Research Report

Growth in Reading Performance During the First Four Years in School

Donald A Rock

Research & Development

October 2007
RR-07-39

Growth in Reading Performance During the First Four Years in School

Donald A Rock

ETS, Princeton, NJ

October 2007

Abstract

This study addressed concerns about the potential for differential gains in reading during the first 2 years of formal schooling (K-1) versus the next 2 years of schooling (1^{st}–3^{rd} grade). A multilevel piecewise regression with a node at spring 1^{st} grade was used in order to define separate regressions for the two time periods. Empirical Bayes estimates of individual growth curves were developed, and smoothed estimates were plotted for selected demographic groups.

Key words: Reading growth, multilevel, longitudinal, piecewise regression, learning curves, early childhood, empirical Bayes, cognitive development

Table of Contents

List of Tables

List of Figures

Introduction

This study, using a national longitudinal database, investigates the shapes of the learning curves for selected groups of children based on demographic classifications. The study also examines gaps in early reading performance between groups based on gender, race/ethnicity, and parental education at time of entry to kindergarten and assesses whether those gaps become larger or smaller during the time span covering fall of the kindergarten year to spring of third grade. The transition from first to third grade is particularly critical, since it involves moving from preliteracy, with its emphasis on phonemic awareness and phonics instruction, to reading for comprehension. The National Research Council concluded in the report *Preventing Reading Difficulties in Young Children* (Snow, Burns, & Griffin, 1998) that academic success (as defined by graduation from high school) can be predicted with reasonable accuracy from a child's third grade reading performance. Other studies supporting the importance of third grade reading as a precursor of later academic success include the work of Anderson, Hiebert, Scott, and Wilkinson (1985) and the longitudinal growth curve analysis of Francis, Shaywitz, Stuebing, Shaywitz, and Fletcher (1996). The Francis et. al. study, as well as others (Fletcher et al., 1984; Stanovich & Siegel, 1994), focused on the question of whether differential reading performance in the early years is best characterized by a developmental lag model or a deficit model. The present study is more general in nature and is not focused on any specific model of reading development but essentially documents the shape of the developmental trajectories and explores how they differ by subpopulations of interest based on a nationally representative sample.

The National Reading Panel in the report on *Teaching Children to Read* National Institute of Child Health and Human Development (2000) carried out an extensive meta-analysis of reading instructional approaches and concluded that phonemic awareness instruction and phonics instruction in the first 2 years of formal schooling are critical to later reading performance. Because of the important instructional transition from phonemic awareness and phonics during the first 2 years of schooling to comprehension skills in the second and third grade, this study will investigate group changes in growth trajectories taking place during the first 2 years of schooling and compare these early trajectories with subsequent growth trajectories spanning spring first grade to spring third grade.

Sample

The study children were participants in the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) sponsored by the U. S. Department of Education, National Center for Education Statistics. A nationally representative sample of over 20,000 children who entered kindergarten in the fall of 1998 was surveyed and subsequently followed. The children were scheduled to be tested at entry to kindergarten and retested in spring kindergarten, spring first grade, and spring third grade. Approximately 17,400 children had reading test data at entry to kindergarten. This reduction in sample size was mostly due the fact that children who spoke English as a second language (ESL) did not take the reading test if they did not pass an English language screening test. A subset of the 17,400 children (n = 14,483) was selected based on whether they had panel data and whether they also provided school information at each of the subsequent retesting sessions.

Using the school information, the sample was further subdivided into those who remained in the same school from kindergarten through third grade versus those who attended two or more schools during this period. Sixty-one percent (n = 8,906) of the 14,483 children remained in the same school and had school sector information and panel data on the reading tests. This study is based on this sub-sample of 8,906 children who were tested in reading in fall kindergarten and who stayed in the same school from fall kindergarten to spring third grade. This number varies across testing occasions and shows some increase as one moves from fall kindergarten and up through the grades, since some ESL children passed the screener test and then could be tested in reading in the later follow-ups. The multilevel methodology used here does not require complete data on the repeated testings. It is somewhat surprising to note that there was little difference in initial reading performance at entry to kindergarten between those who remained in the same school and those who moved between schools. Children who stayed in the same school had a mean reading performance of 27.52, versus 27.36 for their more mobile counterparts (t = 0.36; p = .71).

