IDENTIFYING DIFFERENCES BETWEEN TWO GROUPS OF HIGH-NEEDS HIGH SCHOOLS

June 2008

Prepared by

Kerry Englert, Ph.D.
Zoe A Barley, Ph.D.

McREL

June 2008

This project was conducted for the Institute of Education Sciences (IES) under Contract ED-06-CO-0023 by the Regional Education Laboratory Central administered by Mid-continent Research for Education and Learning. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement of the U.S. Government.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The authors are grateful for the outstanding support provided by the team of researchers who contributed to this project. In particular, we would like to acknowledge Dawn Fries for her critical analysis of state standards and in maintaining and developing the sampling protocol and databases. Mya Martin-Glenn also provided substantive expertise in statistical analyses and methodology. Appreciation is also extended to Michael Phillips, whose management and tracking of school achievement and assistance with the literature review was invaluable.

SUMMARY

This study examined differences on nine factors between a group of the highest-performing high-needs high schools (HPHN) and a group of the lowest performing high-needs (LPHN) high schools using teacher responses from a national dataset. The factors are: shared mission and goals, professional development, collaboration among teachers, assessment and monitoring, parent involvement, safe climate, orderly climate, and support for teacher influence at the school level and at the classroom level. High-needs high schools were schools with greater than 50% of their students eligible for free or reduced price lunch. Meaningful and statistically significant differences in seven of the nine factors suggest that teachers' perceptions of these factors in highest-performing high schools differ from teachers' perceptions in lowest performing high schools. These differences were identified through a descriptive analysis and do not mean that aligning teacher perceptions in low-performing high schools to match those in high-performing high schools will result in improved achievement. The factors with significant differences, however, do suggest areas for further research.

High school student achievement is one of five priority needs identified for the Central Region states (Colorado, Kansas, Missouri, Nebraska, North Dakota, South Dakota, and Wyoming). During McREL's 2007 regional needs assessment activities, 88 percent of educators surveyed in state and local meetings, and 56 percent of randomly selected respondents to a regional survey, indicated that improving high school students' achievement was of great importance. To address this need, McREL researchers examined mean differences between 248 high-performing, high-needs high schools (HPHN) and 40 low-performing, high-needs high schools (LPHN) on each of nine identified factors.

There is also a growing concern among institutions of higher education and employers that high school students do not have the knowledge and skills they need to be successful in college or the workplace. Post-secondary institutions find that high school graduates arrive less prepared than they have in the past, and many incoming freshman require remedial courses in reading, writing, and mathematics (Parsad & Lewis, 2003). Similarly, a picture of employers' concerns about the quality of entry-level workers emerges from a recent spate of studies that identify significant skill gaps among entering workers. For example, a recent skills-gap report from the National Association of Manufacturers/The Manufacturing Institute (Eisen, Jasinowski, & Kleinert, 2005) states that the majority of American manufacturers are experiencing a serious shortage of qualified employees. Finally, Dougherty, Mellor, & Smith (2006) argue that all students should receive preparation that increases their likelihood of success, in part to "equalize opportunities across socioeconomic backgrounds so that no group of citizens is denied the benefits of a strong education" (p. 2).

Following a review of the literature on effective high-needs high schools, McREL initially identified eight factors (*shared mission and goals, instructional guidance, professional development, collaboration among teachers, assessment and monitoring, parent involvement, support for instruction*, and *safe and orderly climate*) that appear to be linked to high-performing, high-needs high schools. Researchers then identified and theoretically matched survey items from the School and Staffing Survey (SASS)

dataset that are associated with these factors (See Appendix B for the survey items associated with each factor). The factor *instructional guidance* was deleted because of a lack of appropriate survey items in the SASS dataset. Additional statistical analyses were conducted on the SASS data to verify that the theoretical decisions were correct. These analyses indicated that two of these factors (*support for teacher influence* and *safe and orderly climate*) were better represented in these data if split into two factors each (support for teacher influence became *at school level* and *at classroom level* and safe and orderly became *safe*, and *orderly*), resulting in a total of 9 factors. We then confirmed sufficient reliability for the identified factors to proceed with analyses.

Researchers then generated scores based on the combination of responses each teacher gave for the items categorized within the particular factor. Differences between the highest- and lowest-performing schools for each factor were then calculated and expressed as effect sizes.[1]

The study yielded several findings. There were statistically significant differences between highest- and lowest-performing schools for eight of the nine selected factors, with teachers in the high-performing schools reporting more favorably on each.

Several of the effect size differences are notably large. The largest of these differences, with an effect size of 1.29, was *safe climate.* Another factor with a large effect size was *parent involvement* (0.70). Five additional factors with statistically significant between-group differences had smaller effect sizes: *support for teacher influence* at the *school* (0.37) and *classroom levels* (0.33), *orderly climate* (0.37), *assessment and monitoring* (0.27), and *collaboration* (0.27). The remaining two factors had effect sizes less than 0.2, considered the lower boundary of a small effect. Based on the items that match each factor, the findings suggest the following:

- The differences between the highest- and lowest performing schools on the variable *safe climate* indicate that the highest performing high-needs schools are more likely to be free of instructional interruptions (including those associated with student misbehavior).
- The differences evidenced for the *parent involvement* factor indicate that teachers felt that there were higher levels of parental support and a more positive relationship between parents and teachers in HPHN schools when compared with parents and teachers in LPHN schools. It may seem that parental involvement would be important primarily for elementary schools where parents of younger children, for various reasons, may be more inclined to participate in their child's education. However, research (see, for example, Fan & Chen, 2001; Jeynes, 2007) supports the notion that positive parental and teacher relationships, coupled with parental support of the school, are of continued importance in students' high school education.

[1] An effect size (or a standard deviation unit) is way of quantifying the difference between two groups with a single number, using a standard deviation (in this case a weighted average of the group standard deviations) to represent the difference. Use of this measure in this instance does not presume a treatment/control design.

- In terms of *support for teacher influence* at both the *classroom* and *school* level and collaboration, teachers in the highest- performing schools report that they have more influence in a range of aspects of school practice and policy than do teachers in low-performing schools. They also report more autonomy in their classrooms to make curricular decisions and determine instructional practices.
- The differences in *orderly climate* indicate that the highest-performing high-needs schools are more likely to 1) enforce policies on student behavior 2) reward appropriate behavior, and 3) discipline students for inappropriate behavior.
- Differences were seen on the factor of *assessment and monitoring*, which indicated that teachers in the highest-performing schools were more likely to use assessment data not only to monitor individual student performance, but also to adjust curriculum and identify areas in their own content knowledge that they might need to strengthen. And finally differences on *collaboration* indicate teachers in high performing high-needs schools report they work together more frequently.

McREL researchers believe that the findings from this study could serve as a framework to guide future research on high-needs high schools. These findings and conclusions suggest areas for further exploration by high-needs high school administrators and teachers. In order to determine clearly what changes in high school practice are effective in improving student performance in high needs high schools, McREL recommends research using an experimental study design to examine the efficacy of programs that feature one or more of the factors. Additional research could also examine a broader range of definitions and review additional factors that might relate to student achievement.

WHY THIS STUDY

This study examines how a sample of the highest-performing, high-needs (HPHN) high schools differ from lowest-performing, high-needs (LPHN) high schools on variables hypothesized to be related to student achievement. The data for this report were taken from the teacher section of the national 2003-2004 Schools and Staffing Survey (SASS). The results of this study suggest additional areas for research on high-needs schools and are intended to stimulate discussions among practitioners and policymakers about how high school education could be improved in high-needs high schools.

High school student achievement is one of the five priority needs identified by McREL for the Central Region. The need for guidance on improving high schools, especially for those with large numbers of students in poverty, was reconfirmed through McREL's 2007 regional needs assessment process. Eighty-eight percent of educators surveyed in state and local meetings and 56 percent of respondents to a regional random survey indicated that improving student achievement in high schools was of great importance. Further, the McREL Board of Directors has designated high school improvement as a continuing priority need.

High school improvement also has become a recognized educator priority nationwide. With the passage of the No Child Left Behind Act (NCLB, 2001) public attention has most often been drawn to elementary and middle schools. In 2003, however, the President and the U.S. Department of Education emphasized the need to close the achievement gaps among subgroups of our high school students (U.S. Department of Education, 2005). Other reports (see, for example, Achieve, Inc & National Governors Associations, 2005; and Alliance for Excellent Education, 2003), have also emphasized the importance of improving high schools. This attention by the President, the nation's governors, and U.S. business leaders has "put high school reform front and center in the education reform movement" (Barton, 2005, p. 2). Despite this recent spotlight, research on high school teaching and learning remains limited (NRC Institute of Medicine, 2003).

The purpose of this study is to identify factors that differ on teacher responses to a national survey between the highest-performing high-needs (HPHN) high schools and the lowest-performing high-needs (LPHN) high schools. The report is intended to further an exploration of what constitutes an effective high school. This study provides one basis for future research that would use an experimental design to examine programs using the factors found in this study to be perceived as different by teachers in the highest- versus lowest-performing, high-needs high schools.

