to Hallinger (2005), Witziers et al. (2003) tested the numerous avenues of well-researched principal effects using a meta-analysis of studies conducted in European school systems: dimensions of instructional management (Hallinger, 1989), leader behavior and situational characteristics (Hoy & Miskel, 1991), and transformational leadership (Leithwood & Jantzi, 1990). They found that school leadership does have a positive and significant effect on student achievement though the effect sizes are very small, approximately .07. Results concerning the subdimension effect of strongly defining and communicating mission showed the strongest effect size (Z = .19). The conclusions of the Dutch study were similar to those of the body of U.S. effective schools and instructional leadership research—the overall principal effect is indirect and small.

Both Hallinger (2005) and Witziers et al. (2003) left two key questions unanswered. First, would the effect be significantly larger if the principal interacted directly with individual students? Second, could any modest indirect effects
created by shaping the school’s mission be of enough practical value for principals facing the well-publicized requirement of having 100% of their students achieve proficiency on state tests by 2014?

**Getting Personal: Direct Principal Effects**

The new NCLB notion of principal leadership in which specific achievement goals are expected to be met within specific amounts of time compels principals to find strategies and tactics that will result in direct effects on student achievement not only to open opportunities for children but also to save their schools from embarrassment and, potentially, to save their own jobs. Nettles and Herrington (2007) summarized the new imperative:

> The traditional policy focus regarding student achievement has been on classroom factors (e.g., scientifically based curricula and teacher quality), and appropriately so, but the national focus is now turning to what a principal can do to improve student achievement. This is a significant redirection, because actions taken to better understand and improve the impact of principals on the achievement of students in their schools has the potential for widespread benefit, as individual improvements in principal practice can impact thousands of students. (p. 724)

Principals are taking more direct action in areas that they hope will increase their students’ achievement. Nettles (2005) conducted a study of the relationship between the principal’s role in the implementation of the Reading First program and the reading achievement of 34,000 first graders in Florida; this study sheds light on the direct effect of the behaviors of 388 principals on student achievement (Nettles, 2005). Nettles found that the principals’ increased implementation of effective reading intervention practices correlated with students’ gains in additional words per minute and accelerated rates of fluency. Significant relationships between increased principal implementation of assessment practices and reading achievement were found among student subgroups, including limited English proficient students and students with learning disabilities.

These encouraging results suggest that even more targeted interventions with specific subgroups of students can be the best strategy for achieving AYP. Principals could focus their time as an instructional leader on those at-risk students and achieve overall school success. Guidance on how best to approach students is found in the literature concerning student perceptions of the principal’s instructional leadership (Gentilucci & Muta, 2007). Students perceive and can explain how the visibility of the principal or the pattern of interaction...
between the principal and the students affects the students’ level of effort and achievement. Gentilucci and Muta (2007) asked two very important and relevant questions: Do students perceive that leadership behaviors of principals have a direct effect on their learning and academic achievement? If yes, what specific leadership behaviors do students perceive most positively affect their academic achievement? They found that eighth grade students, regardless of their ethnicity, SES, English proficiency, or academic ability, reported higher levels of motivation and achievement when their principal engaged them in individual, instructionally focused discussions about their class work. In addition, students in classes in which a principal engaged in discussion with teachers and students during a formal classroom observation reported higher levels of motivation and achievement compared to students in classes in which principals observed passively. Specifically, students appreciated these principals’ accessibility to engage in formal and informal conversations with them. The principals were also perceived to be very approachable and interested in their academic and nonacademic progress rather than distant and police-like. Such relationships with their principals resulted in feelings of trying harder in their academic work. However, their study was an ethnographic study with a small sample, and the achievement measures were not the same for all of the students. They also did not describe the kinds of behaviors that led the students to conclude that the principals were accessible and approachable.

