



A Study of MCAS Achievement and Urban Special Education
The Impact of K to 8 School Configuration on Achievement

July 2005



Acknowledgements

The University of Massachusetts Donahue Institute extends its sincere gratitude to the organizations and individuals who have contributed to the conduct of this ongoing study, including: the Massachusetts State Legislature, whose vision created this opportunity for research; our valued partners at the Massachusetts Office of Educational Quality and Accountability, particularly Executive Director, Dr. Joseph Rappa; and the administrators and teachers from the Boston, Cambridge, Lynn, and Somerville school systems who so generously gave of their time and shared their insights with us.

Research Team Members

Steven Ellis, MPA	Research Manager, UMass Donahue Institute
Robert Gaudet, Ed.D.	Lead Quantitative Researcher, UMass Donahue Institute
Matt Hoover, MA	Research Analyst, UMass Donahue Institute

University of Massachusetts Donahue Institute
Research and Evaluation Group
100 Venture Way, Suite 5, Hadley, MA 01035-9462
P: (413) 587-2400 F: (413) 587-2410
www.donahue.umassp.edu



Table of Contents

Acknowledgements

Table of Contents

Executive Summary

I

I. Methods and Findings: MCAS Data Analysis

1

Introduction

1

Analysis of Student Achievement

1

Main HLM Effects

2

Cross-level Interaction Effects

3

Identifying K-8 Schools of Interest

3

II. Methods and Findings: District and School Interviews

5

Introduction

5

Summary of Interview Findings

5

Concerns

7

School Attributes

7

Conclusion

8

Appendix: Full Results of Hierarchical Linear Modeling Analysis

9

1. Methodology: HLM in Brief

9

2. Source Data

10

3. Results

10

4. Summary Points

15



Executive Summary

In academic year 2003-2004, the University of Massachusetts Donahue Institute, in collaboration with the Office of Educational Quality and Accountability, conducted an extensive analysis of the Massachusetts Comprehensive Assessment System (MCAS) achievement of urban elementary and middle school students with special needs. That research identified high performing urban districts and schools throughout the Commonwealth, among which was an apparent over-representation of schools utilizing a Kindergarten through grade 8 (K-8) configuration.¹ This observation, coupled with concerns regarding the challenges faced by students with special needs as they transition to middle school, led us to ask:

Does the K-8 school design have a positive effect on the achievement of students with special needs?

To answer this question, the research team analyzed 2004 English language arts (ELA) and Math MCAS achievement data. The objective was to isolate grade configuration as a school-level variable and understand its influence on MCAS achievement, while accounting for individual student level factors such as poverty, English proficiency, and special needs status. This required hierarchical linear modeling (HLM), an analytical tool that is increasingly used in educational research because it accounts for the nested structure of students, who function within larger units - schools - that have unique attributes of their own, which may interact with student characteristics.

Four of the five student level characteristics tested through the HLM analysis—free or reduced lunch eligibility, limited English proficiency, non-white, and special needs—displayed a consistently negative relationship with MCAS achievement (gender displayed a mixed impact). Analysis also revealed that some of these student groups—most notably students with special needs or with limited English proficiency—tend to perform better in smaller school settings, as a negative achievement impact is associated with increasing school size. With this established, the question of the impact of K-8 school configuration on student achievement could be answered in context.

With regard to the achievement of students who do not possess any of our target characteristics, the impact of K-8 configuration was negative on all exams, although only the grade 4 ELA and Math findings were statistically significant. K-8's effect on students eligible for free or reduced lunch or with limited English proficiency were also generally negative at grade 4, but mixed on the grade 7 ELA and grade 8 Math exams; while the findings for non-white students suggest K-8 has a positive impact on all tests, but are not statistically significant.

In the context of these mixed and unclear findings, tests of the effect of K-8 on the achievement of students with special needs were compelling. The findings were uniformly positive, with K-8 school configuration displaying a statistically significant, positive association with SPED student achievement on three of four tests (grade 4 ELA and Math, grade 7 ELA). Findings suggest a positive effect on the grade 8 Math test as well, but this finding was not statistically significant.

Having identified that a K-8 grade configuration is associated with improved MCAS achievement by students with special needs, the next step in research was to understand why. To answer this question, we conducted district and school level performance analyses to determine which urban schools, in which districts, demonstrated the most consistent and positive MCAS achievement among students overall, and among those with special needs, in particular. We identified three districts - Cambridge, Lynn, and Somerville - that had produced exemplary

¹ See *A Study of MCAS Achievement and Promising Practices in Urban Special Education: Data Analysis and Site Selection*, available at http://www.donahue.umassp.edu/news/one-item?item_id=12748.



achievement overall using the K-8 configuration. We also identified several schools in Boston that exhibited strong performance.