Preliminary Identification of Factors Differentiating HPHN and LPHN Schools

A McREL study (McREL, 2005) of high-needs elementary schools identified 13 factors that differentiate a set of the highest-performing, high-needs schools from the lowest-performing, high-needs schools. The current study sought to match survey items from the SASS with these 13 factors. Matching survey items were found for only seven of the original 13 factors.

In developing that study we drew from Heck, Larsen, and Marcoulides (1990) and Scheerens and Bosker (1997) among others (see Appendix A for more detail). Table 1 indicates the factors present in these two studies as they compare to the 13 factors in McREL's study (2005). Factors are not operationally defined in exactly the same way in each study. For the current study, the definitions were dependent on the nature of the SASS items available.

As noted in Table 1, the current study includes only seven of the original 13 factors (see Appendix A). For most of the seven, the SASS items limited us to a more restrictive definition of the factor than was described in the literature. We conducted analyses (see Appendix D) to confirm that the selected SASS survey items did indeed match the factors identified from the literature. These statistical analyses led us to divide two of the original seven factors: *support for teacher influence* was split into *school level* and *classroom level*, and *safe and orderly climate* was split into *safe* and *orderly*, resulting in a total of 9 factors. Definitions for the 9 factors derived from the SASS items are provided in box A.

Box A

Shared mission and goals center around framing, communicating, and enlisting engagement in a common school mission. The principal's values and beliefs are known to teachers and are aligned with the school's mission and goals.

Professional development focuses on the perceived usefulness of the professional development that teachers receive.

Collaboration among teachers fosters the sharing of work and expertise among teachers and creates a sense of affiliation and support. Teachers are encouraged to work together in one another's classrooms by observing, providing feedback, and coordinating instruction.

Support for teacher influence: school level means that principals and administrators share leadership responsibilities with staff and create ownership of norms, values, mission, and expectations.

Support for teacher influence: classroom level means that teachers have influence in matters related to their classroom instructional practices.

Safe and orderly climate: orderly indicates a school that has policies in place that clearly articulate rules and codes of behavior, along with associated rewards and punishments. Further, students, faculty, and staff understand and consistently follow the policies.

Safe and orderly climate: safe indicates the school does not experience frequent disruptions due to student behavior.

Assessment and monitoring, means teachers place strong emphasis on using assessment results to determine students' progress toward learning critical content. Teachers make instructional decisions based on student assessment results.

Parent involvement is the degree to which there is a positive and productive relationship between the school's staff and parents; not only how involved parents are in the school, but also whether teachers feel supported by parents.

Table 1 A Comparison of School Factors Found in Other Studies to Those Used in This Study

Factor in McREL study (2005)	Heck, Larsen, and Marcoulides (1990)	Scheerens and Bosker (1997)	Used in this study
shared mission and goals	X	X	X
professional development			X
collaboration among teachers		X	X
support for teacher influence	X	X	X
safe and orderly climate	X	X	X
assessment and monitoring		X	X
parent involvement		X	X
instructional guidance*			
organizational change*	X	X	
clarity of learning goals*			
opportunity to learn cognitively challenging content*		X	
individualization [responsive instruction]*	X	X	
academic press for achievement*		X	

* factors not included in this study for lack of SASS items

The study addressed the following research question:

- How do the highest- and lowest-performing, high-needs high schools differ in terms of school-level factors hypothesized to be related to school performance?

How do the Highest- and Lowest-performing High-needs High Schools Differ on School-Level Factors that are Hypothesized to be Related to Student Performance?

This study examined differences between teachers' perceptions about school-level factors in both the highest- and lowest-performing, high-needs high schools. To establish the appropriate sample for use in the analyses, we began with the complete SASS dataset with the ultimate goal of identifying regular classroom teachers who taught in traditional high schools that were both high-needs and either 'highest-performing' or 'lowest-performing.' To identify "high-needs" schools, we selected high schools from the SASS dataset that had 50 percent or more of their students eligible for free or reduced-price lunch[2]. To identify the highest-performing and lowest-performing groups of schools we used the principal's rating of school performance, as provided by responses to a survey question on school performance on the principal's survey. Finally, we selected schools that did not have specific admissions criteria or that did not serve special populations (see Appendix E). These criteria were all identified through the use of responses to survey items in the SASS dataset.

The final analytic sample included 248 of the highest- performing schools, with 846 teachers, and 40 of the lowest performing schools, with 127 teachers[3]. Before conducting further analyses, we examined background differences in percent of students eligible for free or reduced-price lunches, percent of minority students, percent of LEP students, and urban/rural location between the high-performing and low-performing high-needs samples. We also examined differences in the rigor of state performance between the high- and low-performing schools. Box B shows the results of those comparisons.

[2] 50% of students or greater in a school eligible to receive free and reduced price lunch was chosen based on how the Department of Education defines schoolwide poverty for Title I. In the late 1990s, the schoolwide definition of poverty presented by Title I stipulates that "At least 50% of the children enrolled in the school or residing in the school attendance area are from low-income families" (Department of Education, 2008). While this definition changed in July 2004 to 40 percent (Federal Register 2004), we chose the more strict definition based on the rationale that it was the definition in place when the data were collected and the more rigorous definition would reduce the number of schools that might be misclassified. The SASS data provide numbers of students eligible for participation in the school lunch program as a poverty measure.

[3] The final sample resulted in unequal sample sizes for the two groups, based on the principal's responses to the survey items. Many more principals indicated that schools met all of their state and district standards than those who indicated their schools met none of their state or district standards. An additional 462 principals selected responses that their schools met most or some of the standards. These are not included in this study. The reader should also note that more than one teacher responded from a single school. To prevent this from biasing the results we used analyses that took this into account. See Appendix D.

Box B

Descriptive Comparison of HPHN and LPHN Schools. Descriptive and mean difference analyses were conducted on the percentage of students receiving free or reduced price lunch, the percentage of minority students, the percentage of LEP students, and urban/rural location of the school to determine if particular locations or student characteristics were unequally represented in either the low- or high-performing portion of the sample. For most of the demographic categories, when alpha was set at 0.01, there were no significant differences. However, there was a higher percentage of urban schools in the LPHN group (50 percent vs. 19 percent) and also a higher percentage of minority students (76 percent vs. 43 percent). If alpha is set at the traditional 0.05 level, additional significant differences are found. We carefully examined the influence of these differences in the results. (For more details see Appendix C)We also looked at the differences in the two groups to compare the rigor of the state standards. Our hypothesis was that more LPHN schools would come from states with more rigorous standards and more HPHN schools would come from states with less rigorous standards. We found no significant differences on these comparisons.

Next, researchers examined the mean differences between the highest- and lowest performing high-needs high schools for each of the nine factors. There were statistically significant differences between HPHN and LPHN schools on eight of the nine factors (see Table 2). Effect sizes were then calculated to express the differences between the HPHN and LPHH schools. The largest of these differences was for *safe climate,* with an effect size of 1.29 followed by *parent involvement* (0.70), *support for teacher influence* both at the *school* (0.37) and *classroom levels* (0.33), *orderly climate* (0.37), *assessment and monitoring (0.27)* and *collaboration (0.27).* The remaining factor showed no statistically significant difference between HPHN and LPHN schools[4].

[4] Although this report provides an indication of what differentiates HPHN from LPHN high schools, the design of the study does not warrant claims that the implementation of any of the factors will produce a higher performing school. Additionally, the study was limited to data from the Schools and Staffing Survey. The amount of instructional time, individualization of instructional practices, and the degree to which high expectations were present, all of which may be important in differentiating high- and low-performing schools, were not represented within the SASS data and were therefore not studied.

Table 2: Means and Effect Sizes for Nine School-Level Factors Hypothesized to be Related to Student Performance: High- vs. Low-performing High-needs High Schools.

	HPHN Mean (SD) (N=846)[5]	LPHN Mean (SD) (N=127)	F Value	Effect Sizes
shared mission and goals	0.02 (0.93)	-0.15 (0.81)	4.59*	0.19
professional development (HPHN N=130/LPHN N = 26)	-0.01 (0.83)	0.07 (0.94)	0.15	-0.09
collaboration	0.03 (0.93)	-0.22 (0.99)	7.70**	*0.27*
teacher influence – school	0.04 (0.90)	-0.29 (0.85)	17.13**	*0.37*
teacher influence – classroom	0.04 (0.93)	-0.27 (1.08)	9.30**	*0.33*
safe climate	0.12 (0.70)	-0.77 (0.64)	190.93**	*1.29*
orderly climate	0.04 (0.93)	-0.30 (0.92)	14.09**	*0.37*
assessment and monitoring	0.03 (0.93)	-0.22 (1.03)	6.90**	*0.27*
parent involvement	0.08 (0.87)	-0.53 (0.82)	57.46**	0.70

Note: α= .05 *p<.05, **p<.01

Table 3 presents the seven factors that have effect sizes greater than 0.2 in descending order by effect size. Effect sizes of 0.2 – 0.5 are considered small, 0.5 to 0.8 considered medium, and greater than 0.8 considered large.

