The Development of Mathematics Achievement in Secondary School

Individual Differences and School Effects

Lingling Ma; Kelly D. Bradley

University of Kentucky

The Development of Mathematics Achievement in Secondary School

Individual Differences and School Effects

ABSTRACT

The present study focuses on the effect of students' own characteristics on the their mathematics performance and progress. School context and climate characteristics, as well as the cluster effect of school are considered. Using the Longitudinal Study of American Youth (LSAY), where students were repeatedly measured and clustered within schools, a 3-level multilevel model is applied. Given that some demographic information, such as parent academic push, does not remain constant, the variations of these variables between the waves of the longitudinal study are taken into account. The relationship between student initial mathematics achievement and growth trajectory are also examined. The results provide a frame of reference to compare changes over time given more recent national panel studies.

The Development of Mathematics Achievement in Secondary School

Individual Differences and School Effects

Students' mathematical achievements in secondary school have an influential effect on their performance in college and their future careers. Having a solid background in mathematics helps students develop sophisticated perspectives and offers more career options. The importance of mathematical learning has repeatedly been emphasized by educators and politicians (Wilkins & Ma, 2002). Both teachers and parents have paid attention to students' performance in mathematics and their progress every year. Politicians have also called for improving students' overall performances and closing students' achievement gaps. Until teachers and parents recognize what factors influence their students' mathematics achievement and improvement, they will be unable to help them make substantial academic progress.

Educators have relied on many sources of information and focused on various factors that might affect students' mathematical achievements, including students' own backgrounds, peer environment, and parental involvement (Young, Reynolds & Walberg, 1996). In Ma and Klinger (2000), student individual characteristics, gender, age, ethnicity, and their family characteristics, marital status, socioeconomic status, were investigated. Some interaction effect was considered by Muller, Stage and Kinzie (2001) where they looked at the interaction of race-ethnicity and gender.

More than a decade ago, it was criticized by Willms and Raudenbush (1989) that lacking of adequate statistical control over school characteristics had been a chief limitation for research on school effects. Still, the influence of school on students' mathematics progress has often been overlooked. School characteristics can often be classified into two sets of variables. One set describes the context of a school: school enrollment size, school location, and percentage of free-lunch students. The other set of school level variables, often referred to as ”evaluative variables,” are associated with school climate, attempting to describe the inner working of school life, for example, school organization and expectations of students, parents, and teachers. Previous studies (Ma, 1999; Wilkins and Ma, 2002) have neglected to address this in detail.

Muller et al. (2001) points out that a more dynamic approach to experiences in academic achievement is needed. Wilkins and Ma (2002) called for further detailed longitudinal studies. Cross-sectional data considering achievement have been a main source of information. Regardless of the cost, a panel study could “show the precise patterns of persistence and change in intentions” and “eliminate the confusion by showing the change of the sample in cross section study” (Babbie, 2002. p. 98-99). A panel study should be used in order to increase the explanatory and predictable power. Wilkins and Ma (2002) studied students’ initial mathematics achievement, annual progress and their relationship in the middle schools and high schools separately and reported that students who had higher initial status tended to grow faster than those from a lower starting point. Still, the cluster effect associated with schools was neglected. It is essential to explore the relationship when other environmental factors are considered.

OBJECTIVES

This study examines students’ initial mathematics performance and their annual progress in the secondary school. By studying the relationship of initial mathematics status and the students’ growth rates, the pattern of change is reported. This study emphasized the impact of student and school on mathematics achievement. Students’ individual, peer and family characteristics are used to explain both initial math achievement and growth trajectories. School context variables and school climate variables are included in this study, and variances of student's motivation and attitude will be taken into account.

This research focuses on the students' mathematics performance in secondary school. Specifically, the following research questions will be answered:

1. What are students' initial status and the rate of growth during secondary school?
2. Do the initial status and the growth trajectory differ by student or school characteristics? Will the interactions affect students' mathematics performances? Will the variation within students as related to motivation and attitude influence their academic progress? Will the variation within schools influence students' academic performance?
3. Is there an existing pattern between the initial status and growth trajectory?