Method

As pointed out above, this study is interested in possible changes in the rate of acceleration, or more likely de-acceleration, in the children's learning trajectories in reading as they move from phonics, emphasized in kindergarten through first grade, to reading comprehension, emphasized in the second and third grades. The reading measure used in ECLS-

K was designed to measure developmental growth in reading and was vertically scaled using the methods of item response theory (Embretson & Reise, 2000; Lord, 1980) within a two-stage adaptive testing situation (Cleary, Linn, & Rock, 1968). The use of individually administered adaptive tests insured that there would be little chance of having floor or ceiling effects, the bane of all longitudinal change measurement. The ECLS-K reading test was also criterion referenced, marking eight hierarchically ordered reading milestones.

In order to describe this possible change in growth rates, a piecewise regression model (Seber & Wild, 1989) with a node marking the average age at the third testing point (spring first grade) was formulated and estimated using multilevel software MLwiN (Goldstein, 1998) and STATA (Stata Corporation, 2005) for the graphics. The MLwiN multilevel or mixed model approach with three levels was appropriate here, because ECLS-K used a multistage sampling procedure, sampling schools and then children within schools and finally taking repeated measures on children. The multilevel estimation procedure took into consideration the clustering effects at each stage and thus provides more honeststandard errors. This formulation of the piecewise regression allows for different growth rates for the period from kindergarten entry to spring first grade and for the period from spring first grade to spring third grade. An alternative formulation would be a simple quadratic function, which would include age at time of testing as both a first order term and a squared term. While the simple quadratic model requires fewer parameters, it is not quite as easy to interpret as are changes in linear slopes centered about a particular critical point in time (i.e., the node at spring first grade). Also, not surprisingly, estimates based on the quadratic regressions did not fit as well with respect to mean squared residuals as did those based on the piecewise regression model.

The general form of the piecewise regression used here can be described by the following equation:

$$Y_{ijk} = \beta_{0ijk} + \beta_{1jk} age87_{ijk} + \beta_{2jk} age2_{ijk} + \beta_{3jk} \operatorname{int} 2_{ijk} \qquad (1)$$

where each term is defined as follows:

β_{0ijk} is the constant and is random at both the school and child levels and within child. It is the predicted reading mean for a child who is just under 86.70 months old.

β_{1jk} is the slope of reading on age between entry to kindergarten and spring first grade and is random at both school and child levels.

β_{2jk} is the *change* in slope of reading on age between spring first grade and spring third grade and is random at both school and child levels.

β_{3jk} is the predicted additive constant to β_0 indicating the intercept associated with the slope of reading on age from spring first grade to spring third grade.

Y_{ijk} is the reading scale score for the i_{th} occasion for the j_{th} child in the k_{th} school.

The following coding scheme makes the above interpretation possible. Age87 is the child's age deviated from the mean age of the children at time of testing in spring first grade (mean age at time of testing in first grade was 86.70 months). Age2 is the age centered around the spring first grade mean but converted to 0's for ages below the mean. Int2 is coded 1 if age => 86.70 and 0 otherwise.

There are three primary questions that the above general model and its variations can speak to. First is the test of the coefficient β_{2jk}, which will indicate if there is a significant increase or decrease in the acceleration in rate of learning as a child transits from preliteracy (K-first grade) to more complex reading comprehension skills taking place between spring first grade and spring third grade. Second, since β_{1jk} and β_{2jk} are random at both school and child levels, one can compare and test the between-school variance of the rates of learning before spring first grade and after spring first grade. If the between-school variance associated with the pre spring first grade rate of learning is greater than that between spring first grade and spring third grade, then one of many possibilities would be that the quality of schooling is particularly important during the preliteracy period when it is contrasted with that of the later literacy period. Similarly, if the between-child variation in learning rates is greater during the period from kindergarten to first grade than that for the period from spring first grade to spring third grade, the data would suggest that there is a reduction in the fan-spread effect as one goes from the first 2 years of schooling to the second 2 years of schooling.

Third, background variables will be introduced in the above model as fixed effects, including main effects and interactions, with the slopes describing rates of learning. For example, if a dummy variable for a child's gender is introduced and crossed with the slopes and intercepts, one can test whether girls and boys have the same learning rates before and after spring first grade. If there are differential rates of gain for the two gender groups, one can ask whether the differential growth rates vary by the material being learned. That is, do boys have proportionately greater problems learning the reading comprehension tasks emphasized in Grades 2-3 compared to the phonics-based tasks emphasized during the K-1 instructional sequence? Then, based on the above models, empirical Bayes smoothed estimates of learning curves will be plotted for gender, race/ethnicity, parent education, and groupings defined by the skill level in which the children were making gains.