Table 3. Effect Sizes in Descending Order for Seven School Level Factors with Greater than 0.20 Effect Sizes

	HPHN Mean (SD) N=846	LPHN Mean (SD) N=127	Effect Sizes
safe climate	0.12 (0.70)	-0.77 (0.64)	*1.29*
parent involvement	0.08 (0.87)	-0.53 (0.82)	*0.70*
orderly climate	0.04 (0.93)	-0.30 (0.92)	*0.37*
teacher influence – school	0.04 (0.90)	-0.29 (0.85)	*0.37*
teacher influence – classroom	0.04 (0.93)	-0.27 (1.08)	*0.33*
assessment and monitoring	0.03 (0.93)	-0.22 (1.03)	*0.27*
collaboration	0.03 (0.93)	-0.22 (0.99)	*0.27*

[5] Sample sizes for factors were the same with the exception of Professional Development which had missing data due to the nature of the items.

DISCUSSION

This study was designed to identify factors that differentiate highest-performing, high-needs high schools and lowest-performing, high-needs high schools, using a nationally available dataset. While the study could not show these factors to have a causal influence on school performance, several factors were found to be more prevalent in the HPHN schools. This information is expected to contribute to a better understanding of the differences between HPHN and LPHN high schools for policy makers and educators attempting to improve high-needs high schools. Meaningful differences between HPHN and LPHN schools were found for seven of the nine factors.

The differences in *safe climate* and *orderly climate* indicate that a HPHN school is more likely to be free of instructional interruptions, including those associated with misbehavior (safe climate) It also is more likely to enforce behavior policies and provide rewards and punishments (orderly climate). This result is similar to school research done at the elementary school level that indicates a school environment characterized by a respect for others, with fewer disciplinary issues and disruptions promotes a more academic climate (Evers & Bacon, 1995, McREL, 2005). This type of environment may allow students and teachers to be more productive and focused on academic goals. Raudenbush and Bryk (1989) also found a relationship between schools with fewer disciplinary actions and disruptions and smaller academic differences between white and minority students.

Our analysis of the *parent involvement* factor indicates that teachers felt that there were higher levels of parental support and a more positive relationship between parents and teachers in HPHN schools than in LPHN schools. It may seem that parental involvement would be more likely to differ at the elementary school level, given that parents of younger children may be more inclined to participate in their child's education. Indeed, past studies have found parent involvement to be particularly important in the early grades (McREL, 2005). However, the current research as well as other research in this area (Fan & Chen, 2001; Jeynes, 2007) supports the notion that positive parental and teacher relationships coupled with parental support of the school, remain an important factor in a student's high school education.

Data on *teacher influence* at both the classroom and school level and the degree of *collaboration* among teachers indicate that teachers in high-performing schools report that they have more influence in a range of aspects of school practice and policy. They also have more autonomy in their classrooms to make curricular decisions and select instructional practices. This supports the notion that in schools where teachers' professionalism is valued, there is a degree of autonomy in their classrooms, and free discussions between teachers supported. These teachers may also be able to adjust their teaching to better serve the needs of their students. This supports previous research conducted by Cooper, Ponder, Merritt, and Matthews (2005), which concluded that a school focus on collaborative leadership, as opposed to schools in which principals micromanage details, was related to school success.

The analysis of *assessment and monitoring* indicates that teachers in HPHN schools were more likely to use assessment data not only to monitor individual student performance, but also to adjust curriculum and to identify areas in their own content knowledge that they might need to

strengthen. Previous research indicates that use of assessment data can be critical to improving achievement levels in schools with high percentages of minority and high poverty students (Railsback, Reed, & Boss, 2001 and Dolejs, 2006). In other words, information presented in assessment results can provide important information to help a teacher identify his or her students' strengths and weaknesses and thereby implement the most effective curriculum and monitor performance.

CONCLUSIONS

Overall, the results suggest that HPHN high schools are significantly different in a number of ways from LPHN high schools. HPHN teachers reported more favorable perceptions for a number of the factors as represented by the SASS survey items. These teachers reported that their schools were safer and more disciplined. Parents in these schools were more supportive; the relationships with teachers were more likely to be positive and productive.

STUDY LIMITATIONS

Although this study describes what differentiates HPHN from LPHN high schools, there are limitations that apply to the conclusions. First, the design of the study does not permit causal claims; that is, it is not possible to conclude that implementation of any of the factors will produce a higher performing school. Future work might use a randomized controlled trial with manipulation of one or more of the factors to result in statements about changes that lead to improvement in high needs high schools.

Additionally, the study was limited to data from the SASS survey items and therefore the breadth and depth of the factor definitions were limited. For example, the factor that measured professional development was comprised of items that focused on teachers' perceived usefulness of professional development, resulting in a professional development factor that was limited in scope. If additional items had been included that measured such things as rigor or quality of professional development, differences between the HPHN and LPHN teachers could have been examined. In addition, we were unable to examine instructional factors such as instructional time, individualization of instructional practices, or how high expectations might affect school performance. These factors appear to be related to HPHN elementary schools (McREL, 2005) and may potentially be important for high schools as well.

Finally, we checked for whether the differences in demographics between HPHN and LPHN schools might have by themselves produced the differences in factor scores. We believe they do not (see Appendices C and D). Nonetheless, there are significant differences in demographic characteristics that do influence the results in ways we are unable to measure using the analytic methods we have selected. However, for schools in districts with standards that are more stringent than the state standards, mis-assignment of schools to performance level could have occurred. In other words, in the case of a school with stringent district standards in an easy standards state, if the school did not meet the district standards the principal would have responded that the school had met *no* standards, and that school would have been assigned to LPHN. If however, the principal responded on the basis that the school passed all state standards, the school would have

been assigned to HPHN (regardless of whether or not it passed the more stringent district standards). Thus, some HPHN schools may have been assigned LPHN status; however, this would be a small group, given that only nine schools with easy state performance standards were categorized as low-performing. If a state had average or stringent state standards, even relatively high district standards would not have been mis-aligned.

Additional Research

McREL researchers believe that this study provides information on substantial differences in the characteristics between HPHN and LPHN schools. This study provides descriptive information and a potential framework to serve as a foundation for future research. In fact, additional research is crucial to understanding and expanding upon these findings. Future studies could employ designs to test hypotheses about the influence of school factors, or could use individual student data (rather than teacher and principal perceptions) to develop a hierarchical model based on separate student, teacher and school level data which might more clearly link individual student achievement to teacher and school characteristics and to control for demographic differences. Such research would not indicate causes of improvement in student achievement. Research using experimental designs to link factors as well as combinations of factors to achievement is needed in order to know how changes might lead to improvement in high-needs high schools.

APPENDIX A: IDENTIFICATION OF SCHOOL LEVEL FACTORS FOR THE STUDY

A 2005, McREL study of high-needs elementary schools identified 13 factors that differentiate high-performing, high-needs schools from low-performing, high-needs schools. The current study sought to align survey items from the SASS with these 13 factors. Matching survey items were found for only seven of the original 13 factors. The final factors, which differed slightly from the original factors given the nature of the SASS data, are defined below.

BACKGROUND FOR THE THIRTEEN FACTORS

Effective schools research generally examines academic success broadly, across different socioeconomic contexts. Although several researchers have examined effectiveness in high-poverty schools (for example, Brookover, Beady, Flood, Schweitzer, & Wisenbaker, 1979; Teddlie, Stringfield, Wimpelberg, & Kirby, 1989), research guided by the emphases of NCLB, such as performance measured by assessments linked to standards, is sparse. A predominant methodology in effective schools research has been case study research; however, these studies have not compared demographically comparable low- and high-performing schools. Thus, potential explanations for success in these schools remain unverified. This study's methodology, while not a case study, does lack the rigor needed to determine which factors influence success.

A number of studies were used to identify the original 13 factors described in the 2005 McREL study. Heck, Larsen, and Marcoulides (1990) surveyed principals and a random selection of their teachers in 118 public elementary and high schools in California that for three years scored above or below their comparison band. They used data from schools in which the principal and at least four teachers responded. This research built on the body of research from instructional leadership and school effectiveness. Their study suggested broad areas under which factors could be lodged. They linked leadership through two mediating variables, school climate and school instructional organization, to student achievement. In a later work, Hallinger, Bickman, and Davis (1996) is a secondary analysis of data collected from 87 elementary schools over three years in Tennessee that participated in the School Incentives Improvement Program. Each year, more than 1300 teachers (>90 percent) completed questionnaires. The study identified grouping for instruction, clear mission, student opportunity to learn, and high teacher expectations as factors.

Scheerens and Bosker (1997, pp.135–136) in their 1997 book *The Foundations of Educational Effectiveness*, defined school effectiveness using student achievement as an outcome measure, adjusted for prior achievement and/or relevant student background characteristics. They review both conceptual work and empirical studies to derive their set of effectiveness enhancing factors. They posit 13 general effectiveness enhancing factors: 1) achievement, orientation, and high expectations; 2) educational leadership; 3) consensus and cohesion among staff; 4) curriculum quality/opportunity to learn; 5) school climate; 6) evaluative potential (for example. monitoring pupil's progress and use of evaluative results) , 7) parental involvement, 8) classroom climate, 9) effective learning time; 10) structured instruction; 11) independent learning; 12) differentiation (e.g. special attention for pupils at risk); and 13) reinforcement and feedback.