Having identified schools that had successfully implemented K-8, we then arranged interviews to discuss the relative strengths and weaknesses of K-8 configuration with top administrators and teachers at each site. In sum, 27 administrators and teachers were interviewed. Analysis of those interviews revealed that these educators did not attribute success to the K-8 configuration per se, but did believe that it made it easier to teach all students effectively. Interviewees generally felt that K-8 offered the following advantages over more traditional, segregated elementary and middle school configurations:

- K-8 engenders a shared responsibility for learning across all grade levels because teachers are connected to students for a longer period of time. Teachers grades K-8 work in the context of a single system, increasing accountability for performance.
- Communication among staff is often better across grade levels in a K-8 configuration. With children in the same school from K-8, teachers of middle school aged students can easily communicate with teachers who had the child in earlier grades.
- There is real collaboration and a continuity of instruction/curriculum from grade K-8, which is difficult to support through more traditional grade configurations, where elementary and middle school grades are more likely to reside in separate buildings.
- K-8 eliminates certain aspects of a student's transition from elementary into middle school, which can be of particular importance to students who are sensitive to such transitions.
- K-8 schools are more supportive of building community, creating stability, and making the school feel like a family. This is possible, in part, because students, school staff and parents maintain connection over a longer period of time, allowing relationships to develop.

Notably, many of these attributes correspond closely to practices reported in our *Summary of Field Research Findings: A Cross Case Analysis of Promising Practices in Selected Massachusetts Urban Public Schools*.² In addition, that report noted the profound concern of many educators regarding the impact of middle school transitions on students with identified or emerging special needs.

The interview process also identified some potential shortcomings of K-8 configurations. These include:

- The challenge of meeting the needs of students at widely different developmental and educational levels.
- The limitations that accompany the relatively small size of middle school grades, which limit peer group size, options for courses and teachers as compared to larger, traditional middle schools.
- The possibility that students are so nurtured in a K-8 environment that they may have difficulty making the transition to high school.
- Although some interviewees suggested that K-8 facilitates or can extend parental involvement, which often wanes during the middle school years, other respondents felt that K-8 offered no real advantage to developing parent relationships.

Interview findings suggest that K-8 is no magic bullet for student success, overall, and that effective schools will show good results, regardless of grade configuration. However, the advantages of K-8, as reported by these educators, may make these schools more responsive to the needs of students with special needs, who—evidence suggests—are particularly sensitive to the transition to traditional middle or junior high school environments.

² Published by the University of Massachusetts Donahue Institute, October 2004. Available at http://www.donahue.umassp.edu/docs/?item_id=12695.

I. Methods and Findings: MCAS Data Analysis

Introduction

Over the past two years, the University of Massachusetts Donahue Institute, in collaboration with the Massachusetts Office of Educational Quality and Accountability, has studied the MCAS achievement of special education students in districts with urban characteristics.³ This research included a comparative analysis of the MCAS achievement of students with special needs at both the district and individual school levels. School level analyses suggested a possible over-representation of schools using a Kindergarten through 8 (K-8) grade configuration among the top performers. In addition, school level interviews suggested concerns regarding the effects of the transition to middle school upon students with special needs and other student sub-groups. These observations led us to ask, “Does the K-8 school design have a positive effect on the achievement of students with special needs (and, if so, why)?”

To answer these questions, we first needed to identify all public schools utilizing a K-8 configuration in the Commonwealth of Massachusetts, a task enabled by data available at the Massachusetts Department of Education’s web site (www.doe.mass.edu). In total, we identified 114 schools that utilize a K-8 configuration. These data were merged into existing MCAS data sets containing 2003 and 2004 student level results on four MCAS exams, including the grade 4 English language arts (ELA), grade 4 Math, grade 7 ELA, and grade 8 Math. Analyses were performed to investigate whether school configuration has a measurable impact on student achievement.

Analysis of Student Achievement

A preliminary analysis used multiple linear regressions to determine whether K-8 schools were indeed more likely to produce higher MCAS scores than traditional elementary and middle school. Regression looks at the overall impact of individual factors on an outcome. For example, a regression can test the effect of poverty (Free and Reduced Lunch rate) on MCAS scores. Results suggested that the analysis was worth pursuing on a more sophisticated level, which would allow us to determine how individual student-level factors actually affect MCAS outcomes in the context of actual schools.

The next step in this analysis was to use hierarchical linear modeling (HLM), one of several analytic strategies used to simultaneously model nested data. By *nested*, we mean that information is collected on units that are a subset of a larger unit. For example, employees are nested within businesses; students are nested within classrooms; et cetera. HLM is increasingly used in educational research because of the nested structure of students, who are nested within larger units - schools - which have unique attributes of their own that interact with student characteristics.

The strength of HLM is that it allows the researcher to look at outcomes as a function of both individual factors and the context in which those factors operate. In this study, we looked at the impact of individual factors (Free/Reduced Lunch; Special Education Status; Limited English Proficiency [LEP]; Gender; and Race) in a school context that was defined by Total School Size and K-8 School Arrangement. With HLM, you can determine how individual student factors operate in each school in affecting outcomes, which were MCAS scores in this case.

³ A summary of publications can be found at <http://www.donahue.umassp.edu/publications/publications.htm>



We used HLM to determine how achievement is influenced by Level 1 factors (student characteristics such as Free/Reduced Lunch; Special Education Status; LEP status; gender; and race) and how it is influenced by Level 2 factors (Total School Size; K-8 School Arrangement). Finally, we looked at Cross-Level Interactions, which reflect how various factors actually influence MCAS achievement in the real school setting, with specific interest in assessing whether the K-8 design was a positive factor in student achievement. Our analysis looked at the impact of K-8 on student sub-groups – special education students, LEP students; Students eligible for Free/Reduced Lunch; Non-White students; and Male students.