In educational effectiveness research, multilevel models have become popular since these models take account of the hierarchical structure of the data. In the social sciences, hierarchical structured data arise routinely where the students are nested within the schools (Young et al., 1996). The multilevel structure could not be ignored, as the independent assumption of many traditional statistical analyses is violated (Muller et al., 2001). Multilevel modeling was used as the method of analysis to solve the dilemma. Though it is a relatively new approach to the analysis of hierarchically structured data, it is a refined version of multiple regression. Similar to multiple regression, it can be used to look at potentially interesting differences. The multilevel modeling can also be used to explore differences in mathematical growth trajectories (Ma, 1999).

Less obvious applications of multilevel models are longitudinal research and growth research where several distinct observations are nested within individuals (Hox, 1995). In this study, the same students were measured more than once in the longitudinal studies and the students were nested in schools. Therefore, a three-level hierarchical linear growth model is applied. The first level is to model the students' mathematics scores on their grade levels. At the second level, student-level variables are added to model the initial status and the rate of growth. The third level of the model includes two between school equations that regressed the average initial status and average rate of growth in mathematics on several school-level covariates. MLwin is a windows-based statistical software package developed by the Multilevel Models Project for the analysis of multilevel models. It is used to analyze the three level model in this study.

METHODS

Data

The data for the present study came from the second cohort of the stratified national probability sample of 52 schools in the Longitudinal Study of American Youth (LSAY). Beginning in the fall of 1987, the LSAY was a longitudinal panel study of public middle and high school students. About 60 seventh graders were randomly selected in each of the 52 schools, and the total sample size was 3116 students. Students were tracked from Grade 7 to Grade 12, taking mathematics and science achievement tests and completing student questionnaires annually. With a focus on mathematics and science education, background information from parents, peers and teachers was also included in the study (Miller & Hoffer, 1994). This LSAY project guaranteed anonymity, providing aggregated data. These LSAY data are available on a CD-ROM. After selecting the variables needed, the raw data could be converted into a SPSS formatted data file by the electronic codebook system. The SPSS file was then exported into an MLwin worksheet for analyses.

Variables

Measurement of student mathematics achievement refers to student mathematical test scores from Grade 7 to Grade 12, using as dependent variables. The mathematics test contained questions drawn from the fields of basic mathematics, Algebra, Geometry, Quantitative literacy. Those scores were imputed scores, which were non-aberrant observed scores. They were stored as continuous variables and were comparable across grade levels within each school subject. Some data were missing because some students were absent during testing or they dropped out. Nevertheless, all the available data could be used for hierarchical linear modeling.

The difference in student academic achievement could be typically explained by students' individual characteristics and their family characteristics. For this reason, gender, race-ethnicity, the interaction of gender and race-ethnicity, age, number of parents, parents' socioeconomic status, number of siblings, parent push and students' attitudes are used in this study as explanatory variables at the student level. Considering students' mathematics achievements could be influenced by their peer academic push, therefore, peer push is also included as an explanatory variable.

Gender came from student self-reports obtained in the fall of 1987. Female was recoded as a dichotomous variable comparing females with males, with 0=male and 1=female. Using the recorded month and year of birth, age was calculated as the number of months since birth. The LSAY identified the ethnic background of students by six categories: (a) Hispanic, (b) Black, (c) White, (d) Asian, (e) Native American, and (f) others. Four dummy variables were created to represent race-ethnicity with White as the baseline category against Hispanic, Black, Asian, and Others including both Native American and others. As main explanatory variables for a student's social background, parents' SES were standardized composite variables constructed based on parents' self-reported education and occupational status, as well as student reported household possessions (Miller & Hoffer, 1994). Marital status and the number of children were obtained from the parent interviews. There were five categories of marital status: married, widowed, divorced, separated, and never married. One dummy variable was created to represent the number of parents from marital status, with married as 1 and other categories as 0, comparing both-parent families with single-parent families based on the base-year (1987 -1988) data. The number of siblings was created based on the number of children, which is a continuous variable.