Selection of SASS Items to Measure Factors

We categorized items from the SASS teacher survey into the 13 factors from the original McREL study; however, we found items to match only seven of the 13. Further, these survey items didn't necessarily align to all of the original aspects of those seven.[6] In many cases, the survey items only covered a part of the original factor. We then asked a content alignment specialist external to the project to review the assignment. Items aligned to the seven were confirmed as appropriate. We then used item analysis to determine how the items loaded on the factors. Using the theoretical work of the content alignment, as well as the item loadings, we finalized the set of items to be used for each factor. (Subsequent factor analysis separated two of the seven factors into two each as noted below. Thus, the number of factors in final set used in the study was nine.)

The factors that are hypothesized to be related to school performance and the SASS item concepts that aligned to them are listed below. (See Appendix C for a specific item map of the items selected from the SASS survey for the seven factors).

Factor 1. Shared Mission and Goals

Three SASS questions measure shared mission and goals: 1) the degree to which principals communicate expectations, 2) the degree to which there is a shared belief system, and 3) the degree to which the principal communicates the vision he or she has for the school.

Thus, *shared mission and goals* deals with framing, communicating, and enlisting engagement in a common mission. The principal's values and beliefs are known to teachers and are aligned with the mission and goals.

Factor 2. Professional Development

The four items in the SASS dataset that evaluate the quality of professional development focus on its perceived usefulness. Teachers were asked if they participated in professional development specific to 1) the content they taught, 2) the use of computers for instruction, 3) reading instruction, and 4) classroom management. If they had participated, they were asked the degree to which the professional development was useful.

Thus, *professional development* focuses on the usefulness of the professional development that teachers receive.

Factor 3. Collaboration

Five questions were used to measure the construct of Collaboration. These questions include the degree to which teachers: 1) collaborated with one another, 2) participate in scheduled collaborations, 3) whether they were observed in their classrooms by other teachers, 4) whether

[6] A complete list of items for each factor can be found in Appendix C.

they coach or mentor others, and 5) how much they coordinated the content in their classrooms across other teachers' classrooms.

Thus, *collaboration* fosters the sharing of work and expertise among teachers and creates a sense of affiliation and support. Teachers are encouraged to work together in one another's classrooms by observing, providing feedback, and coordinating instruction.

Factor 4 and 5. Support for Teacher Influence

School-level teacher influence was defined by six items that focused on teachers' ability to affect their school's policies. Specifically, teachers were asked how much influence they had on 1) setting performance standards, 2) selecting curriculum, 3) determining professional development content, 4) hiring decisions, 5) setting discipline policies and 6) establishing the budget. Teachers were also asked questions that were specific to the degree to which they were allowed to direct their own classrooms. These questions included 1) control over selecting instructional materials and content, 2) deciding on appropriate instructional techniques, and 3) determining how to best evaluate and grade students.

Thus, *support for teacher influence* emphasizes shared leadership responsibilities among principals, administrators, and staff, and shared ownership of norms, values, mission, and expectations. Teacher influence also incorporates the degree to which teachers have influence in matters related to their classroom instructional practices (for example, in selecting textbooks and curriculum).

In subsequent factor analysis the support for teacher influence split into a school factor and a classroom factor.

Factor 6 & 7. Safe and Orderly Climate

Six SASS questions address safe and orderly climate. Two questions ask about the extent to which certain types of problems (for example, tardiness, bullying, disorder in classrooms, and misbehavior) occur in the school. Two other questions are related to the principal or teachers enforcing rules for student conduct (these two questions were later separated out to become *orderly climate*). The fifth question asks whether student misbehavior interferes with teaching and the sixth whether student tardiness and cutting classes interferes with teaching.

Thus, a *safe and orderly climate* is most frequently characterized as one that supports school safety and provides a disciplined environment. A school with an orderly climate has policies in place that clearly articulate rules and codes of behavior, along with associated rewards and punishments. Further, students, faculty, and staff understand and consistently follow the policies. A safe school does not experience frequent student disruptive behavior.

In subsequent factor analysis safe and orderly climate split into safe and orderly.

Factor 8. Assessment and Monitoring

Three SASS questions addressed teachers' use of assessment and monitoring. These questions measured the extent that teachers used information from state or district student achievement test scores in planning and teaching.

In *assessment and monitoring*, teachers place strong emphasis on using assessment results to determine students' progress toward learning critical content. Teachers make instructional decisions on student assessment results.

Factor 9. Parent Involvement

There are two questions that address parent involvement. The first is the degree to which lack of parental involvement is an issue at the school and the second is the degree to which the teacher feels supported by parents for the work that they do.

Parent involvement, then, is the degree to which there is a positive and productive relationship between the school's staff and parents; not only how involved parents are in the school but also whether teachers feel supported by parents.

Other Studies of the Influence of Factors on Achievement

Extant research on factors that influence student achievement has been conducted largely as a part of the evaluations of comprehensive school reform (CSR) models. These models range from relatively straightforward instruction and curriculum in a given content area, to systemic, all-encompassing school reform models. The Comprehensive Reform School Program funded by Congress defined 11 components of a CSR program (U.S. Department of Education, 2002) The 11 components are a mixture of criteria (e.g. being based on scientific research) and practices (e.g. conducting annual evaluations) as well as four components that relate to the factors of this study. These are: teacher and staff professional development (which relates to the current study's factor 3: professional development), support for shared leadership (which relates to factor 4: support for teacher influence: school), measurable goals and benchmarks (which relates to factor 8: assessment and monitoring), and meaningful involvement of parents (which relates to factor: 9 parent involvement).

Borman, Hewes, Overman, & Brown (2003)[7] conducted a meta-analysis of the research conducted on 29 of these Comprehensive School Reform models. They coded three of the four CSR components that related to factors in this study (all except support for shared leadership) but used the results to better describe the set of 29 CSR models and not to analyze the influence of those components on achievement. Of their six programs with the strongest evidence, two programs—

[7] Borman, Hewes, Overman, & Brown (2003) have carefully assessed the quality of the research base for the models and identified the top six (in 2003-2004) for strength of evidence of effectiveness. The What Works Clearinghouse has reported on the evidence for four models, two of which are in Borman's top six. The Comprehensive School Reform Quality Center has a top group of seven models three of which have been reviewed by the What Works Clearinghouse and three of which are in Borman's top six.

Expeditionary Learning and the School Development Program—included two of the three components that were coded. When we examined two accessible research studies of the five that met Borman's inclusion criteria for the School Development Program, we found that the studies did not examine the influence of the factors on student achievement, but the studies treated them as independent program effects, e.g. the school climate improved parallel with improvement in student achievement. This result is similar to the result from this study; the greater presence of certain factors in high-performing high schools is an independent effect. Instituting the factor could not be said to improve school performance.

Studies are now underway under the aegis of IES that are designed to meet scientific standards of study quality. Some of these may include a rigorous look at how factors such as the ones in this study influence student performance.

APPENDIX B: ITEM MAP OF SASS QUESTIONS TO CONSTRUCTS

Factor	SASS Survey Item #	Questions
Shared Mission and Goals	63a	The principal lets staff members know what is expected of them.
	63j	Most of my colleagues share my beliefs and values about what the central mission of the school should be.
	63k	The principal knows what kind of school he/she wants and has communicated it to the staff.
Professional Development[8]	40	Overall, how useful were these professional development activities (in the past 12 months) to you? c. specific to and concentrating on the content of subject you teach in the last 12 months.
	41	c. focused on computers for instruction in last 12 months
	42	c. focused on reading instruction in last 12 months
	43	c. focused on student discipline and management in the classroom in last 12 months
Collaboration	47	In the past 12 months, did you do any of the following? b. Participate in regularly scheduled collaboration with other teachers on issues of instruction c. Observe, or be observed by, other teachers in your classroom (for at least 10 minutes) d. Act as a coach or mentor to other teachers or staff in your school, or receive coaching or mentoring
	63	l. There is a great deal of cooperative effort among the staff members.
	63	r. I make a conscious effort to coordinate the content of my courses with that of other teachers.
Support for Teacher Influence (School)	61	How much actual influence do you think teachers have over school policy AT THIS SCHOOL in each of the following areas? a. Setting performance standards for students at this school, b. Establishing curriculum, c. Determining the content of in-service professional development programs, e. Hiring new full-time teachers, f. Setting discipline policy,

[8] Not all teachers received all types of professional development. If a teacher did not receive that specific professional development, the data for that item became system missing.