It should be kept in mind that the fixed effects in a multilevel model reflect slopes and intercepts that essentially are pooled within estimates. For example, within a particular school, all children are assumed to have the same regression function (i.e., the mean within-school regression, which is then averaged across schools). When one includes interaction effects with say, gender groups as fixed effects, then one averages within-school differences in the two within-school regression lines across schools and tests the differences between the average within-school regression lines for males and females The differences between subgroup regressions based on the fixed effects (i.e., the within-school part of the overall multilevel mode), may or may not completely reflect what is happening in terms of differential growth rates, unless the subgroups in question have the same proportional distributions across schools that vary in performance levels.

The model-based empirical Bayes estimates, however, allow for different regression functions for individual children, albeit shrunken towards the school mean or subgroup mean within the school, but they also combine this information with that from the between-school regressions and as a result have a smaller mean square error and thus are preferable for plotting individual points and then fitting a smoothing function to represent the different group growth curves. The upshot of all this is that, while it is convenient to do statistical tests on the fixed effects and the variability around the fixed effects (the random components), it is the empirical Bayes plots that best describe what is going on, since they combine both pieces of information in estimating the growth curves. When one finds quite different patterns of rates of gain comparing the fixed-effects results with the empirical Bayes smoothed estimates, it is likely to be due to disproportionate distributions of the subgroups in question across schools varying considerably in performance. This latter situation is, however, a reflection of the real world.

Results and Discussion

Figure 1 presents the results of the estimation of Equation 1 using the MLWIN multilevel software.

$$\text{reading}_{ijk} \sim \text{N}(XB,\ \Omega)$$
$$\text{reading}_{ijk} = \beta_{0ijk}\text{cons} + \beta_{1jk}\text{age87}_{ijk} + \beta_{2jk}\text{age2}_{ijk} + \beta_{3jk}\text{int2}_{ijk}$$
$$\beta_{0ijk} = 63.942(0.386) + v_{0k} + u_{0jk} + e_{0ijk}$$
$$\beta_{1jk} = 1.978(0.014) + v_{1k} + u_{1jk}$$
$$\beta_{2jk} = -0.380(0.021) + v_{2k} + u_{2jk}$$
$$\beta_{3jk} = 6.731(0.219) + v_{3k} + u_{3jk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \\ v_{3k} \end{bmatrix} \sim \text{N}(0,\ \Omega_v)\ :\ \Omega_v = \begin{bmatrix} 87.654(5.819) & & & \\ 2.709(0.206) & 0.101(0.008) & & \\ -3.390(0.296) & -0.128(0.012) & 0.211(0.019) & \\ 6.173(2.360) & 0.312(0.088) & -0.482(0.126) & 10.029(1.986) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \\ u_{3jk} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 399.395(8.479) & & & \\ 12.807(0.326) & 0.472(0.014) & & \\ -18.254(0.462) & -0.653(0.021) & 1.157(0.034) & \\ -43.290(5.614) & -1.172(0.278) & 0.234(0.293) & 98.798(6.886) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 27.403(0.765) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 306517.800(39656 of 42432 cases in use)

Figure 1 **.Multilevel estimates of the rate of gain (β_1) before spring first grade and the change in that rate (β_2) after spring first grade.**

Inspection of the fixed effects and their associated regression coefficients (standard errors in parenthesis) indicates that the average child's rate of gain during the period from kindergarten entry to spring first grade is equal to 1.98 score points per month (β_{1jk} = 1.98, χ^2_{1df} = 19027.27, p = .00). During the transition from spring first grade to spring third grade, the rate of gain slowed significantly to 1.60 score points per month (($\beta_{1jk} + \beta_{2jk}$ = 1.98 + (-0.38)). The reduction in slope or rate of gain was also highly significant (χ^2_{1df} = 328.70; p = .00). It would appear that the transition from mastering phonemic awareness and phonics skills to reading for comprehension is not a trivial step in the typical child's early reading development.