Parent mathematics push is an equally weighted average of two variables. The variables included are 1) my parents expect me to do well in mathematics 2) my parents think mathematics is important. Peer mathematics push is an equally weighted average of four variables. The variables included are, 1) my friends like math; 2) my friends do well in math; 3) my friends hope to become scientists, doctors, engineers, or mathematicians; 4) my friends know how to write computer programs. There are three variables related to students' own attitudes toward mathematics: they enjoy mathematics; they are good at mathematics; they usually understand what they are doing at mathematics. Since these background information about students, their peers and their parents varied from wave to wave, instead of only including the information obtained in the first year, the mean value of parents' mathematics push and variance; the mean value of peers' mathematics push and variance are used as explanatory variables at the student level. The mean and variance of three factors of students' attitudes are also included.

Considering the above review, school context and school climate are utilized in the study. School enrollment size, location, and percentage of free-lunch students were used as independent school context variables. School location had three categories: urban school, suburban school, and rural school. Dummy recoding of school location created two variables with urban school as the baseline category against which suburban school and rural school were compared. The percentage of students eligible for federal lunch assistance was used to measure a school's socioeconomic composition. Other school climate variables, such as principal leadership, academic press, teachers' commitment, teaching experience, and extracurricular activities are included. All the variables are examined for extreme data, with corrections or deletions.

Data Analysis

In student growth studies, an example of hierarchical structured data occurs when repeated measurements over time are taken from individuals, who are in turn grouped within schools. Such structures are typically strong hierarchies since the variation within students in much smaller than the one between students. Here the repeated measurement constitutes the level 1 unit, with students representing level 2 units in a 3-level structure where the level 3 units are schools. The existence of such data hierarchies is neither accidental nor ignorable. Failure to consider the hierarchical nature of the data leads to unreliable estimation of the effectiveness of school policies and practices. Once the groupings are established, the group and group members both influence and are included by the group membership. In all instances mentioned above, the responses are no longer independent of each other. This factor may invalidate many of the traditional statistical analysis techniques, which assume the independence of the responses. Multilevel modeling is developed specially to account for correlated response variables at multiple levels; hence, it solves the dilemmas in the analysis of hierarchical data.

Few studies have focused on the nature of learning as a process of change over time. Although some researchers have considered longitudinal data, they used at most two time points, with the first measure functioning as a control for prior achievement in models predicting subsequent achievement (Wilkins & Ma, 2002). The multilevel model could estimate not only students' status but also their rate of growth in one subject. Furthermore, the effects of student characteristics and school composition on students' status and rate of growth could also be examined via the multilevel model for repeated measures data.

Null model

First, the level-one model is a simple linear growth model without any student-level variables and school-level variables. It is to model students' outcome scores on their grade levels:

$$Y_{ijk} = \alpha_{0jk} + \alpha_{1jk} * Grade_{ijk} + \varepsilon_{ijk}$$

where Y_{ijk}, α_{0jk}, and α_{1jk} represent the score, the initial status, and the rate of growth for j th student at i th year in k th school, respectively. And $Grade_{ijk}$ is the time at grade i for student j in school k. It is assumed that the errors ε_{ijk} are independent and normally distributed.

This model assumes that response variables are linearly related to time within each subject. However, growth may not be linear for all students over this age range. Non-linearity parameters such as the quadratic term need to be added to the model (Rasbash, 2002). Although adding parameters to a growth model can improve model accuracy, doing so increases the complexity of the model and should be done only when the advantages conferred by improved accuracy overweigh the disadvantages associated with greater complexity (Boyle & Willms, 2001).

At level 2, the intercept and slope from the level 1 model become dependent variables, modeled in two separate equations as a function of student-level variables. However, before any student-level variables are added into the equations, the initial status and the rate of growth in mathematics are only described as an average value (fixed effect) plus a variation (random effect). This approach provides an opportunity to examine not only the average values of initial status and rate of growth in mathematics, but also their variances and covariance. The estimates have been adjusted for measurement and sampling errors. This kind of simple models is named unconditional models in that no level 2 explanatory variables for either α_{0jk} or α_{1jk} have been introduced (Muller et. al., 2001).