Main HLM Effects

The following table illustrates Main HLM Effects for grade 4 MCAS results under the HLM work. Level 2 factors are Total School Size and K-8 School Arrangement. (This material, explanatory information, and the balance of the HM analysis for other grades and tests can be found in the appendix.)

Table 1: Grade 4 English HLM Models: Outcome is English raw score

	Model 1 (School Size)	Model 2 (K-8 Structure)	Model 3 (Size and K-8)	Model 4 (Free/Reduced Lunch)	Model 5 (All Predictors)
Level 2 Predictors					
Total School Size	-.002 **		-.0009	-.0009	-.0009
K-8 School Arrangement		-2.68***	-2.66 ***	-2.66 ***	-2.67 ***
Level 1 Predictors					
Free or Reduced Lunch				-4.47 ***	-3.23 ***
Special Education Status					-7.24 ***
Limited English Proficient					-4.74 ***
Gender: Male					-1.95 ***
Race: Non-White					-1.86 ***

Table 1 presents the results of the grade 4 English test. Model 1 is an HLM run with no level 1 predictors (student characteristics) and just the total school size as a level 2 predictor (Total School Size and K-8 School Arrangement) for school performance. The -.002 regression coefficient in the “Model 1” column indicates that for every additional increase of 1 student to the total size would predict a decrease in performance of .002 points.⁴ Model 2 is just the impact of K-8 school arrangement only. This indicates that compared to non-K8 schools, students in these schools are expected to perform 2.68 points worse.

The Model 3 shows the effect of combining both level 2 predictors (total school size and the K-8 arrangement). Interestingly, school size is no longer a significant predictor of performance when the K-8 arrangement is known. This indicates that the effect of school size is partially explained by the presence of a K-8 arrangement. Since the effect of the K-8 arrangement only changed slightly, this indicates that there is a unique impact of the K-8 school arrangement on performance. Model 4 shows the impact of the level 1 predictor of socio-economic status (whether the student is on free or reduced price lunch). Students on free or reduced lunch are expected to perform nearly 4 and a half points worse than students not on free or reduced lunch.

The last model (Model 5) is the full model with all level 2 and level 1 predictors. This shows that the single biggest predictor of performance is special education status. Special education students are expected to perform over 7 points worse than non-special education students. The final model also shows that LEP status decreasing performance as well as gender and race (Males and non-Whites are expected to score close to 2 points worse compared to females and whites). Please refer to the appendix for detail on HLM and the other test analyses.

⁴ This effect was significant at the .01 alpha level (“*” = .05, “**” = .01, and “***” = .001).

Cross-level Interaction Effects

The final HLM analysis looked at Cross-level Interaction Effects, that is the impact of individual student factors as they actually operate in schools. We looked at K-8 schools and non-K-8 schools. Our outcome indicator was grade 4, 7, and 8 MCAS scores.

- On 3 of the 4 tests, special education status had a significant interaction with K-8 school arrangements. Special education students generally perform better in schools with a K-8 arrangement type.
- On 3 of the 4 tests, special education status had a significant interaction with total school size. The impact of adding 1 additional student disproportionately impacts special education students more than non-special education students.
- On 3 of the 4 tests, limited English proficiency status had a significant interaction with total school size. The impact of adding 1 additional student disproportionately impacts Limited English Proficiency (LEP) students more than non-LEP students. In summary, special education students appear to do better in a K-8 school on most tests. (See Appendix for more information on the HLM model.)

Four of the five student level characteristics tested through the HLM analysis—free or reduced lunch eligibility, limited English proficiency, non-white, and special needs—displayed a consistently negative relationship with MCAS achievement (gender displayed a mixed impact). Analysis also revealed that some of these student groups—most notably students with special needs or with limited English proficiency—tend to perform better in smaller school settings, as a negative achievement impact is associated with increasing school size. With this established, the question of the impact of K-8 school configuration on student achievement could be answered in context.

With regard to the achievement of students who do not possess any of our target characteristics, the impact of K-8 configuration was negative on all exams, although only the grade 4 ELA and Math findings were statistically significant. K-8's effect on students eligible for free or reduced lunch or with limited English proficiency were also generally negative at grade 4, but mixed on the grade 7 ELA and grade 8 Math exams; while the findings for non-white students suggest K-8 has a positive impact on all tests, but are not statistically significant.

Identifying K-8 Schools of Interest

Despite the HLM findings, not all K-8 schools outperform expectations. There is great variability among the schools in the study – leadership, resources, curriculum, culture – all of which can have a profound influence on student outcomes. In order to identify exemplary schools for field study we did the following.

We first compared predicted MCAS Proficiency Index (PI) and pass rate scores to actual grade 4, 7, and 8 MCAS scores, with the predicted score based on percentage of Free/Reduced Lunch, Limited English Proficiency status, and school attendance rate, variables that have powerful influences on achievement. We then sorted K-8 schools by their over-performance – the difference between predicted and actual PI and predicted and actual pass rate performance.