Factor	SASS Survey Item #	Questions
		g. Deciding how the school budget will be spent
Support for Teacher Influence (Classroom)	62	How much actual control do you have IN YOUR CLASSROOM at this school over the following areas of your planning and teaching? a. Selecting textbooks and other instructional materials, b. Selecting content, topics, and skills to be taught, c. Selecting teaching techniques, d. Evaluating and grading students, e. Disciplining students, f. Determining the amount of homework to be assigned
Safe Climate	63	d. The level of student misbehavior in this school (such as noise, horseplay or fighting in the halls, cafeteria or student lounge) interferes with my teaching.
	63	s. The amount of student tardiness and class cutting in this school interferes with my teaching.
	64	To the best of your knowledge how often do the following types of problems occur with students at this school? a. Physical conflicts among students b. Robbery or theft c. Vandalism f. Possession of weapons g. Physical abuse of teachers h. Student racial tensions i. Student bullying j. Student verbal abuse of teachers k. Widespread disorder in classrooms l. Student acts of disrespect for teachers m. Gang activities
	65	To what extent is each of the following a problem in this school? a. Student tardiness b. Student absenteeism c. Student class cutting
Orderly Climate	63	h. My principal enforces school rules for student conduct and backs me up when I need it.
	63	i. Rules for student behavior are consistently enforced by teachers in this school, even for students who are not in their classes.
Assessment &	55	To what extent do you use the information from your students'

Factor	SASS Survey Item #	Questions
Monitoring		state or district achievement test scores: a. to group students into different instructional groups by achievement or ability? b. to assess areas where you need to strengthen your content knowledge or teaching practice? c. to adjust your curriculum in areas where your students encountered problems?
Parent Involvement	63	e. I receive a great deal of support from parents for the work I do.
	65	To what extent is each of the following a problem in this school: h. Lack of parental involvement

APPENDIX C: DESCRIPTIVE COMPARISONS OF HPHN VS. LPHN SCHOOLS

Descriptive and mean difference analyses were conducted on student demographics (percent of students eligible for free or reduced price lunch, percent of minority students, and percent of LEP student) and locale (e.g. urban, rural) to determine if particular locations or student characteristics were unequally represented in either the low- or high-performing portion of the sample. Multiple comparisons of the means were made to ensure the comparability of the HPHN and LPHN samples.

We decided to set alpha at 0.01 for the entire study based on the number of tests of statistical significance tests that were planned (many of which were correlated). For example, with regard to demographics, we conducted 3 t-tests and six Chi-square tests. Using the Bonferroni adjustment (alpha divided by the number of tests (in this case 0.05/3 or 0.05/6), we obtained alphas of 0.017 and 0.008 (Abdi, 2007). For the analyses of the mean differences of the nine factor scores, the Bonferroni adjustment resulted in an alpha of 0.006. We therefore decided to set alpha at 0.01, a midpoint between the three alphas. We also note those differences that were significant at the more conventional alpha level of 0.05.

We first looked at location for LPHN and HPHN schools (Table C-1). The difference between the two groups is significant with a higher proportion of LPHN schools in urban areas.

Table C-1. Comparison of Locale of the High Schools

Performance level / School Locale	HPHN	LPHN	Total	Chi-Square
	Number of schools (%)			
Urban	48 (19%)	20 (50%)	68 (24%)	13.87**
Suburban	60 (24%)	8 (20%)	68 (24%)	0.24
Rural	140 (56%)	12 (30%)	152 (53%)	4.55*
Total	248 (100%)	40 (100%)	288(100%)	

α=.01. *p<.05. ** p<.01.

We then looked at HPHN/LPHN differences in school-level student demographics. The only significant difference at the 0.01 level was for the percentage of racial and ethnic minorities with a higher percentage in LPHN schools (see Table C-2). We examined all of the significant differences by comparing means of subgroups of the two main groups, separating them into high and low groups, e.g. high-minority HPHN schools and low-minority HPHN schools and high-minority LPHN schools and low-minority LPHN schools. The HPHN schools continued to outperform the LPHN schools under the varying demographic conditions. (See Appendix D for the analyses.)

Table C-2. Comparison of School Level Student Demographics

Demographic Characteristic	Mean (SD) HPHN (n=248)	Mean (SD) LPHN (n=40)	T Statistic
Average Percentage racial/ethnic minority	43.29	76.20	5.87**
	(33.4)	(29.3)	
Average Percentage limited-English-proficiency	3.41	10.90	2.43*
	(6.6)	(19.3)	
Average Percentage National School Lunch Program	70.74	77.78	2.52*
	(16.3)	(16.9)	

α=.01. *p<.05. ** p<.01.

Finally, the stringency of state standards was examined to determine if there was an equal proportion for the two groups of schools of states where state performance standards may be considered either 'easy', 'average' or 'stringent'. Categorization of the states came from the report "Mapping 2005 State Proficiency Standards Onto the NAEP Scales" (NCES, 2007) which mapped state student assessment data onto NAEP scale scores in an attempt to establish the rigor of state standards relative to NAEP standards. The intent was to offer "a credible indicator of the relative stringency of the (state) standards" (p. 1). The NCES study accomplished this by comparing schools whose students took the NAEP and calculating the proportion of those students who met the state's proficiency standard relative to the proportion who were proficient on NAEP. Then, using the schools and students who took the NAEP, NAEP scores were estimated for the entire state. Finally, a point on the NAEP score scale was determined based on the "estimated proportion of students in the state scoring above that point equaled to the estimated proportion of students in the state meeting the state's own performance standard" (NCES 2007, p. 3). All states were mapped onto the NAEP scale scores relative to the estimated rigor of their state's performance standards. For the purposes of this study, we focused on the assessment of 8th grade math and reading scores, the highest grade included in the report. We have noted where states in our study fell relative to each other on the NAEP scale scores to be able to examine whether the sample was proportionate between the three stringency categories. Consideration was given to the scores that NAEP uses to define its proficiency levels. States that fell below a score equivalent of 240 on the NAEP reading assessment (p. 13, Figure 3) or 260 on the mathematics assessment (p. 15, Figure 5) were classified as states with "easy" performance standards. This level was approximately at the NAEP 'basic' score. States that fell between 240 and 255 on the NAEP reading equivalent scores and between 260 and 280 on the NAEP mathematics equivalent scores were classified as states with "average" standards[9]. States that were above the score of 255 on reading or 280 on mathematics for the NAEP score equivalents were classified as "stringent." Once states were

[9] Ideally, we would have liked to use the NAEP proficiency score to differentiate between the "average" and "stringent" categories. However, no states in our sample fell above the reading NAEP proficiency cut score. It was therefore necessary to use a somewhat lower score (i.e. compared to the NAEP proficiency cut scores for reading and mathematics) to create the second and third groups: average and stringent.

classified as easy, average, or stringent, the schools were separated into the six categories (see Table C-3). An analysis conducted to examine whether the proportion of easy, average, and stringent states in high- vs. low-performing schools was different proved to be non-significant at the 0.01 level ($\chi^2 = 5.84$, $p<0.01$).

Table C-3: Comparison of Groups on Rigor of State Performance Standards

Performance level	HPHN	LPHN	Total
Level of State Assessment Rigor	**Number of schools (%)**		
Stringent Performance Standards	65 (29%)	5 (16%)	70 (27%)
Average Performance Standards	75 (33%)	17 (55%)	92 (36%)
Easy Performance Standards	87 (38%)	9 (29%)	96 (37%)
Total[10]	227 (100%)	31(100%)	258(100%)

The results of the demographic descriptive analyses indicated statistical differences at the 0.01 level on two of the demographic characteristics between the HPHN and LPHN samples. These were higher percentages of schools with minority students and a higher percentage of schools in urban areas. And at the 0.05 level on key characteristics such as percentages of students in poverty or with limited English proficiency. There were no differences found on the rigor of the states' standards. These differences in demographic characteristics may have contributed to but were not the sole cause of the differences we found between HPHN and LPHN high-needs high schools on the factors. We are not able to determine to what degree that is the case.

[10] Based on the NCES (2007) report of NAEP 2005 data, 9 states had no 8th grade data for both math and reading. Therefore, there were 31 schools in the 9 states that were not included in the analysis.

APPENDIX D: ANALYSES

INITIAL ANALYSES

Using the identification from Item 28 for HPHN and LPHN schools, there were 846 teachers in high-performing high schools and 127 teachers in low-performing high schools Power analyses were then conducted to determine if the sample and predicted effect sizes would be sufficient to detect group differences. Because the primary analyses used data at the teacher level, numbers of teachers were used to calculate power. We calculated power based on the number of teachers in these two categories using an alpha of 0.01 and an effect size of 0.515 based on an average of effect sizes generated from Englert, Fries, Martin-Glenn, and Douglas (2007). The estimated power was 0.99, which gave us confidence that we would be able to find group mean differences (including effect sizes) on the factor scores if they were present.

FACTOR ANALYSES OF SASS ITEMS – PRINCIPAL COMPONENT ANALYSES

In order for the researchers to examine if the theoretical matching of survey items to the factors was justified statistically, we conducted a series of factor analyses. This process also allowed us to generate factor scores for each teacher in the dataset. The factor scores were generated using the regression approach in SPSS, which results in a distribution of factor scores with a mean of zero and a standard deviation of one. Using factor scores for analyzing mean differences between the HPHN and LPHN groups was more appropriate than using item means, as factor scores provide "estimates of the values that would be produced if the underlying constructs could be measured directly" (Tabachnick & Fidell, 2007, p. 650).

We used principal components analysis (PCA) as the extraction method. As Fabrigar, Wegener, MacCallum and Strahan (1999) note, "the objective of PCA is to determine the linear combinations of the measured variables that retain as much information from the original measured variables as possible" (p. 275). A varimax rotation was used to increase the interpretability of the analysis.

Determinations were made regarding the number of factors to retain based on a combination of statistical data and theoretical information. We first examined the eigenvalues. A generally accepted criterion is retaining only those eigenvalues greater than 1.0 (Stevens 1996). However, this can lead to retaining values that might not be theoretically meaningful. In order to confirm the importance of a factor we also examined the percent variance explained by the factor as well as examining which items comprised each factor to determine if there was a sufficient theoretical justification for the factor. Factors that had an eigenvalue of 1.0 or greater were examined for both percent variance explained and theoretical value.