Therefore, the unconditional level 2 models are:

$$\alpha_{0jk} = \beta_{00k} + \mu_{0jk}$$

$$\alpha_{1jk} = \beta_{10k} + \mu_{1jk}$$

where β_{00k} and β_{10k} represent the average initial status of students' mathematics performance and average rate of growth at secondary school k, and μ_{0jk} and μ_{1jk} represent random errors from the students.

Following the same pattern, the unconditional level 3 models are:

$$\beta_{00k} = \gamma_{000} + \nu_{00k}$$

$$\beta_{10k} = \gamma_{100} + \nu_{10k}$$

where γ_{000} and γ_{100} represent the average initial status of students' mathematics performance and average rate of growth at all secondary schools participating the study, and

ν_{00k} and ν_{10k} represent random errors at from the schools. The null model provides a measure of the variances within and between students and schools.

Full Model

The second step of analysis is to introduce between-student and between-school covariates, establishing a complex full growth model. The purpose is to use those covariates to explain the variation between students in schools and between schools regarding the initial status and rate of growth in mathematics. The student variables are added to the null model separately. Only the significant variables are retained to determine which one has a significant effect on the academic measures in the presence of other variables. Those student-level variables, which have a significant relative effect on the academic measures, are kept in the final model. It is similar to the forward elimination method in multiple linear regression analysis. A similar procedure is applied to the school-level variables at the third stage.

Student-level variables $X_{.jk}$ are added to the second-level multilevel model to model the initial status α_{0jk} and the rate of growth α_{1jk}. Thus, the conditional level-2 models are

$$\alpha_{0jk} = \beta_{00k} + \sum_p \beta_{01pk} X_{pjk} + \mu_{0jk}$$

$$\alpha_{1jk} = \beta_{10k} + \sum_q \beta_{11qk} X_{qjk} + \mu_{1jk}$$

where the parameters β_{00k} and β_{10k} represent the expected initial status and rates of growth for k th school after controlling student-level variables. β_{01pk} describes the relationship between the initial status of students' mathematics achievement α_{ojk} and student-level variable X_{pjk} at school k. β_{11qk} measures relationship between the student rate of growth α_{1jk} and student-level variable X_{qjk}. And μ_{0jk} captures the difference between each person's estimated initial status α_{ojk} and the average initial status β_{00k}, and the residual μ_{1jk} captures the difference between each person's estimated rate of growth α_{1jk} and the average rate of growth β_{10k}.

The conditional level 3 models are specified as follows:

$$\beta_{00k} = \gamma_{000} + \sum_{s} \gamma_{00s} Z_{sk} + \nu_{00k}$$

$$\beta_{10k} = \gamma_{100} + \sum_{t} \gamma_{10t} Z_{tk} + \nu_{10k}$$

Where the parameters γ_{000} and γ_{100} represent the expected average initial status and rates of growth for all the schools after controlling both student-level variables and school-level variables. γ_{00s} describes the relationship between the initial status of students' mathematics achievement β_{00k} and school-level variable Z_{sk} after controlling student-level variables. γ_{10t} measures relationship between the initial status of student growth rate in mathematics β_{10k} and school-level variable Z_{tk}. And ν_{00k} captures the difference between each school's estimated initial status and the average initial status γ_{000}, and the residual ν_{10k} captures the difference between each school's estimated rate of growth and the average rate of growth γ_{100}.

RESULTS

The unadjusted model contained neither student-level variables nor school-level variables. The results from MLwin are listed in Table 1. Hence, the correlation between the rates of growth among students is reported as 0.35, whereas the correlation between the rates of growth among schools is reported as 0.39.

Table 1: Mathematics Achievement Effect (Unadjusted Model)

	Mathematics Achievement	
	Effect	SE
Initial status	50.79	0.62
Rate of growth	3.40	0.08

Variance covariance matrix:

$$\Omega_v = \begin{pmatrix} 18.70 & 0.39 \\ 0.81 & 0.23 \end{pmatrix}$$

$$\Omega_f = \begin{pmatrix} 75.12 & 0.35 \\ 4.66 & 2.34 \end{pmatrix}$$

Note: The lower triangles of these matrices contain the variance and covariance; the upper triangles contain the correlations.