We then factored in how much better a school had performed than predicted. If a school's MCAS score was in the top standard deviation of over-performance (generally in the top quarter of performance), it received two points. If it was not, but still over-performed, it received one point. We added the points and came up with total scores for the K-8 schools we studied. The top scorers were then examined to see if they had produced consistently high achievement across most grades and tests or if they produced strong scores for specific grades and tests. We were especially interested in schools that did well with special education students in grade 7 and 8, since students in these middle grades tend to have poorer MCAS performance. Those schools that met our criteria were selected for field research.

After this analysis, we noted that three districts had top-performing K-8's (Cambridge; Somerville; Lynn) and that



several Boston schools were strong. Please note that not all schools produced strong results across the board. Some schools did well in elementary grades and not as well in middle grades; for others the converse was true. Schools selected for possible study included:

Boston: Raphael Hernandez

Cambridge: Maria Baldwin; Peabody; Graham & Parks

Lynn: Cobbet; Robert Ford

Somerville: East Somerville Community; Winter Hill Community

There were other schools that demonstrated solid performance compared to predicted scores, but resource and time constraints limited our choices. There were also some schools that, while demonstrating solid achievement, had too few students taking MCAS to produce reliable statistical results.



II. Methods and Findings: District and School Interviews

Introduction

While there is little established national research on the effectiveness of the K-8 school design in boosting achievement⁵, our HLM analysis suggested that there were positive effects for special education students. There are over 100 K-8 schools in Massachusetts today and with more being planned and built. Thus, both the results of the HLM analysis and the fact that many educators and policy makers are embracing the K-8 design led us to develop a research protocol that would help us understand why the K-8 school design in our research was associated with better special education MCAS scores.

In order to identify the specific elements of the K-8 educational experience that enhanced achievement, we selected four districts for further investigation based on district and school level performance analyses. We conducted district and school level performance analyses to determine which urban schools, in which districts, demonstrated the most consistent and positive MCAS achievement among students overall, and among those with special needs, in particular. The districts chosen —Cambridge, Lynn, Somerville and Boston - had produced exemplary achievement by their K-8 schools.

We were able to speak with administrators and teachers at the following districts and schools: Boston - Raphael Hernandez; Somerville - East Somerville Community; Winter Hill Community; and Lynn - Cobbet; Robert Ford. In Cambridge, due to time constraints, we were not able to meet with school personnel, but we were able to speak with district administrators.

We interviewed 27 educators, including superintendents and senior staff; principals; specialists; and teachers. Our interview protocol; was designed to determine if there were salient characteristics about the grade configuration of K-8 schools that contributed to success. We also were interested in understanding what practices and policies were in place at these relatively high-achieving schools.

Summary of Interview Findings

Was there a specific reason why districts in the study had adopted the K-8 design? We did not have a big enough sample size to make any authoritative analyses, but speaking with educators about the reasons behind embracing the design were instructive. In most cases K-8's complemented more traditional K-5 or K-6 designs. Cambridge, which has had all pre-secondary schools in the K-8 design for decades, was the exception here. Districts offered a range of motivations for their use of the K-8 model. Often in the districts, parents' concerns were the initial catalyst for making a school K-8. Often a "bad middle school experience" or fear of such was the reason cited by parents for wanting to add middle school grades to elementary schools. Again there is little compelling national research available on the issue of the reasons behind the popularity of the K-8 design, and in our interviews, there was no consistent answer to the question about why the district had adopted the configuration. In short, there seem to be no research-based reasons for moving towards the K-8 configuration even in districts that embrace the design.

There was no consistent pattern to the physical layout of the K-8 schools in the study. Some had different grade clusters in different wings of the building; others had grade clusters on different floors. One school, the Robert Ford in Lynn, has two buildings two miles apart comprising the K-8 school. Regardless of physical layout, all schools operated under a single administrative team.

⁵ Pardini, Priscilla, *School Administrator*, v 59 n 3, March 2002, p. 6-12
Available online at http://www.aaa.org/publications/sa/2002_03/pardini.htm



In the interviews, several points consistently emerged about the K-8 design.

- All teachers have responsibility for MCAS success throughout the student's academic career. The grade 3 or 4 teacher knows that what he or she teaches will affect later success, and that the teacher will see the child in later years in the same school. There is a real shared responsibility for learning.
- Communication among staff is better in a K-8 configuration. Since the children can be in the school for K- through grade 8, upper-school teachers can easily communicate with teachers who had the child as a youngster. This can enhance the information flow about the individual needs of students. For students who are in special education, this can be especially helpful.
- There is real K-8 collaboration. Besides the informal contacts between teachers in different grades, many schools organize professional development activities to connect across elementary and middle school. This helps build a continuity of instruction that may not be present in more traditional grade configurations.
- K-8 allows smoother transition from elementary school into middle school. While students still have issues with the physical and psychic changes that come with adolescence, they do not have to uproot themselves from one school to go to another. Some interviewees noted that young people have enough stress in their lives without bringing a new school into the mix.
- Keeping an orderly school is easier. Older students tend to behave better when there are younger students in the school, especially when these younger students may be their siblings. K-8 does not eliminate discipline problems, particularly among younger students, but it is seen to have a calming effect on the school. Most schools have "Buddy" programs that match older students with younger ones for reading or tutoring. This is seen as a way to give middle school students responsibility for helping others as well as a way to help younger students learn.
- Fewer middle school students in a K-8 school means that you do not have a critical mass of adolescents in one place at one time. Instead of having 600 pre-adolescent students in one setting, as is the case in a middle school, in a K-8 you often have only 150 to 200 or so such students. This size group is seen as easier to handle than the larger populations of middle schools.
- K-8 schools are more supportive of building community, creating stability, and making the school feel like a family. Students literally grow up together, and teachers may see some students for 8 or 9 years. Teachers know the students.