The results for each study factor in terms of the components that were identified in the principal components analysis are discussed below.

- *Parent involvement*
  The results for this factor showed one component (#1) with an eigenvalue

greater than 1.0 thus confirming our hypothesis that the two items represented one factor (see Table D-1). Coefficient Alpha for these two items was a relatively low 0.61 which was most likely due to the small number items comprising the factor. It is important to note that we realized we could have taken an average of the two items in this component and standardized them. However, following a discussion with our statistical consultant[11], we determined that a factor analysis was appropriate, as these items were an imperfect measure of the construct and we believed that there was shared variance between the items.

Table D-1. Percent Variance Explained and Eigenvalues – Parent Involvement

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.442	72.099	72.099	1.442	72.099	72.099
2	.558	27.901	100.000			

- *Safe and orderly climate*
  These 9 items together had a relatively high coefficient alpha of 0.87. The factor analyses showed that eigenvalues of the first two components (#1, #2) were greater than 1.0, so we decided to examine the value of splitting the factor into two separate factors. The percent of variance explained was 51.92 and 12.22 for the first and second factor which increased to 44.98 and 19.16 for the rotated solution (see Table D-2). Upon review of the items that comprised each component (Table D-2a), we saw that they were loading on two constructs that we believed to be theoretically distinct. Therefore, we decided to break this into two factors: safe (coefficient alpha 0.89) and orderly (coefficient alpha 0.66). Two items measured the orderly component and eight questions measured the safe component.

Table D-2. Percent Variance Explained and Eigenvalues – Safe and Orderly Climate

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.673	51.918	51.918	4.673	51.918	51.918	4.048	44.983	44.983
2	1.100	12.220	64.138	1.100	12.220	64.138	1.724	19.155	64.138
3	.816	9.061	73.199						
4	.687	7.632	80.831						
5	.504	5.604	86.435						
6	.459	5.101	91.536						
7	.302	3.353	94.889						
8	.285	3.162	98.050						

[11] Statistical consultant to the project, Dr. Edward Wiley, Assistant Professor at the University of Colorado was consulted on this issue.

9	.175	1.950	100.000						

Table D-2a. Factor Loadings for Rotated Component Matrix

Factor	Item	Component Loadings	
		1	2
orderly	63 h My principal enforces school rules for student conduct and backs me up when I need it.	.121	.864
	63 i Rules for student behavior are consistently enforced by teachers in this school, even for students who are not in their classes.	.262	.792
safe	63 d To what extent do you agree or disagree with each of the following statements? The level of student misbehavior in this school (such as noise, horseplay or fighting in the halls, cafeteria or student lounge) interferes with my teaching.	.598	.334
	64 (a, f, g, i, m) To the best of your knowledge how often do the following types of problems occur with students at this school? (Physical conflicts among students, Possession of weapons, Physical abuse of teachers, Student bullying, Gang activities)	.738	.231
	64 (b, c, h, j, k, l) To the best of your knowledge how often do the following types of problems occur with students at this school? (Robbery or theft, Vandalism, Student racial tensions, Student verbal abuse of teachers, Widespread disorder in classrooms, Student acts of disrespect for teachers)	.752	.314
	63 s To what extent do you agree or disagree with each of the following statements? The amount of student tardiness and class cutting in this school interferes with my teaching.	.667	.207
	65 a To what extent is each of the following a problem in this school? Student tardiness	.816	.115
	65 b To what extent is each of the following a problem in this school? Student absenteeism	.815	.110
	65 c To what extent is each of the following a problem in this school? Student class cutting	.849	.137

- *Assessment and monitoring*
  The factor of assessment and monitoring was comprised of only one component with an eigenvalue greater than 1.0 (see Table D-3). The coefficient alpha was 0.83 which was relatively high given the small number of included items. We thus concluded that only a single factor was represented.

Table D-3. Percent Variance Explained and Eigenvalues – Assessment and Monitoring

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.242	74.718	74.718	2.242	74.718	74.718
2	.625	20.831	95.549			
3	.134	4.451	100.000			

- *Collaboration*
  Again, only one component had an eigenvalue greater than 1.0 and thus we concluded the data were adequately represented by that single factor (see Table

D-4). The coefficient alpha was low at 0.31. This might have been the result of the small number of items comprising the factor. An examination of the individual items comprising the factor showed that the coefficient would not have markedly increased with the deletion of any items. We decided to retain all of the items in the factor because together, they theoretically represented a more complete notion of collaboration. We also considered the relationship between reliability and power for this factor. In spite of the more stringent alpha which might have had a negative impact on the power of the statistical tests, we ultimately found that there were still statistically significant differences between HPHN and LPHN schools.

Table D-4. Percent Variance Explained and Eigenvalues – Collaboration

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.288	42.942	42.942	1.288	42.942	42.942
2	.977	32.575	75.517			
3	.734	24.483	100.000			

- *Professional development*
  For the factor of professional development, only one component had an eigenvalue greater than 1.0. We concluded that the data represented a single factor (see Table D-5). Additionally, the coefficient alpha was sufficient at 0.80.

Table D-5. Percent Variance Explained and Eigenvalues – Professional Development

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.515	62.863	62.863	2.515	62.863	62.863
2	.638	15.955	78.818			
3	.483	12.068	90.886			
4	.365	9.114	100.000			

- *Support for teacher influence*
  The coefficient alpha for the items was relatively high at 0.81. However, the results of the factor analysis showed two components had eigenvalues greater than 1.0 and the percent of variance accounted for was 34.3 and 17.9 respectively (see Tables D-6, D-6a). Upon review of the items, we determined that it would be appropriate to divide this factor; the first six items measuring teacher influence at the school level (coefficient alpha 0.80) and the next five items measuring teaching influence at the classroom level (coefficient alpha 0.77). We felt that teachers might have autonomy in their own classrooms but might not necessarily have school level influence so measuring these individually might give us additional insight regarding how these factors played out in high- versus low-performing schools.

Table D-6. Percent Variance Explained and Eigenvalues – Support for Teacher Influence

Component	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.772	34.288	34.288	3.038	27.617	27.617
2	1.960	17.882	52.109	2.694	24.492	52.109

Table D-6a. Factor Loadings for Rotated Component Matrix

Factor	Item	Component Loadings	
		1	2
Support for teacher influence: school	61.a. How much actual influence do you think teachers have over school policy AT THIS SCHOOL in each of the following areas? Setting performance standards for students at this school	.679	.237
	61.b. How much actual influence do you think teachers have over school policy AT THIS SCHOOL in each of the following areas? Establishing curriculum	.679	.289
	61.c. How much actual influence do you think teachers have over school policy AT THIS SCHOOL in each of the following areas? Determining the content of in-service professional development programs	.670	.169
	61.e. How much actual influence do you think teachers have over school policy AT THIS SCHOOL in each of the following areas? Hiring new full-time teachers	.675	-.003
	61.f. How much actual influence do you think teachers have over school policy AT THIS SCHOOL in each of the following areas? Setting discipline policy	.781	.106
	61.g. How much actual influence do you think teachers have over school policy AT THIS SCHOOL in each of the following areas? Deciding how the school budget will be spent	.702	-.043
Support for teacher influence: classroom	62.b. How much actual control do you have IN YOUR CLASSROOM at this school over the following areas of your planning and teaching? Selecting content, topics, and skills to be taught	.182	.600
	62.c. How much actual control do you have IN YOUR CLASSROOM at this school over the following areas of your planning and teaching? Selecting teaching techniques	.067	.797
	62.d. How much actual control do you have IN YOUR CLASSROOM at this school over the following areas of your planning and teaching? Evaluating and grading students	.040	.809
	62.e. How much actual control do you have IN YOUR CLASSROOM at this school over the following areas of your planning and teaching? Disciplining students	.264	.615
	62.f. How much actual control do you have IN YOUR CLASSROOM at this school over the following areas of your planning and teaching? Determining the amount of homework to be assigned	.016	.695

- *Shared mission and goals*
  There was only one component with an eigenvalue greater that 1.0 with 61.57 percent of the variance accounted for (see Table D-7) and the coefficient alpha was adequate given the small number of items in the factor at 0.68. We concluded that this solution adequately represented the data.

Table D-7. Percent Variance Explained and Eigenvalues – Shared Mission and Goals

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %

1	1.847	61.572	61.572	1.847	61.572	61.572
2	.774	25.798	87.370			
3	.379	12.630	100.000			

Ultimately, following our review of our results, we identified a total of 9 factors for our subsequent research question analyses.

ANALYSES OF MEAN DIFFERENCES ON FACTOR SCORES

In this study, teachers were 'nested' within schools resulting in teacher data that were not strictly independent. We considered whether the use of a multi-level model was necessary. On average, there were fewer than five teachers per school in our sample. Because there were relatively few teachers per school and McREL (2005) had found small Intraclass Correlation Coefficients (ICCs), indicating that only a small portion of the variance is accounted for between schools (0.01-0.12) a multi-level model was determined to be unnecessary. However, in order to account for the clustered design of our sample, Proc SurveyReg and Proc SurveyMeans procedures in SAS were used to generate accurate standard errors that would reflect the nested design.