The instructional leadership literature further suggests that principals should maintain a safe and orderly environment, provide a clear mission, involve stakeholders, monitor school progress, maintain an instructional focus, hold high expectations for student performance, and support professional development of teachers (Nettles & Herrington, 2007). The argument can be made that principals who embrace the three components most directly applicable to students, namely, their ability to communicate a mission directly to students, to involve students in monitoring their own progress, and to communicate high expectations directly to individual students, may have the best chance of affecting the achievement of certain subgroups of students significantly. A new set of empirical research should test specific dimensions of instructional leadership that may have the most direct effects on student achievement, especially when principal actions are conducted specifically and individually with the students who need the achievement gains the most. Thus, the research questions for this study are the following: Do principals have a significant direct effect on an individual student’s achievement, as indicated by the difference between the student’s projected score and the actual score achieved by the student on the Pennsylvania State System of Assessment (PSSA) Reading Test, when they communicate a mission directly to the student, involve the student in monitoring his or her own progress, and communicate high
expectations directly to the individual student in an achievement-based discussion? And if so, to what degree is the effect?

**Method**

**Population and Sample**

Table 1 presents the participant characteristics for the population and final sample for this study. A total of 66 students represented the school district’s entire population of nonproficient eighth grade students as determined by their seventh grade 2008 PSSA Reading Test scores. The students were randomly assigned in equal numbers to an experimental or control condition. Of the 66 students in the population, 25 did not participate in the study because their parents did not give consent. Differences by gender of the participants and nonparticipants were significant, $\chi^2(1) = 6.76, p < .01$; 16 of 34 males participated in the study, whereas 25 of 32 females did so. Similarly, high SES students (34 of 48), those not qualifying for free or reduced-priced lunches, participated at a significantly higher rate than low SES students (7 of 18), $\chi^2(1) = 5.67, p < .05$. No differences were found for ethnicity. Furthermore, no differences were found between the experimental and control groups for gender, SES, ethnicity, or prior achievement scores.

The principals who engaged the students consisted of the principal and the two assistant principals in the school in which the students were enrolled. Principal 1 was a 51-year-old White, middle-class male. Principal 2 was a
31-year-old White, middle-class male. Principal 3 was a 54-year-old White, middle-class female. The three principals implemented both conditions of the independent variable. Each principal was randomly assigned discussions with 11 students from the experimental group and 11 from the control group. Because of participant nonconsent to participate, Principal 1 held discussions with 7 participants in the experimental condition and 6 in the control condition, Principal 2 with 5 in the experimental condition and 5 in the control condition, and Principal 3 with 8 in the experimental condition and 10 in the control condition. Thus, 20 and 21 students completed the experimental and control conditions, respectively.

**Research Design and Treatment Conditions**

This study incorporated an experimental design to address the specific research question. The independent variable consisted of two conditions: (a) an experimental condition in which students engaged in achievement-based discussions with the principal prior to the eighth grade 2009 PSSA Reading Test and (b) a control condition of no treatment prior to the eighth grade 2009 PSSA Reading Test. Students in the control condition had similar discussions with the principal as students in the experimental condition, but the discussion took place after the eighth grade 2009 PSSA Reading Test and focused on the importance of reading achievement in subsequent grades. The PSSA is the annual standardized test developed, scored, and reported to all Pennsylvania public schools by the PSSA. The PSSA Reading Test is given to all students in Pennsylvania public schools in Grades 3–8 and 11 and is used to determine a school’s AYP under the requirements of NCLB.

In the experimental condition, students met twice individually with a principal during the month immediately prior to the eighth grade 2009 PSSA Reading Test. During the first meeting, the principal and the student engaged in a 15-minute achievement-based discussion. The discussion protocol focused on six components: (a) introductions and general discussion to put the student at ease, (b) a statement of the school mission and the principal’s high expectations for students’ improved reading performance, (c) a review of the student’s individual achievement report from the seventh grade 2008 PSSA Reading Test, including identification of the student’s overall level of performance, areas of relative strength, and areas of relative need, (d) identification of the Pennsylvania Value Added Assessment System (PVAAS) projected score for the student, (e) collaboratively setting a goal for the student’s percentile score on the eighth grade 2009 PSSA Reading Test, and (f) expressions
of appreciation, support, and encouragement to the student by the principal. In the second meeting, the principal conducted a follow-up discussion within a week prior to the mid-March administration of the 2009 PSSA Reading Test. The principal recorded information from the discussions on a form and submitted the form of each meeting to the first author.