There was no consensus about the advantage of K-8 in some areas.

- People interviewed had mixed views of how much the K-8 configuration increases parental involvement. Parental involvement is very strong at the Hernandez school in Boston, but that may have more to do with the quality of the school and its reputation as an outstanding two-way bilingual facility than the fact that it is K-8. Staff in other schools suggested that even with K-8 it was still difficult to get strong parental involvement, although many thought that having younger children in the same school as older ones keeps parents connected to the school longer. Several administrators mentioned that the quality of parental involvement ranged widely – from just showing up at a food event to actively participating in helping the school improve. This intensity of involvement seemed to be idiosyncratic in that it depended on school factors other than the K-8 configuration.

The Robert Ford School in Lynn has an interesting approach to involving parents. The school offers a variety of programs for adults - ESL; GED; and citizenship programs. Since many parents come to the



school at night for their own programs, they develop a strong connection to the school that brings them back for their children.

- Most people saw no specific advantages of K-8 concerning student sub-groups. There was a general sense that K-8 is good for everyone, especially if the school is organized for success. Students who are more fragile or at-risk likely benefit from the family and community aspects of a good K-8 school, but this effect is not limited to sub-groups.

Concerns

While there were no strong disadvantages seen in the K-8 design, there were some concerns raised by educators.

- It is critical to hire appropriate staff, including a principal with middle school experience. Schools led by elementary school leaders with no middle school background would likely not be successful. It is also important to have middle school certified teachers in the upper grades.
- Because there are fewer middle school students in a K-8, there are fewer middle school teachers and fewer courses available. This may be problematic in terms of the variety of course selection and the ability of the school to move a student who is having difficulty with one teacher to another teacher.
- Having fewer middle school students in the school means that there are more limited opportunities to make friends and bond with peers. This is not a problem for everyone, but in a small K-8, the number of upper school students is limited.
- Students nurtured in a K-8 environment may have difficulty making the transition to high school, particularly if they will be moving on to a large urban high school.

School Attributes

Not surprisingly, many of the exemplary K-8 schools identified in this report shared many of the attributes of effective schools identified in the larger MCAS achievement study. Good schools, regardless of design, share similar elements that contribute to success.

- There is a strong commitment to aligning the curriculum with the MCAS frameworks. Systems have developed resources to help individual schools do this, and principals and teachers are very aware of the need to relate teaching to the Massachusetts Frameworks.
- Professional development is seen as critical to supporting effective teaching. Schools (and their districts) see professional development as embracing a wide range of skills – curriculum development and content; differentiated instruction; English Language Learners – as well as the basics of ELA and math.
- Educators use data routinely to inform changes in curriculum and in professional development. Data is embraced as a valuable tool to improve teaching and learning,
- Strong, dynamic leadership at the district and school levels is seen as critical to producing better achievement.
- There is a merging of special education and regular education in terms of instruction and professional development. The special education silos that once were part of school structure are disappearing.



- Most of the schools had strong support services, both in terms of counselors and in terms of specialists to support subject areas and sub-groups.
- There is a strong MCAS focus, with many opportunities to help students – before and after school; Saturday classes; Summer courses. Schools also placed students in small tutoring groups to help them prepare for the tests.
- Finally, there was recognition in these schools that the education world has changed over the past ten years. One administrator categorized this as “a profound new business” of standards-based education that is causing shifts in school culture.

Conclusion

There was a general consensus among interviewees that the K-8 model was preferable to more traditional designs, but interviewees were quick to point out that there is nothing magical about grade configuration. If the school is not organized and managed for success, grade arrangement is relatively unimportant. National research on the effectiveness of the K-8 arrangement has been inconclusive. There simply is little hard data to confirm any positive effect related to K-8 design. Some researchers have noted that educational practices and policies that support achievement – team teaching; looping; cross-grade grouping – are more likely to be found in K-8 than in middle schools.⁶

The educators we interviewed were generally satisfied with the K-8 structure and feel that it has a positive effect on achievement, but the impact may have more to do with the civility, stability and community aspects of K-8 institutions than with the technical grade arrangement. The achievement bounce that we see in these schools also may be related to good practices and policies in the districts and schools. Somerville has managed to keep its pre-school and all-day kindergarten programs despite budget cuts. Such offerings help prepare children to learn and be successful in school. The Robert Ford School in Lynn has a very powerful parental engagement program that offers adults needed educational services, thus cementing the bond between parents and their children’s school. The schools in the study that were more effective generally had implemented policies and practices that would likely make any school more effective, but it may have been easier for these pedagogical innovations to take root in the K-8 environment.

⁶ See “Effects of School Restructuring on the Achievement and Engagement of Middle Grade Students,” by V. E. Lee and J. B. Smith, highlighted in an article in *American Association of School Administrators*, March 2002, http://www.aasa.org/publications/sa/2002_03/pardini_resources.htm.