The next step of the analyses was to examine how the factor scores differed between the highest- and lowest-performing groups. First, we ran overall mean differences on the factor scores for teachers in HPHN and LPHN schools. This resulted in eight of the nine factors being statistically significantly different at $p<0.05$. We then calculated effect sizes for each of the differences, weighted by sample size, which showed that there were small to medium effects present.

We next decided to examine if the differences could be attributed to demographic differences in the HPHN and LPHN samples. In other words for example, might the differences be attributable to the simple fact that there were significantly more high-minority and urban schools in the LPHN sample.

As indicated in Appendix C, the two groups (HPHN and LPHN) differed significantly at alpha <0.05 on four demographic characteristics: proportion of urban schools, percent of minority students, percent limited English proficient (LEP) students, and percentage of students qualified to receive free or reduced price lunch. We therefore compared the means for a subgroup from the HPHN and LPHN groups that represented the largest potential differences for the characteristics and thus would most conservatively identify where the greatest impact of the demographics might be. That is, we compared urban HPHN schools to non-urban LPHN schools, high-minority HPHN schools to low minority LPHN schools, high LEP HPHN schools to low LEP LPHN schools and high FRL HPHN schools to low FRL LPHN schools. Tables D-8 through D-11 present these figures. In the right-hand column we have noted the direction of the difference in these means. For the purposes of clarifying how the group means compared, we decided to take a conservative approach and examine the more 'at-risk' high performing group with the more 'low-risk' low performing group. From this comparison, we could better understand if the overall differences in the means presented in the original report were due to demographic differences. One would expect the HPHN group to continue to be higher or at least equal to the LPHN mean when these

groups are compared. If such is the case, we would infer that the demographics are not the sole source of the differences in the factor scores. Note: we have only examined the seven factors for which the differences represented an effect size of 0.20 or greater when considering the entire sample. The HPHN schools continued to outperform the LPHN schools under the varying demographic conditions.

Table D-8: Factor Scores Means and Effect Sizes by Demographics and Performance Level – Higher versus Lower Percentages of Students Receiving Free or Reduced Price Lunch

	Low FRL		High FRL		Group Comparison of Means [12]
	HP Mean	LP Mean	HP Mean	LP Mean	LPHN Low FRL vs. HPHN High FRL
Sample Size	**462**	**37**	**384**	**90**	
Safe Climate	0.33	-0.44	-0.14	-0.91	LPHN/LFRL<HPHN/HFRL
SD	0.64	0.61	0.78	1.37	
Parent Involvement	0.18	-0.52	-0.04	-0.54	LPHN/LFRL<HPHN/HFRL
SD	0.86	0.85	0.78	0.85	
Orderly Climate	0.10	-0.24	-0.02	-0.32	LPHN/LFRL<HPHN/HFRL
SD	0.86	1.03	0.98	0.85	
Teacher Influence – School	0.13	-0.17	-0.06	-0.35	LPHN/LFRL<HPHN/HFRL
SD	0.86	0.91	0.98	0.85	
Teacher Influence – Classroom	0.10	-0.17	-0.03	-0.31	LPHN/LFRL<HPHN/HFRL
SD	0.86	1.16	0.98	1.04	
Assessment and Monitoring	0.08	-0.31	-0.22	-0.19	LPHN/LFRL<HPHN/HFRL
SD	0.86	0.97	0.98	1.04	
Collaboration	0.05	-0.04	0.01	-0.30	LPHN/LFRL=HPHN/HFRL
SD	0.86	0.73	0.98	1.04	

Table D-9: Factor Scores Means and Effect Sizes by Demographics and Performance Level – Higher versus Lower Percentages of Minority Students

	Low minority		High Minority		Group Comparison of Means
	HP Mean	LP Mean	HP Mean	LP Mean	LPHN Low Minority vs. HPHN High Minority

[12] Means were said to be approximately equal if they were within 0.05 points of each other.

Sample Size	**462**	**16**	**384**	**111**	
Safe climate	0.49	-0.37	-0.34	-0.83	LPHN/LM=HPHN/HM
SD	0.64	0.52	0.78	0.63	
Parent involvement	0.38	-0.24	-0.28	-0.57	LPHN/LM=HPHN/HM
SD	0.86	0.92	0.78	0.84	
Orderly climate	0.02	-0.25	-0.03	-0.30	LPHN/LM<HPHN/HM
SD	0.86	1.04	0.98	0.95	
Teacher influence – School	0.19	-0.77	-0.13	-0.23	LPHN/LM<HPHN/HM
SD	0.86	0.92	0.98	0.84	
Teacher Influence – Classroom	0.06	-0.45	0.02	-0.24	LPHN/LM<HPHN/HM
SD	1.07	1.36	0.78	1.05	
Assessment and Monitoring	0.09	0.21	-0.04	-0.29	LPHN/LM>HPHN/HM
SD	0.86	0.88	0.98	1.05	
Collaboration	0.11	-0.26	-0.06	-0.22	LPHN/LM<HPHN/HM
SD	0.86	0.80	0.98	1.05	

Table D-10: Factor Scores Means and Effect Sizes by Demographics and Performance Level – Higher versus Lower Percentages of Limited English Proficient Students

	Low LEP		**High LEP**		**Group Comparison of Means**
	HP Mean	**LP Mean**	**HP Mean**	**LP Mean**	**LPHN low LEP vs. HPHN high LEP**
Sample Size	**394**	**34**	**452**	**93**	
Safe Climate	0.33	-0.70	-0.07	-0.80	LPHN/LLEP<HPHN/HLEP
SD	0.79	0.58	0.64	0.68	
Parent Involvement	0.16	-0.29	0.01	-0.62	LPHN/LLEP<HPHN/HLEP
SD	0.79	0.87	0.85	0.77	
Orderly Climate	0.04	-0.05	0.05	-0.38	LPHN/LLEP<HPHN/HLEP
SD	0.99	0.93	0.85	0.96	
Teacher Influence – School	-0.01	-0.24	0.09	-0.32	LPHN/LLEP<HPHN/HLEP
SD	0.79	1.11	0.85	0.77	
Teacher Influence – Classroom	-0.01	-0.49	0.09	-0.19	LPHN/LLEP<HPHN/HLEP
SD	0.99	1.46	0.85	0.96	
Assessment and Monitoring	0.06	-0.03	0.01	-0.30	LPHN/LLEP=HPHN/HLEP
SD	0.79	0.99	1.06	1.06	
Collaboration	-0.10	-0.10	0.15	-0.27	LPHN/LLEP<HPHN/HLEP

SD	0.99	0.93	0.85	0.96	

Table D-11: Factor Scores Means and Effect Sizes by School Locale and Performance Level

	Non-Urban		**Urban**		**Group Comparison of Means**
	HP Mean	**LP Mean**	**HP Mean**	**LP Mean**	**LPHN non-urban vs. HPHN Urban**
Sample Size	**676**	**68**	**170**	**59**	
Safe Climate	0.29	-0.54	-0.56	-1.04	LPHN/NU=HPHN/U
SD	0.78	0.66	0.78	0.61	
Parent Involvement	0.14	-0.12	-0.14	-0.55	LPHN/NU=HPHN/U
SD	0.78	0.74	0.91	0.92	
Orderly Climate	0.08	-0.42	-0.11	-0.15	LPHN/NU<HPHN/U
SD	1.04	0.82	0.91	1.00	
Teacher Influence – School	0.04	-0.39	0.06	-0.18	LPHN/NU<HPHN/U
SD	0.78	0.82	0.91	0.92	
Teacher Influence – Classroom	0.07	-0.34	-0.07	-0.19	LPHN/NU<HPHN/U
SD	1.04	1.32	0.91	0.92	
Assessment and Monitoring	0.08	-0.38	-0.15	-0.05	LPHN/NU<HPHN/U
SD	1.04	0.91	1.04	1.15	
Collaboration	0.00	-0.45	0.17	0.04	LPHN/NU<HPHN/U
SD	1.04	0.99	0.91	1.00	

APPENDIX E: SAMPLING METHODOLOGY

To establish the appropriate sample for use in the analyses, we began with the complete SASS database. Our goal was to identify regular classroom teachers who taught in traditional (not solely serving special populations) high schools that were both high-needs and either 'high-performing' or 'low-performing.' High schools were categorized as high-need by selecting schools that had 50 percent or more students who qualified for free or reduced price lunch. Schools were categorized as either high-performing or low-performing high schools based on principals' answers to Item 28 of the SASS principal survey. We chose to eliminate the schools where the principal rated the school as passing 'most' or 'some' of the state and district performance standards. This was done to produce a sample with two distinctly separate high- and low-performing groups. Ideally, student achievement data on a test such as NAEP would have provided us with a comparable variable with which to assess school performance. However since not all schools participate in NAEP, we chose to select only the highest and lowest schools based on the principal rating.

Item 28. Which of the following best describes this school's performance last year?

a. Passed all district and state performance standards
b. Passed most district and state performance standards
c. Passed some district and state performance standards
d. Passed no district and state performance standards

High schools whose principals selected response *a.* were categorized as high performing; those whose principals selected response *d.* were categorized as low performing. Data from schools in which principals selected the middle two categories (*b. and c.*) were not used.