The first 4 x 4 matrix (Ω_v) shown in Figure 1 presents the variance-covariance matrix including standard errors (in parenthesis) of the school-level random components associated with the coefficients in Equation 1. For example, 87.65 is an estimate of the school-level variance of the intercept just before the node in the piecewise regression, while 0.10 is the school-level variance of the slope of reading on age between kindergarten and spring first grade, and 0.21 is the school-level variance of the change in slope as one goes from spring first grade to spring third grade. The last main diagonal element is the school-level variance of the intercept at the node in the piecewise regression. The second 4 x 4 matrix (Ω_μ) is the corresponding matrix for the child-level random components.

With respect to a comparison of the variability at both the school level and child level of the rate of learning between kindergarten and spring first grade with that of spring first to spring third grade, a different parameterization from that of Equation 1 must be used. Equation 2 shows the re-parameterization:

$$Y_{ijk} = \beta_{1jk} age1 + \beta_{2jk} age2 + \beta_{3jk} \operatorname{int}1 + \beta_{4jk} \operatorname{int}2 \qquad (2)$$

where age2 and int2 remain the same and age1 is centered around age 86.70 and coded zero for ages => 86.70, and int1 is coded 1 for ages below 86.70 and zero thereafter.

This estimation was carried out using the no-constant option for the fixed effects. In this re-parameterization, β_{1jk} is the slope of reading on age (rate of gain) for the period kindergarten through spring first grade, while β_{2jk} is the slope of reading on age (rate of gain) for the period spring first grade to spring third grade. It should be remembered that the regression weight for age2 in Equation 1 estimated the change in slope and not the slope itself. To get at and test the relative variability of rates of learning before and after spring first grade, we need to directly estimate the slopes β_1 and β_2 and their variability across schools and students, as is done in Equation 2. Table 1 presents the school-level and child-level variance-covariance matrices that apply to the slopes and intercepts before and after spring first grade based on Equation 2.

The first two main diagonal elements in Table 1 at both the school level (.10 and .05) and

Table 1

School and Student Level Variance-Covariances of the Rates of Gain Slopes and Intercepts Before and After Spring First Grade

		School level		
Age1	10 (.01)			
Age2	- .03 (.01)	.05 (.01)		
Int1	2.71 (.27)	-.68 (.15)	87.68 (5.82)	
Int2	3.02 (.22)	-.85 (.17)	93.88 (6.00)	110.11 (7.04)
		Student level		
Age1	.47 (.01)			
Age2	-.18 (01)	.32 (.01)		
Int1	12.81 (.33)	-5.46 (.25)	399.53 (8.48)	
Int2	11.64 (.32)	-6.39 (.27)	356.35 (7.18)	411.88 (8.41)

child level (.47 and .32) indicate that between-school variability and between-child variability within schools with respect to rates of gain are decreasing as children go from their first 2 years of schooling to the second 2 years of schooling. Both these reductions in variance are statistically significant (χ^2_{1df} = 24.24, p = .00 and χ^2_{1df} = 75.01, p = .00) respectively. If one transforms the variance-covariance matrices to correlation matrices, the correlation between the rate of gain before spring first grade with that after first grade is -.37 at the school level and -.46 at the child level. Clearly the greater the rate of gain during the period before spring first grade, the lesser the gain during the period following spring first grade. The reduction in variability in rate of gain after first grade at both the school level and between children within schools is, of course, consistent with the de-acceleration in the rate of gain following first grade. The negative correlations between slopes pre and post first grade at both the school level and the child level add additional evidence for slower growth for children at upper levels of the vertical scale and thus explain the decrease in variability. This leads to the question: is the slowing of growth for children at the upper end of the scale due to ceiling effects in the test, or is it primarily coming from the children moving from mastering preliteracy phonics-based tasks to mastering more complex comprehension skills with a larger language component? The adaptive reading test used in ECLS-K was developed to measure growth along a behaviorally anchored scale with eight ascending milestones, or proficiency levels, where the hierarchy reflects both cognitive demand and instructional sequence. Using procedures outlined in the ECLS-K user manual (National Center for Education Statistics, 2004) and elsewhere, Rock and Pollack (2002), one can uniquely identify for each child where on the test score scale they are making their gains and with respect to which of the proficiency levels they are showing their maximum growth.