Appendix: Full Results of Hierarchical Linear Modeling Analysis

1. Methodology: HLM in Brief

Hierarchical linear modeling is one of several analytic strategies used to simultaneously model nested data. By *nested*, we mean that information is collected on units that are a subset of a larger unit. For example, employees are nested within businesses; students are nested within classrooms; et cetera. Hierarchical linear modeling (HLM) is increasingly used in educational research because of the nested structure of students, who are nested within larger units (schools, which have unique attributes of their own). A brief discussion of level 1 and level 2 Units, as well as Fixed and Random Effects, will provide a better basis for understanding how HLM works.

Level 1 and Level 2 Units

In HLM designs, there can be 2 level or even 3 level analytical models. In the two level model (the most common design); information on the “lowest” unit is called “level 1” data. In the educational context, this is often information related to student characteristics (performance, socio-economic status, primary language, etc.). These level 1 characteristics are nested within a larger context (level 2 units), such as the characteristics of a school or community.

The regression equation used in HLM designs specifies that the outcome is a function of both level 1 predictors as well as the unique influence of being nested within a level 2 unit. In our regression model, we analyzed school performance as a function of both individual (level 1) predictors as well as of the particular features (level 2) of the schools these students were nested within (such as K-8 grade configuration and total school size).

Fixed vs. Random effects

In regression analysis a distinction is made between fixed and random effects. The term “fixed effect” refers to assumptions that are made about the independent variables and the error distribution of those variables. In most regression equations, we assume that the values of independent variables are “fixed” to that study. For example, in an experimental drug design where groups are given 0 mg of a drug, 5mg of a drug, and 10mg of a drug, we make no inferences as to how other values of the independent variable influence an outcome. A “random effects” model assumes that the values of an independent variable are taken as a random sample from all possible values of that variable, and therefore inferences can be made outside of the confines of that study.

In HLM regression, the level 1 and the level 2 units are assumed to be fixed effects. However, the level 1 intercepts and slopes are assumed to vary randomly across groups. For example, the influence of socioeconomic status (SES) on performance is assumed to be fixed across all level 1 units. However, when we model how SES influences performance in a particular school, we assume that the slope is dependent on the unique impact that SES has within that particular school (the SES influence on performance is now both a function of how all students perform based on SES as well as how each school impacts the SES – performance relationship).

In HLM, we can specify which level 1 slopes we are assuming to be fixed and which we are assuming to be random. A regression model that uses both fixed and random effects is called a “mixed” model. It is possible to constrain all level 2 units to where only fixed effects will be generated. In our regression models, we assumed a random design in which each particular school is assumed to be randomly selected from a pool of possible schools and the variability within each school is included in the model.

2. Source Data

Student-level 2003-2004 MCAS achievement data were provided by the Office of Educational Quality and Accountability, which received the data from the Massachusetts Department of Education. We analyzed MCAS achievement data for the English language arts (ELA) and Math exams in grades 4, 7, and 8 (Grade 4, ELA and Math, grade 7 ELA, and grade 8 Math). The outcome measure was the raw score of each student on each of the available tests.

Available level 1 (student profile) data included information on student background and socio-economic factors, including free or reduced lunch status, special education status, limited English proficient status, gender, and race. The most useful level 2 (school level) data available included total school size and the K-8 school arrangements.

Separate data runs were performed for each grade and test combination. Separate HLM analyses were run showing the changes in regression coefficients when additional level 1 or 2 factors are added to the model. Main HLM effects, showing the direct influence of the level 1 and level 2 predictors, as well as cross-level interactions, are presented below.

3. Results

The following are the results of two distinct HLM analyses. The first of these present *main HLM effects*, these are the influence that contextual factors (level 2) and individual factors (level 1) have on the outcome of interest. The main HLM effects assume each contextual and individual factor have an independent influence on the outcome of interest. For our analysis, the contextual factors of school size and K-8 arrangements as well as the individual student-level factors (poverty, LEP, etc.), each have a separate influence on school performance. Each variable is independent of each other and they each contribute some information about the outcome.

A different way of viewing the impact that level 2 factors have on performance is to see if there are any interactions with the level 1 factors that might disproportionately impact performance. This way of viewing the influence on the outcome is called the *cross level interaction* effects. Here, each factor is not assumed to be independent. Since context can have a disproportionate influence with or without the presence of certain individual factors, all of the level 2 variables are tested across all level 1 variables. For example, the contextual variable of school size might only have a strong impact on performance for specific demographic groups. It is possible that the main effect of school size is small or not significant, but for special education students, school size has a big impact on performance. In this example, school size *interacts with* special education status in explaining school performance.

A. Main HLM Effects

Grade 4 English

Table 1: Grade 4 English HLM Models: Outcome is English raw score

	Model 1	Model 2	Model 3	Model 4	Model 5
Level 2					
Total School Size	-0.002 **		-0.0009	-0.0009	-0.0009
K-8 School Arrangement		-2.68***	-2.66 ***	-2.66 ***	-2.67 ***
Level 1					
Free or Reduced Lunch				-4.47 ***	-3.23 ***
Special Education Status					-7.24 ***
Limited English Proficient					-4.74 ***
Gender: Male					-1.95 ***
Race: Non-White					-1.86 ***

Table 1 presents the results of the grade 4 English test. Model 1 is an HLM run with no level 1 predictors and just the total school size as a level 2 predictor for school performance. The -.002 regression coefficient indicates that for every additional increase of 1 student to the total size would predict a decrease in performance of .002 points. This effect was significant at the .01 alpha level (“*” = .05, “**” = .01, and “***” = .001). Model 2 is just the impact of K-8 school arrangement only. This indicates that compared to non-K8 schools, students in these schools are expected to perform 2.68 points worse.