At student level, gender, age, mother's socioeconomic status, father's socioeconomic status and racial ethnicity have significant relative effects on the students' initial status of mathematics achievement. Therefore, within students, the initial status in mathematics is viewed as dependent on students' gender, age, number of parents and their racial ethnicity.

$$\alpha_{0jk} = \beta_{00k} - 3.073(Hispanic)_{jk} - 4.483(Black)_{jk} + 5.456(Asian)_{jk} - 4.063(Others)_{jk}$$
$$+ 0.871(Female)_{jk} + 1.058(Mothsei)_{jk} + 1.046(Fathsei)_{jk} - 0.275(Age)_{jk} + \mu_{0jk}$$

Within students, the rate of growth in mathematics is viewed as dependent on students' gender, age, number of parents and their racial ethnicity. The rate of growth is also dependent on parents' mathematics push and students' own attitudes, such as their enjoyment in learning mathematics and their self-esteem and confidence in learning mathematics.

$$\alpha_{1jk} = \beta_{10k} - 0.222(Hispanic)_{jk} - 0.420(Black)_{jk} + 0.184(Parents)_{jk} - 0.182(Female)_{jk}$$
$$+ 0.996(Parentpush)_{jk} - 0.245(MeanEnjoyment)_{jk} - 0.197(SDEnjoyment)$$
$$+ 0.437(MeanGoodAtMath) + 0.297(MeanUnderMath) - 0.043(Age)_{jk} + \mu_{1jk}$$

At the school level, school overall parental involvement status and the percentage of free lunch had effects on mathematics achievement at Grade 7. As to students' improvement in mathematics performance, only general support for mathematics had significant effect at the school level.

$$\beta_{00k} = \gamma_{000} + 1.389(ParentInvolvement)_k - 0.053(Freelunch)_k + \nu_{00k}$$

$$\beta_{10k} = \gamma_{100} + 0.648(GeneralSupport)_k + \nu_{10k}$$

Table 2: Mathematics Achievement Effect (Adjusted Model)

	Mathematics Achievement	
	Effect	SE
Initial status	50.84	0.35
Rate of growth	3.52	0.09

$$\Omega_v = \begin{pmatrix} 5.18 & -0.43 \\ -0.36 & 0.14 \end{pmatrix}$$

$$\Omega_f = \begin{pmatrix} 62.38 & 0.18 \\ 1.94 & 1.94 \end{pmatrix}$$

Note: The lower triangles of these matrices contain the variance and covariance; the upper triangles contain the correlations.

DISCUSSION

From the null model, students were found to have initial mathematics achievement score 50.79 at and grown average 3.40 points annually. After controlling for student and school characteristics, ”typical” students were found to have grown 3.52 points annually in their mathematics achievements starting from 50.84.

This study examined a variety of factors traditionally related to secondary mathematics achievement and growth. Many of them have been identified significantly related to secondary mathematics achievement and growth. Based on the full model, the gender gap in mathematics achievement appears early in secondary school, where female students were found to have a higher initial mathematics scores than male students. However, gender differences in mathematics achievement become less substantial as students progress though secondary school. Gender differences in mathematics achievement are declining as male students showed significant greater gains than females in mathematics through secondary school.

Asian American and White students showed higher mathematics achievement scores, as well as greater mathematics achievement gains, than their Hispanic and African American counterparts during secondary schools. In addition, these racial-ethnic differences on the mathematics tests were much substantial than gender difference and racial-ethnical differences tend to increase with age. As none of the interaction between gender and racial-ethnicity existed, gender difference within racial-ethnical categories are similar, as well as the achievement differences across racial-ethnical categories within female or male students.

Students from lower SES families were found to have lower initial mathematics achievement scores. Even though, these lower SES students didn’t show significant less growth in mathematics achievement over time. Therefore, the performance gap between lower- and higher- SES students hasn’t been widened by the time they reach 12th grade.