Table 2 presents the percentages of children who are making their maximum growth at each of the criterion-referenced points during the period spanning spring first to spring third grade, the time span where one would expect ceiling effects if present. For example, 43% (in the total column) were making their gains on the vertical scale in the neighborhood of simple literal inference. This being the modal percentage by far, one might infer that children who are on grade level are likely to be making their greatest gains at this proficiency level. Similarly, the 20.6% who are making their maximum gains in the two highest levels, extrapolation and evaluation, are likely to be doing above-grade work. It would be fairly safe to conclude that many of the differences in percentages reflect what is emphasized in the curriculum that spans the first through the third grade. If one computes the mean scale score for the children who made their maximum gains in the highest level (evaluation), the mean (139.2) and standard deviation (3.7) put them about 2.4 standard deviations below a maximum score.

Table 2

Percentage of Children Making Maximum Gains at Each of the Criterion-Referenced Points Between Grades 1 and 3

Locus of maximum gain	%
Letter recognition	1.0
Beginning sounds	0.2
Ending sounds	0.8
Sight words	6.6
Comprehension of words in context	27.4
Literal inference	43.3
Extrapolation	17.9
Evaluation	2.7
Total	100.0

Another way of looking at potential ceiling effects for the highest proficiency level is in terms of the skewness of the distribution of gain scores. The skewness index for the gain scores for those making their gains at the highest level was .02, with a probability of .85. It would seem that the curriculum emphasis and the complexity of the literacy skills required for gains at the very upper ends of the scale contributed to the slowing down of growth near the top of the scale. It is also felt that the moving from phonics to beginning comprehension and finally to out-of-grade literacy tasks that may require an increasing language component may also have contributed to a structural or developmental ceiling.

These early results call into question not so much the quality of schooling for below- and on-grade children, but the possibility that the curriculum may need to direct more attention to emphasizing vocabulary development to facilitate growth in reading comprehension and help with the transition from phonics to more language dependent reading skills. The next section, which looks at differential growth rates for ethnic/racial groups, will help throw some light on the possibility that an increased emphasis on language as one moves up the test score scale may be contributing to the leveling off of growth rates.

Differential Growth Rates for Racial/Ethnic Groups

Figure 2 presents the results of the multilevel analysis of the piecewise regression model with the addition of dummy variables for racial/ethnic groups and their interactions with the slopes and intercepts. This procedure provides fixed-effects tests of differential gains by the different racial/ethnic groups both before and after first grade.

Inspection of the coefficients associated with the interactions of the ethnic groups with age87 (e.g., ag87xwht, ag87xhisp, ag87xasian, and ag87xothr) reflect deviations of the respective groups’ rates of gain during the first 2 years of schooling from that of the base group consisting of the Black children. Similarly, the coefficients associated with the interactions of the ethnic groups with age2 (e.g., age2xwht, age2xhisp, age2xasian, and age2xothr) reflect deviations of the respective groups from the change in growth rates for the Black children as they go from the first 2 years of schooling to the next 2 years of schooling. For example, the White children’s group rate of gain for the first 2 years of schooling was 0.24 score points per month more than that of the Black children. The ratio of 0.24 to its standard error in parenthesis equals

$$\text{reading}_{ijk} \sim \mathrm{N}(XB, \Omega)$$

$$\begin{aligned}\text{reading}_{ijk} = {} & \beta_{0ijk}\text{cons} + \beta_{1jk}\text{age87}_{ijk} + \beta_{2jk}\text{age2}_{ijk} + \beta_{3jk}\text{int2}_{ijk} + 6.279(0.792)\text{white}_{jk} + 0.069(0.900)\text{hisp}_{jk} + 9.805(1.391)\text{asian}_{jk} + 3.919(1.373)\text{other}_{jk} + \\ & 0.244(0.031)\text{ag87xwht}_{ijk} + 0.079(0.037)\text{ag87xhisp}_{ijk} + 0.314(0.059)\text{ag87xasian}_{ijk} + 0.156(0.054)\text{ag87xothr}_{ijk} + -0.046(0.049)\text{age2xwht}_{ijk} + \\ & 0.091(0.057)\text{age2xhisp}_{ijk} + -0.426(0.084)\text{age2xasian}_{ijk} + -0.141(0.078)\text{age2xothr}_{ijk} + 1.322(0.591)\text{int2xwht}_{ijk} + -0.591(0.734)\text{int2xhisp}_{ijk} + \\ & 1.345(1.079)\text{int2xasian}_{ijk} + -0.270(1.013)\text{int2xothr}_{ijk}\end{aligned}$$