The third model shows the effect of combining both level 2 predictors (total school size and the K-8 arrangement). Interestingly, school size is no longer a significant predictor of performance when the K-8 arrangement is known. This indicates that the effect of school size is partially explained by the presence of a K-8 arrangement. Since the effect of the K-8 arrangement only changed slightly, this indicates that there is a unique impact of the K-8 school arrangement on performance. The fourth model shows the impact of the level 1 predictor of socio-economic status (whether the student is on free or reduced price lunch). Students on free or reduced lunch are expected to perform nearly 4 and a half points worse than students not on free or reduced lunch.

The last model (model 5) is the full model with all level 2 and level 1 predictors. This shows that the single biggest predictor of performance is special education status. Special education students are expected to perform over 7 points worse than non-special education students. The final model also shows that LEP status decreasing performance as well as gender and race (Males and non-Whites are expected to score close to 2 points worse compared to females and whites).

Grade 4 Math

Table 2: Grade 4 Math HLM Models: Outcome is Math raw score

	Model 1	Model 2	Model 3	Model 4	Model 5
Level 2					
Total School Size	-.002 *		-.004	-.0004	-.0004
K-8 School Arrangement		-3.11***	-3.07 ***	-3.07 ***	-3.07 ***
Level 1					
Free or Reduced Lunch				-4.76 ***	-3.35 ***
Special Education Status					-7.07 ***
Limited English Proficient					-3.35 ***
Gender: Male					1.08 ***
Race: Non-White					-1.94 ***

Table 2 shows the HLM run for the grade 4 math. The same basic structure (the variables added at each model) is followed throughout the main effects analysis. Model 1 shows the effect of school size is the same as that of the grade 4 English test. This effect also goes away when both level 2 predictors are added to the model (Model 3). The effect of being in a K-8 school is stronger in the grade 4 math test than it was in the grade 4 English test. Students in K-8 schools are expected to perform worse on the grade 4 math test (by over 3 points) than students in other school arrangements. The effect of school arrangement is very stable across models.

Model 4 shows the effect of free and reduced lunch, and this effect is very strong (these students are close to scoring 5 points worse on the grade 4 math test. In the final model, the impact of free and reduced lunch decreases (this is because some of the explanatory power of free and reduced lunch status is also explained by other demographic factors). The factor that has the most impact on scores for this model is special education status. Compared to non-special education students, special education students are expected to perform over 7 points worse on the grade 4 math test. In the grade 4 math test, males are expected to perform higher than females (a reversal of the grade 4 English test). Also, non-whites are predicted to score close to 2 points worse than whites on this test.

Grade 7 English**Table 3: Grade 7 English HLM Models: Outcome is English raw score**

	Model 1	Model 2	Model 3	Model 4	Model 5
Level 2					
Total School Size	0.001		0.001	0.001	0.001
K-8 School Arrangement		-.63	-.63	-.63	-.62
Level 1					
Free or Reduced Lunch				-5.12 ***	-3.30 ***
Special Education Status					-9.74 ***
Limited English Proficient					-8.21 ***
Gender: Male					-2.10 ***
Race: Non-White					-1.82 ***

The grade 7 English results are presented in table 3. In these models, both of the level 2 predictors were non significant. School size and K-8 arrangements have very little impact on performance. In the final model, the strongest determinants of performance are special education status and limited English proficiency. Free and reduced lunch status, gender, and race also have significant negative effects on performance. Most of the middle schools are larger than the elementary schools, therefore the effect of 1 additional student might not be as strong as it is at the earlier grades. The effect of being in a K-8 arrangement was also not statistically significant.

Grade 8 Math**Table 4: Grade 8 Math HLM Models: Outcome is Math raw score**

	Model 1	Model 2	Model 3	Model 4	Model 5
Level 2					
Total School Size	0.001		0.001	0.001	0.001
K-8 School Arrangement		-.74	-0.74	-.71	-.72
Level 1					
Free or Reduced Lunch				-5.24 ***	-3.41 ***
Special Education Status					-11.41 ***
Limited English Proficient					-2.95 *
Gender: Male					1.26 ***
Race: Non-White					-2.15 ***

Table 4 above presents the grade 8 math HLM runs for the level 1 and level 2 predictors. The same pattern occurs in terms of the level 2 predictors that was shown at the grade 7 English test. Total school size and school arrangement are both not statistically significant. The effect of free and reduced lunch is highly statistically significant and it remains a strong predictor of performance even after controlling for other socio-demographic factors. The biggest effect on performance for the grade 8 math test is special education status. Students who are in a special education program are predicted to score over 11 points less than non-special education students. Limited English proficiency status is also a negative impact on performance as well as being non-white. Males are predicted to perform better on this test than females by over a point.