In the control condition, students met twice with the principal during the month immediately after the eighth grade 2009 PSSA Reading Test. During the first meeting, the principal and student engaged in 15-minute achievement-based discussion. The discussion protocol focused on five components: (a) introductions and general discussion to put the student at ease, (b) a statement of the school mission and the principal’s high expectations for students’ improved reading performance, (c) a review of the student’s individual achievement report from the seventh grade 2008 PSSA Reading Test, including identification of the student’s overall level of performance, areas of relative strength, and areas of relative need, (d) collaboratively setting a goal for the student’s achievement on high school reading assessments, and (e) expressions of appreciation, support, and encouragement to the student by the principal. In the second meeting, the principal reviewed the discussion from the previous meeting. The principal recorded information from the discussion in the same way as described for the experimental condition.

Dependent Measure

The scores from the eighth grade 2009 PSSA Reading Test were used as the basis for calculating the dependent measure for each of the participants. The reliability of the 2009 PSSA Reading Test for eighth graders ranged from .90 to .94 for each decision category: below basic–basic, basic–proficient, and proficient–advanced (Pennsylvania Department of Education, 2009). Validity tests found that the 2009 PSSA Reading Test generally measured a distinct construct from other subject matter tests (Pennsylvania Department of Education, 2009). Finally, scores of students from previous versions of the PSSA Test have been found to correlate with their scores on often used achievement tests such as the California Test of Basic Skills and the California Achievement Test (Thacker, Dickinson, & Koger, 2004).

The dependent variable was the net gain score in the number of units that each student’s eighth grade 2009 PSSA Reading percentile score was above or below the percentile score predicted for each student by the PVAAS. The PVAAS incorporates a mixed longitudinal model that estimates the growth that a cohort of students may experience during a school year. The PVAAS
provides predicted percentile scores of each individual student’s eighth grade PSSA Reading performance based on his or her performance on his or her previous year’s PSSA Reading performance.

Procedure
In early February 2009, the first author conducted a simulated training of the two discussion protocols with the three middle school principals. The achievement-based discussions for the experimental group took place in late February and early March 2009 prior to the PSSA Reading Test in mid-March. The discussions for the control group took place in late March and April 2009. These discussions took place in a random order to the extent administratively feasible during the days of the week and the periods of the day. The principals followed the protocols for the meetings with students as described in the Research Design section. The first author reviewed these reports to quality control check the integrity of the two conditions. The principals followed state-mandated procedures related to PSSA testing. The Pennsylvania Department of Education scored and reported the data related to students in both conditions in July 2009.

Data Analysis
A $t$ test of independent samples was used to test for differences between the means of the experimental ($n = 20$) and control ($n = 21$) groups with an alpha level of .05. Tests were not conducted on the principal variable for two reasons. For the experimental condition, the number of students participating with each principal was very small ($ns = 5, 7, 8$); thus any test would not have sufficient power. A test of principal effects for the control group was not conducted because the principals did not interact with the students before the PSSA was given as the principals did in the experimental condition; hence, they could not have influenced their behavior.

The initial $t$ test comparing the experimental and control groups was not significant. However, inspection of the data revealed three cases in the control group that were considered outliers because their scores were considerably outside of the interquartile range of the control group. These students were interviewed about their school experience between the 2008 PSSA and 2009 PSSA testing periods. One outlier was removed from the control group because this student’s net gain score was more than two standard deviations above the
A t test was applied again to the experimental group \((n = 20)\) and the control group without the one case \((n = 20)\). The same alpha level was used for this test. A detailed presentation of the data resulting from these analytic steps is presented in the Results section.

**Experimental Check**

Each participant completed a questionnaire within 2 weeks after taking the eighth grade 2009 PSSA Reading Test to determine whether the participant experienced the treatment as the researcher intended. The questionnaire asked what kinds of discussions the participant had with the principal during the weeks before and after the 2009 PSSA Reading Test and how the discussion affected his or her level of motivation to improve his or her PSSA Reading Test performance.

**Results**

Results of the experimental check indicated that students in both conditions received the treatments from all three principals as intended. Students in both conditions self-reported that the principal–student discussions involved direct and focused goal setting for reading achievement. All but one student in the experimental group self-reported that the discussions with his or her principal led to “more” or “a lot more” motivation to improve on the eighth grade 2009 PSSA Reading Test.