$$\beta_{0ijk} = 59.695(0.722) + v_{0k} + u_{0jk} + e_{0ijk}$$

$$\beta_{1jk} = 1.802(0.029) + v_{1k} + u_{1jk}$$

$$\beta_{2jk} = -0.348(0.045) + v_{2k} + u_{2jk}$$

$$\beta_{3jk} = 6.062(0.528) + v_{3k} + u_{3jk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \\ v_{3k} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_v)\ :\ \Omega_v = \begin{bmatrix} 75.776(5.416) & & & \\ 2.277(0.192) & 0.085(0.008) & & \\ -3.107(0.282) & -0.119(0.011) & 0.203(0.019) & \\ 2.144(2.212) & 0.166(0.082) & -0.408(0.122) & 8.625(1.918) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \\ u_{3jk} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 396.095(8.414) & & & \\ 12.732(0.326) & 0.471(0.015) & & \\ -18.160(0.462) & -0.650(0.021) & 1.148(0.034) & \\ -43.444(5.641) & -1.200(0.279) & 0.269(0.294) & 99.740(6.872) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 27.403(0.765) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 305626.500(39608 of 42432 cases in use)

***Figure 2*. Multilevel piecewise regression for race/ethnicity.**

0.24/0.03 = Z of 7.87 (p = .00). In general, we will report the results for the Wald (1943) large sample χ^2 test rather than the Z test because of its ease in making joint tests. The largest increment in rate of gain during the first 2 years of schooling over that of Black children was that of the Asian children (0.31, χ^2_{1df} = 28.88, p = .00) which is not significantly greater statistically than that of the White children (χ^2_{1df} = 1.73, p = .19). The total rate of gain of the White children during the first 2 years of schooling is equal to the Black rate of gain plus the White increment (i.e., β_{1jk} = 1.80 + 0.24, or 2.04 score points per month). Similarly the Asian children's total rate of gain is 1.80 + 0.31, or 2.11 score points per month. Since the Black children are the contrast group, β_{1jk} gives the Black rate of gain (1.80 score points per month) during the first 2 years of schooling.

When the groups are ordered with respect to their rates of gain from high to low during the first 2 years of schooling, we have Asian children with the greatest rate of gain followed by the White children, and then the Other children followed by the Hispanic children and then the Black children. All of these group contrasts with the rate of gain for Black children are statistically significant. Clearly the gap between Asian and Black children, White and Black children, Other and Black children, and to a lesser extent the Hispanic and Black children is increasing over the first 2 years of schooling. The same is true, although to a lesser extent, with respect to the Asian and Hispanic gap (χ^2_{1df} = 18.66, p = .00) and the White and Hispanic gap (χ^2_{1df} = 34.94, p = .00). It should be kept in mind that while there are significant differential gains during the first 2 years of schooling based on the fixed effects, all groups are making significant gains.

The regression weights associated with the interaction terms of the racial/ethnic groups with age2 estimate the change in rates of gain as children move through the second 2 years of schooling. For example, the Black children, who are the base or comparison group, gained 1.80 score points per month during the first 2 years of school, and then their rate of gain declined to 1.80 - 0.35, or 1.45 points per month during the second 2 years of schooling. The change in the rate of gain for Whites in the second 2 years of schooling is 0.05 score points less than the change for Black children, which is not significant (p = .37). Thus the estimated rate of gain for White children for the second 2 years of schooling is equal to the rate of gain for White children during the first 2 years of schooling (2.05) plus the White change contrasted with the Black change (-0.05 + (-0.35)), giving a rate of gain of 1.65 score points per month during the second 2 years of schooling. While this is a relatively large decline, the rate of gain during the second 2 years still favors the White children compared to the Black children (i.e., 1.65 vs. 1.45; χ^2_{1df} = 40.66, p = .00). A more pronounced decline in rate of gain was found for Asian children when contrasted with the Black children base group as they transited from the first 2 years of schooling to the second 2 years. That is, the Asian children's decline in rate of gain was 0.77 score points per month when going from the first 2 years of schooling to the second (p = . 00). Thus the rate of gain for Asian children during the second 2 years of schooling is equal to 2.12 + (-0.43 + (-0.35)), or 1.34 score points per month. This is a relatively large decline in the rate of gain (i.e., going from 2.12 during the first 2 years to 1.34 for the last 2 years). While there was a significant difference in rate of gain favoring Asian children over Black children during the first 2 years of schooling, there is a significant difference in rates of gain between the two groups during the second 2 years of schooling, favoring the Black children. That is, Black children gained 1.80 - 0.35 = 1.45 score points per month during the second 2 years of schooling, while Asians gained 1.34 points per month during the same time period, yielding a significant difference with the rate of gain favoring the Black children. (1.45 - 1.34 = 0.11; χ^2_{1df} = 4.71, p = .03).