B. HLM Cross-Level Interaction Effects

Table 5: Grade 4 English HLM Interaction Effects

	Coefficient	Significance
School Size (Impact per 1 additional student)		
by Free or Reduced Lunch	-.0004	NS
by Special Education Status	-.0033	<.001
by Limited English Proficiency	-.0046	<.01
by Gender: Male	.001	<.01
by Race: Non-White	-.0002	NS
K to 8 Configuration (vs. all other configurations)		
by Free or Reduced Lunch	-.222	NS
by Special Education Status	1.215	<.05
by Limited English Proficiency	-.561	NS
K by Gender: Male	.067	NS
by Race: Non-White	.741	NS

For grade 4 English (Table 5 above) there is a significant interaction effect between school size and special education status, school size and limited English proficiency, school size and gender (males benefit), and K-8 arrangement and special education status.

In terms of school size, adding 1 additional student has small effects - a negative impact for SPED students, and negative impact for LEP students, and a positive impact for males. Students who are in special education are benefited by being in a K-8 arrangement.

Table 6: Grade 4 Math HLM Interaction Effects

	Coefficient	Significance
School Size (Impact per 1 additional student)		
by Free or Reduced Lunch	-.0003	NS
by Special Education Status	-.0031	<.001
by Limited English Proficiency	-.0031	<.05
by Gender: Male	.001	<.01
by Race: Non-White	-.00006	NS
K to 8 Configuration (vs. all other configurations)		
by Free or Reduced Lunch	-.366	NS
by Special Education Status	1.146	<.05
by Limited English Proficiency	-.439	NS
by Gender: Male	-.150	NS
by Race: Non-White	.221	NS

The grade 4 Math test follows the same general pattern of the grade 4 English test. School size disproportionately impacts special education students in a negative direction. Students who have a limited English proficiency are also expected to do worse with additional increase of every student. In terms of the K-8 arrangement, special education students are expected to perform better in these kinds of schools.

Table 7: Grade 7 English HLM Interaction Effects

	Coefficient	Significance
School Size (Impact per 1 additional student)		
by Free or Reduced Lunch	-.0003	NS
by Special Education Status	-.002	<.01
by Limited English Proficiency	-.001	NS
by Gender: Male	-.000008	NS
by Race: Non-White	-.0004	NS
K to 8 Configuration (vs. all other configurations)		
by Free or Reduced Lunch	1.260	<.001
by Special Education Status	1.903	<.001
by Limited English Proficiency	.169	NS
by Gender: Male	-.187	NS
by Race: Non-White	.589	NS

The interaction effects pattern changes for the Grade 7 English test. In terms of school size, only the special education status shows a statistically significant, but small, interaction. There is still a significant positive interaction that the K-8 arrangement has on performance for special education students. Additionally, the K-8 arrangement also has a positive interaction with free and reduced lunch status in these kinds of environments.

Table 8: Grade 8 Math HLM Interaction Effects

	Coefficient	Significance
School Size (Impact per 1 additional student)		
By Free or Reduced Lunch	-.0003	NS
by Special Education Status	-.001	NS
by Limited English Proficiency	-.003	<.05
School Size by Gender: Male	-.0001	NS
School Size by Race: Non-White	-.0003	NS
K to 8 Configuration (vs. all other configurations)		
by Free or Reduced Lunch	.586	NS
by Special Education Status	.857	NS
by Limited English Proficiency	-1.42	NS
by Gender: Male	-.022	NS
by Race: Non-White	.283	NS

For the grade 8 Math test, the only significant interaction effect is for school size with limited English Proficient students. These students are disproportionately impacted by increasing the total school enrollment.

In 3 out of the 4 tests, there was a significant positive interaction with K-8 school arrangement and special education status. Also, in 3 out of the 4 tests, there was a significant negative interaction school size and special education students. Controlling for all other factors, the special education students perform best in smaller schools and K-8 schools.

4. Summary Points

HLM Main Effects

- The K-8 school arrangement has a negative impact on performance for elementary school students on both the grade 4 English and Math tests.
- The K-8 school arrangement is not significantly related to performance for middle school students.
- Total school size was not significantly related to performance for all of the grades and tests reported.
- After controlling for other demographic factors, socio-economic status (reflected by free and reduced lunch status) had a strong negative impact on performance for all tests and grades reported.
- Special education status had a strong negative impact on performance across all tests and grades.
- Males performed significantly higher on the grade 4 Math test and the grade 8 Math test, and females performed significantly higher on the grade 4 English test and the grade 7 English test.
- Limited English proficiency had a significant negative impact on performance for all tests and grades. This effect was strongest on the Grade 4 English and Grade 7 English tests.
- Non-white status had a significant negative impact on performance for all tests and grades.

Cross Level Interaction Effects

- On 3 of the 4 tests, special education status had a significant positive interaction with K-8 school arrangements. Special education students generally perform better in schools with a K-8 arrangement type.
- On 3 of the 4 tests, special education status had a significant interaction with total school size. The impact of adding 1 additional student disproportionately impacts special education students more than non-special education students.
- On 3 of the 4 tests, limited English proficiency status had a significant interaction with total school size. The impact of adding 1 additional student disproportionately impacts limited English proficiency students more than non-limited English proficiency students.