Table 2 presents the descriptive statistics by conditions (control, experimental) and principal (1, 2, 3). The means of net gains for the students were

<table>
<thead>
<tr>
<th>Principal</th>
<th>Frequency</th>
<th>(M)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Experimental</td>
<td>7</td>
<td>1.29</td>
</tr>
<tr>
<td>2</td>
<td>Experimental</td>
<td>5</td>
<td>3.20</td>
</tr>
<tr>
<td>3</td>
<td>Experimental</td>
<td>8</td>
<td>3.38</td>
</tr>
<tr>
<td>All</td>
<td>Control</td>
<td>21</td>
<td>−2.00</td>
</tr>
<tr>
<td>Experimental</td>
<td>20</td>
<td>2.60</td>
<td>8.67</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>0.24</td>
<td>10.25</td>
</tr>
</tbody>
</table>

2.60 percentile points in the experimental group and –2.00 percentile points below their predicted scores for the control group, a net loss. The \( t \) test comparing the two groups was not significant (\( t = 1.45, ns \)).

However, Figure 1 illustrates the different distribution of scores between conditions. Each box plot portrays a 5-point summary for the distribution of scores attributable to each principal: the lowest score, the first quartile, the median, the third quartile, and the highest score. Each box represents the interquartile range and allows for an easy comparison of the shape and spread of the distributions. Figure 1 shows that a sizable difference was found between the medians for the experimental and control conditions. The difference between the medians was as large as the interquartile ranges.

Scores were generally more tightly clustered around the median in the control condition, with the exception of the values for three outlier students, Students 39, 53, and 38. The first author subsequently interviewed each outlier and reviewed his or her previous year’s PSSA Reading Test scores to determine if any situational event led to his or her significant net gain.
Student 39 had a striking 30-point net gain. Student 39 reported that his or her 2008 score was affected by an illness during the test. Further investigation revealed that Student 39 scored at the 58th percentile, 38th percentile, and 23rd percentile on the 2006 PSSA Reading Test, 2007 PSSA Reading Test, and 2008 PSSA Reading Test, respectively. It appears that Student 39 recaptured a level of achievement more reflective of his or her ability before a 3-year decline staring in fifth grade.

Student 53 had a 21-point net gain and reported that her improvement was caused by “really liking her eighth grade language arts teacher.” Further investigation revealed that Student 53 scored at the 45th percentile and 25th percentile on the 2007 PSSA Reading Test and 2008 PSSA Reading Test, respectively. Similar to Student 39, Student 53 appeared to return to a level of achievement more reflective of her performance before a 2-year decline starting in sixth grade.

Student 38 had a 15-point net gain and reported that she “didn’t have the problems in the group home” that she had had the previous year. Student 38’s previous PSSA Reading Test scores indicated that her net gain in 2009 was unprecedented in her testing history. Student 38’s net gain may have been reflective of the academic benefits that accompany an improvement home life. None of the outliers mentioned having a discussion with a principal or not having a discussion with a principal as a factor contributing to their net gains. Their gains all appeared to be situational in nature and unrelated to their treatment or condition.

These three students (38, 39, and 52) had attained scores that we considered outliers within the control group. One student (39) was considered an extreme case because his or her performance was more than two standard deviations above either the experimental or control group. Thus, we removed this student’s score from the control group sample and performed another t test. The net gain scores of the experimental group \( (M = 2.60, SD = 8.67, n = 20) \) were compared with those of the revised control group that did not include the one outlier \( (M = -3.60, SD = 8.83, n = 20) \). The t test was significant \( (t = 2.24, \ p < .05, \text{Cohen’s }d = 0.71) \), indicating a moderately large effect size for the treatment.

**Discussion**

This study is the first step that builds on the findings of Gentilucci and Muta (2007), who found that student perceptions about their principals matter in terms of students’ level of effort and achievement. Similar to Gentilucci and Muta’s (2007) study, the eighth graders in this study reported higher levels of effort and achieved better outcomes regardless of their ethnicity, SES, or
Specifically, the experimental check showed that all but one student in the experimental condition reported more motivation or a lot more motivation to do well on the 2009 PSSA Reading Test as a result of their achievement-based discussions with a principal. The current study goes beyond Gentilucci and Muta’s because it was conducted in an authentic school setting with a sample of students who were labeled as below proficient on the statewide reading test. It also employed a pure experimental design with random assignment of these students to treatment. Finally, the current study’s achievement variable of net gains on the state assessment is a more valid and relevant measure of improvement than student perceptions and self-reported effort.