Younger students were found to perform better in mathematics than older students from the same grade cohort at Grade 8. It was also found that students from both-parents families grew faster than student from single-parent families in their mathematics achievements. This study indicated that the students from single-parent family were disadvantaged in the development of mathematical skills.

Parent mathematics push has a positive effect on the growth trajectory of students’ mathematic achievement in secondary schools. Students were also found to have improved at a faster rate when they have a more positive self-esteem related mathematics learning, such as they think that they are good at mathematics and they think they understand mathematics. Those who enjoyed learning mathematics didn’t improve their mathematics score faster.

At the school level, significant effects were associated with parent involvement and school’s percent of free lunch when initial mathematics achievement was studied. Parent involvement had a positive effect on the initial status of the mathematics achievement, while the percent of free lunch had a negative effect. The percentage of free lunch is an indicator of schools’ socioeconomic status. Resulting that the students from schools with lower socioeconomic status were disadvantaged in their mathematics skills when they entered secondary schools. At the school level, students in schools with a more positive general support toward mathematics grew at a faster rate than others.

CONCLUSIONS

African American and Hispanic students continue to perform far below whites and Asian Americans in terms of their secondary mathematics achievement. Researchers and educational practitioners need to continue to strive to reduce the racial-ethnic and gender gaps. Differences between racial-ethnic groups were generally larger than gender differences within groups. However, further research is still needed regarding the gender differences, especially existing in the growth rate of mathematics achievement at secondary schools.

Parents' socioeconomic status positively related to students' initial status – eighth-grade mathematics achievement. Students who go to schools with lower socioeconomic status usually had lower scores in mathematics. Parents' involvement, especially parents' mathematics push, helped students to improve themselves much faster. School background characteristics – general support toward mathematics had a significant positive effect on the growth trajectory of mathematics achievement. Therefore, this finding implies that schools should provide more support towards mathematics.

This study also found a positive correlation between the rate of growth and initial eighth mathematics achievement status from the null model. This shows that those students with the lowest levels of achievement in eighth grade also gain the least mathematics reasoning and knowledge during their secondary school years. Or, students who had a higher starting point also learned faster. Thus, the mathematics achievement gap continues to widen over time.

This study takes control over the school characteristics and the interaction effect among the student characteristics. It also takes account of the variation of some variables between the waves of the longitudinal study. However, the exact change of students' and schools' background information from year to year is not reflected in this study. Researchers need to systematically examine this issue. In this LSAY project, every year there were some students dropped out of the panel study. For that reason, the results of this study could be distorted when those students were not typical. Also by using the existing LSAY data set, the accuracy of this study may heavily depends on the quality of the data set (Babbie, 2002). This study is limited by only examining the existing variables covered by LSAY. As the LSAY project was conducted in 1980’s, it is necessary to reexamine those research questions on data sets from recent national panel studies; however, this study provides a critical baseline for comparison.

REFERENCES

Babbie, E. (2002). The Basics of Social Research, *2nd Edition.* Wadsworth Publishing.

Ma, X. (1999). Gender Differences in Growth in Mathematical Skills During Secondary Grades: A growth Model Analysis. *The Alberta Journal of Educational Research,* 4, 448-466.

Ma, X., & Klinger, D. A. (2000). Hierarchical Linear Modeling of Student and School Effects on Academic Achievement. *Canadian Journal of Education*, 25, 41-55.

Miller, J. C., & Hoffer, T. B. (1994). *Longitudinal Study of American Youth: Overview of Study Design and Data Recourses.* DeKalb, IL: Social Science Research Institute, Northern Illinois University.

Muller, P.A., Stage, F. K., & Kinzie, J. (2001). Science achievement growth trajectories: Understanding factors related to gender and racial-ethnic differences in precollege science achievement. American Education Research Journal. 38, 981-1012.

Wilkins, J. L., & Ma, X. (2002). Predicting student growth in mathematical content knowledge. *The Journal of Educational Research*, 95, 288-298.

Willms, J. D., & Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement,* 26, 209-232.

Young, D.J., Reynolds, A.J., & Walberg, H.J. (1996). Science achievement and educational productivity: A hierarchical linear model. *The Journal of Educational Research,* 89, 272-278.