We recognized that identifying school performance using this method might be open to principal bias[13]. In order to evaluate the performance of each school identified for the sample, the Common Core of Data (CCD) of the National Center of Education Statistics (NCES) was used. Once schools from the sample were identified, state department of education websites were searched for achievement data for the 2002-2003 year to confirm the performance level[14]. Performance data, which often included norm-referenced and criterion-referenced measures, were available for 24 of the 37 states that had schools in the sample. This resulted in a sample for this analysis of 136 high

[13] However, Englert, Fries, Martin-Glenn, and Douglas (2007) cross-checked principal responses to a similar item with school performance data in state databases and found that principals were generally accurate in reporting the performance levels of their schools.

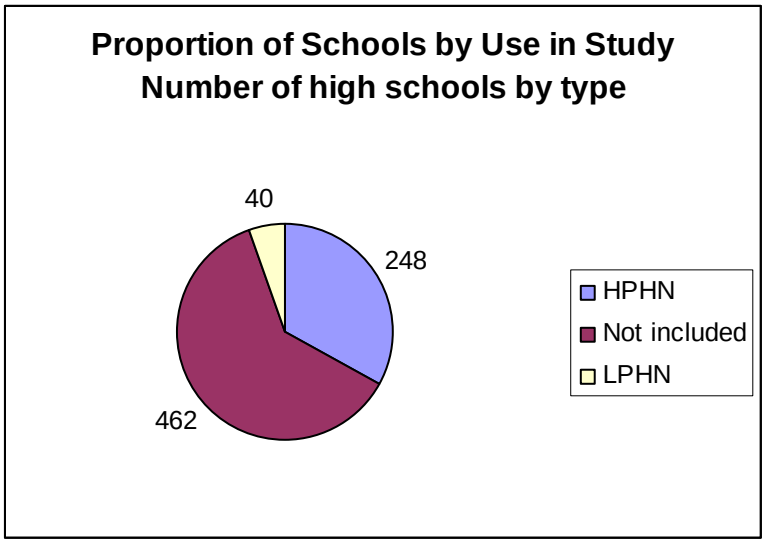
[14] During the process of evaluating and checking schools performance level, we found that 34 schools were classified as elementary or middle schools as compared to the Common Core of Data database. In addition to these 34 schools, another 9 schools were identified as being K-12; two more were identified as 6-12. These schools were all contacted to determine if they contained at least one of the high school grades (9-12). All but two of the schools did indeed have one of the higher grade levels; most (23) were K-12.

schools. We compared principals' ratings for what we classified as HPHN and LPHN to their schools performance based on state standards. If a school fell below the state's standards, the school was classified as LPHN and if they were above they were classified as HPHN. We found a 94 percent agreement rate between principal ratings and performance data.

Next, "regular" high schools were selected. These are schools that were coded as enrolling grades 9, 10, 11, and/or 12. Because K-12 schools are not "traditional" high schools, we considered excluding them from the sample and retaining only traditional (9-12) schools. We decided, however, to keep the K-12 school sample because they were predominantly small rural schools and removing them would have greatly reduced the representation of rural schools in our sample. However, teacher data from these schools were included only if they reported teaching at the 9, 10, 11, or 12th grades.

Finally, schools with student admission criteria were removed as not typical of high-performing schools. These schools were identified using an item in the SASS school survey that specifically asks, 'Does this school have any special requirements for admission other than proof of immunization, age or residence?'. Additionally, schools were eliminated if they were determined to serve special populations, e.g., have only students that have been expelled or suspended from other schools, or contain only special education or vocational education students.

Using the identification from Item 28 for HPHN and LPHN schools, there were 846 teachers in 248 high-performing high schools and 127 teachers in 40 low-performing high schools. There were 462 high-needs high schools whose principal's responses to the item were b or c. Thus, the final sample includes only the highest and lowest schools in terms of reported performance. The graphic below depicts the proportions.



REFERENCES

Abdi, H. (2007). Bonferroni and Sidak corrections for multiple comparisons. In N.J. Salkind (Ed.): *Encyclopedia of Measurement and Statistics.* Thousand Oaks (CA): Sage. pp. 103-107.

Achieve, Inc. & National Governors Association. (2005). *An action agenda for improving America's high schools.* New York: Author.

Alliance for Excellent Education. (2003). *Progress report on American high schools.* Washington, DC: Author.

Barton, P. E. (2005). *One-third of a nation: Rising dropout rates and declining opportunities.* Princeton, NJ: Educational Testing Service.

Borman, G.D., Hewes, G.M., Overman, L.T. & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research* , 73 (2) pp. 125-230.

Brookover, W., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *School social systems and student achievement: Schools can make a difference.* New York: Praeger Publishers.

Cooper, J. E., Ponder, G., Merritt, S., and Matthews, C. (2005). High-performing high schools: Patterns of success. *NASSP Bulletin, 89*(2), 2-23.

Dolejs, C. (2006). *Report on key practices and policies of consistently higher performing high schools.* National High School Center, Washington, DC

Department of Education (Downloaded 2008). Guidance on Schoolwide Programs. http://www.ed.gov/legislation/ESEA/Title_I/swpguide.html

Dougherty, R. Mellor, L. & Smith, N. (2006, May). *Identifying Appropriate College-Readiness Standards For All Students*. National Center for Educational Accountability Issue Brief # 2. Retrieved December 1, 2006 from http://www.nc4ea.org/files/College_Readiness_article-05-03-06.pdf

Englert, K., Fries, D., Martin-Glenn, Douglas, B. (2007). Accountability Systems: A Comparative Analysis of Superintendent, Principal, and Teacher Perceptions. International Journal of Education Policy & Leadership.

Eisen, P., Jasinowski, J. J., and Kleinert, R. (2005). *2005 skills gap report: A survey of the American manufacturing workforce.* Retrieved June 5, 2006 from http://www.deloitte.com/dtt/cda/doc/content/us_mfg_Talent%20Management_112205.pdf.

Evers, J.W. and Bacon, T.H. (1994). *Staff perceptions of effective school components as a means to school improvement and accountability.* Retrieved February 19, 2008 from http://www.eric.ed.gov/ERICWebPortal/ ED379389

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299.

Fan, X.T., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational psychology review, 13*(1), 1-22.

Federal Register: July 2, 2004 (Volume 69, Number 127)

Heck, R. H., Larsen, T. J., & Marcoulides, G. A. (1990, May). Instructional leadership and school achievement: Validation of a causal model. *Educational Administration Quarterly, 26*(2), 94–125.

Jeynes, W.H. (2007). The relationship between parental involvement and urban secondary school student academic achievement: A meta-analysis. *Urban Education, 42*, 82-110.

McREL (2005). Final Report: High-Needs Schools- What does it take to beat the odds? [REL Deliverable 2005-08]. Aurora, CO: Mid-continent Research for Education and Learning.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, §115 Stat. 1425 (2002).

NRC Institute of Medicine. (2003) *Engaging schools: Fostering high school students' motivation to learn.* Washington, DC: The National Academies Press.

Parsad, B., & Lewis, L. (2003, November). Remedial education at degree-granting postsecondary institutions in fall 2000: Statistical analysis report (NCES 2004-010). Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved September 19, 2004, from http://nces.ed.gov/pubs2004/2004010.pdf

Railsback, J., Reed, B., & Boss, S. (2001). *Profiles of progress: What works in Northwest Title I schools.* Northwest Regional Education Laboratory. Retrieved from http://www.nwrel.org/request/sept01/references.html

Raudenbush, S. W., & Bryk, A. S. (1989). Methodological advances in analyzing the effects of schools and classrooms on student learning. In E. Z. Rothkopf (Ed.), *Review of research in education, 15* (pp. 423–475). Washington, DC: American Educational Research Association.

Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness.* Oxford, U. K.: Pergamon.

Schools and Staffing Survey (SASS) – 2003-04 (NCES 2006-313 Revised) U.S. department of Education, National Center for Education Statistics, Washington, DC

Stevens, J. (1996). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Lawrence Erlbaum Associates.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education.

Teddlie, C., Stringfield, S., Wimpelberg, R., & Kirby, P. (1989). Contextual differences in models for effective schooling in the USA. In B. P. M. Creemers, T. Peters, & D. Reynolds (Eds.), *School effectiveness and school improvement: Proceedings of the Second International Conference* (pp. 117–130). Amsterdam/Lisse; Rockland, MA: Swets & Zweitlinger.

The Education Trust (2005). *Gaining traction, gaining ground: How some high schools accelerate learning for struggling students*. Washington, DC: Author

U.S. Department of Education (2002). *Comprehensive School Reform (CSR) Program Guidance.* Retrieved February 21, 2008, from http://www.ed.gov/offices/OESE/compreform

U.S. Department of Education (2005). *No Child Left Behind: Transforming America's high schools.* Retrieved from http://www.ed.gov/about/offices/list/ovae/pi/hsinit/papers/nclb.pdf

U.S. Department of Education (2006). *Charter high schools closing the achievement gap*. Retrieved from http//www.ed.gov/admins/comm/choice/charterhs/index.html

Von Secker, C.E. & Lissitz, R.W. (1997). Estimating school value-added effectiveness: Consequences of respecification of hierarchical linear models. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.