Even with these new procedures and measures, we caution that this study is but a first step. This study was conducted in one suburban school district that enrolled very few low-income or non-White students. Although internal validity issues were controlled within the study, the participation rates in the study, regardless of experimental or control condition, differed depending on the students’ gender and SES level but not by ethnicity for those few students who participated. Given the ages of the students, their parents had to give consent for them to participate. We do not know whether either or both the parents and students were unwilling to give consent to participate. Nevertheless, principals are perceived to be powerful authority figures who can influence student careers, positively and negatively. How do parents and students perceive an invitation to meet the principal, as well intentioned that invitation may be, when the students are labeled as below proficient? The school histories for boys show that they are generally referred and placed in special education more often than girls. Students and their parents from low SES backgrounds may be reticent to participate because they are from families who have very little influence in the schools. We can only speculate about the reasons for the different rates of participation because we do not know the personal histories of the nonparticipating students, which could vary from very productive relationships with principals and teachers to those that result in isolation and rejection from the school.

Nevertheless, the results are promising because the experimental treatment appears to increase the performance of students who were classified as nonproficient. The underlying cause for the increased performance may be attributed to the students’ reported increase in their motivation to do well on the PSSA examination. Our experimental treatment basically helped low-achieving students set goals for performance and provided them with encouragement from an authority figure. How these behaviors affected the motivation
levels of these students is complex and not directly investigated in this study. However, one approach to conceptualizing motivation may be Wigfield and Eccles’s (2000) expectancy-value theory of achievement motivation, in which they suggested that students develop more accurate understanding of how good they may be at school subjects as they progress through school. Thus they develop a more precise sense of their expectancies for future success in those subjects.

This study did not sample students with variations on the variables that Wigfield and Eccles (2000) and others have suggested. However, with a group of students labeled as below proficient, principal intervention may have altered the experimental students’ self-weighting on these variables. The experimental treatment was a very personal experience in which principals discussed each student’s relative needs. They also collaboratively set goals for each student’s achievement on the upcoming test that may have increased the student’s expectancy for success. They expressed appreciation and support and encouragement for the student that may have increased the student’s value of performing well on the examination. Although these mechanisms for increased motivation were not specifically investigated, all but one student in the experimental condition reported that the discussions with the principals led to higher levels of motivation. Furthermore, one of the three outliers from the control group reported that her improvement was caused by “really liking her eighth grade language arts teacher.” Perhaps liking her teacher increased her self-assessment of expectancy for success in the subject and her valuing the subject more than before.

Leithwood and Riehl (2003) may have been too pessimistic when they wrote, “Although the accountability demands appear to be getting the attention of school leaders, principals may not be adjusting their practice in ways that truly benefit students and teachers” (p. 5). Their pessimism reflected the pre-NCLB literature on principal effects and ignored the recent research that emphasizes the importance of principal attention to motivation levels of individual students. This study’s preliminary findings of direct effects should bolster school leaders’ search for alternatives to cynical test preparation curricula and slow-moving school improvement plans.

Indeed, this experimental study found promising results that support the need to develop research programs with larger samples and variations of the specific elements of principal–student discussions that may increase the power of the interventions such as the content of the discussion, the number of discussion contacts, each discussion’s duration, and the type of students involved in the discussion. Studies with larger samples will also mitigate the disproportionate influence of a few outliers on the results, though the presence
of such outliers should alert research and administrators alike to be cautious about relying on yearly improvement as the sole or most significant measure of student achievement or a school’s annual performance.

However, the principals’ position alone or their ethnic background may not determine how they may develop functional relationships with students. Previous research conducted with the early 1970s investigated whether African American children could have tested lower on their individual intelligence tests than White children when White examiners tested them (Solkoff, 1972). African American and White children were randomly assigned to either African American or White examiners. The results showed that African American students tested by White examiners did not score lower than those tested by African American examiners. The findings from that study suggest that position or ethnic status alone do not relate to performance.

Thus, principals should review what they can do to develop a school culture that will help get students to the table. The extent to which a promising intervention has an effect on an entire group of students is lessened if some targeted students do not participate. Perhaps the methods used in the experimental intervention can be more generally applied to all students in a school such as frequently communicating a mission for the school community, monitoring student progress, and communicating high expectations (Witziers et al., 2003). We do not know which of these three principal behaviors had the largest effect in the experimental condition compared to the others or if any of the three principal behaviors actually weakened the discussion’s effect. Perhaps a single behavior may be just as effective as all three in combination. Studies of behavior-specific discussions will further support the efficient use of the principal’s time. Witziers et al.’s meta-analysis concluded that communicating mission had the strongest albeit indirect effect on student outcomes. Communicating mission about student learning could be the first content item to be investigated on its own in the principal’s direct intervention with individual students and more broadly in the school for all students to develop a culture for participating in the targeted intervention.

Achievement-based discussions must also focus on topics beyond reading. Subgroup performance in mathematics is as important as reading to many NCLB-era principals. Mathematics can be an intimidating subject for non-proficient students, resulting in debilitating math anxiety. Studies need to be done on what principals can say and do that may reduce math anxiety and increase the motivation to succeed in math.

Another important set of variables is whether these results could generalize from our sampled suburban students to students from urban communities, to different populations of students at risk, to students of different genders,
and to students at different grade levels. This study took place in a relatively wealthy suburban community in which personal adult–student relationships are more common than in poorer urban communities. For example, Gentilucci and Muta’s (2007) study found that eighth graders, regardless of their ethnicity, SES, English proficiency, or academic ability, self-reported higher levels of motivation and achievement caused by the principal’s accessibility in formal and nonformal conversations. Nettles (2005) found that a principal’s direct action in implementing the Reading First program with urban students showed a significant relationship between implementation practices and reading achievement gains including gains for students with learning disabilities and limited English proficiency. These two studies, together with the results from the present study, show the promise of the direct approach for increasing achievement. However, given the lack of consensus in principal direct-effects research, replication studies are needed in urban settings and smaller suburban and rural settings to identify geographic, ethnic, and economic influences on achievement-based discussion effects.

The long-held belief that male students learn differently than female students should be tested in terms of the effect of achievement-based discussions. Would female students benefit disproportionately from a condition that emphasizes relationships and conversations? Or would males respond more dramatically to a condition that directly employs an authority figure? Achievement-based discussions contain strategies that have been included in studies of the unique needs of both male and female learners. Further achievement-based discussion studies will help clarify both the effect of the discussions and the learning differences between genders.

Finally, the current study focused on the effect of achievement-based discussions on eighth graders in a middle school. Middle school students were assumed to be more articulate than elementary students and more willing to share information than high school students. Given the positive effect of achievement-based discussions found in this study, replication studies should take place with fifth grade students or even younger students. The potential of at least 7 years of interactions with a principal that begin in students’ early elementary school years should not be ignored. Such interactions may develop more positive profiles on the motivational variables that Wigfield and Eccles (2000) proposed. Proficiency rates on PSSA Reading Tests typically decline as cohorts of students progresses through middle school and high school, especially among subgroup populations such as students with disabilities. The decline must be arrested sooner rather than later to give secondary principals the time and tools to improve the performance of larger populations of nonproficient students.
Achievement-based discussions may be an effective approach that will help refocus a principal’s work from leaving no child behind to moving that one child forward. More research is yet to be done to maximize the potency of this approach. This study provides us with some preliminary data to reconsider the potential of the principal’s direct effect on children and their motivation to succeed.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interests with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

References


**Bios**

**Jack P. Silva** is the Assistant Superintendent/Chief Academic Officer for the Bethlehem Area School District. He oversees all aspects of school improvement, curriculum, assessment, instruction, and professional development. His research and practice focus on improving school performance by assisting principals and other educational leaders in their instructional leadership practices with teachers and students.

**George P. White** is professor of educational leadership and Director of the Center for Developing Urban Educational Leaders at Lehigh University. His current work lies at the intersection of leadership preparation, urban school reform, and organizational change.

**Roland K. Yoshida** is professor of education and Provost Emeritus of Lehigh University. Since returning to faculty, he has chaired dissertations of students who have posed interesting questions about leadership per se and the issues that confront contemporary school leaders.