Table 3 presents a summary of the rates of gain based on the fixed effects and also on the more complete empirical Bayes (EB) smoothed estimates for each of the racial/ethnic groups for the first 2 years of schooling and separately for the second 2 years. While the patterns are relatively similar, there are some differences. When the school means are included in the estimation (empirical Bayes estimates), all subgroups show an increment in rates of gain during the first 2 years and less of a decline in the second 2 years. The White children's EB estimates show a greater increment when the EB estimate is contrasted with the fixed estimate during the first 2 years of schooling. This is probably because a greater proportion of the White students are shrinking towards school means that tend to be at the higher performance levels. During the second 2 years of schooling, the Hispanic children show proportionately less benefit from the inclusion of school mean effects in their EB estimates.

Table 4 presents a summary of the significance tests between group rates of gain both before and after spring first grade based on the fixed-effects estimates.

Table 3

Fixed and Model Based Empirical Bayes Estimates of Rates of Gain by Racial/Ethnic Groups for Grades K-1 Versus Grades 1-3

	Slope K–1		Slope 1st–3rd	
Group	Fixed	EB	Fixed	EB
Asian	2.12	2.27	1.34	1.56
White	2.05	2.32	1.65	1.86

Table 4

Fixed Effect Differences in Rates of Gain by Racial/Ethnic Groups, Grade

	White	Asian	Hispanic	Black
		K-1		
White				
Asian	.07 (.18)			
Hispanic	-.16 (.00)	-.24 (.00)		
Black	-.24 (.00)	-.31 (.00)	-.08 (.03)	
		Grades 1-3		
White				
Asian	-.31 (.00)			
Hispanic	-.03 (.31)	.28 (.00)		
Black	-.20 (.00)	.11 (.03)	-.17 (.00)	

Note. Probabilities shown in parenthesis. For the fixed effect differences, column variables are subtracted from the row variables.

The intercepts are not included in this discussion, since they are primarily determined by the slopes in this model, and sensible interpretations of intercept differences both between and within groups depends on the slopes being equal, which is not the case here. While the complicated contrasts moving through the two 2-year spans are necessary for statistical tests of the fixed effects, the plots of the smoothed empirical Bayes predicted growth curves provide a better summary of the group trajectories.

Figures 3 and 4 present separate growth curves for the racial/ethnic groups based on empirical Bayes shrunken estimates from the above model and then plotted using a quadratic smoothing function. The empirical Bayes estimates have the property of minimum mean squared error. Figure 3 shows the growth curves for the White, Asian, and Black children, while Figure 4 contrasts the White and Hispanic children. The vertical lines in the graphs represent the mean age of the children when they were tested at fall kindergarten entry, spring first grade, and spring third grade. The reader should not attempt to interpret the growth curves beyond the mean age vertical lines representing the entry at kindergarten (the leftmost vertical line) and the spring third grade vertical line. As discussed above, the plots in Figure 3 show a steadily increasing gap between White and Black children. However, the most interesting comparison is the contrast between the plots of the White and Asian children. The Asian plot essentially shows the same rate of gain as that of the White children during the first 2 years of school, but then the Asian children's rate of gain during the second 2 years of schooling declines sufficiently for the White children to overtake them by spring of third grade. Figure 4 shows some divergence in the rates of gain between the White and Hispanic children's plots, with the rates of gain favoring the White children and the gap widening between the White and Hispanic children as they progress through the second 2 years of schooling.

The relatively large decline for the Asian children lends credence to the theory that well prepared second-language children who proceed rapidly through the phonics tasks may have proportionately more trouble when they reach more language-based tasks involved in reading comprehension. It should also be pointed out that those children who proceed relatively quickly with their mastery of the phonics-based skills during the first 2 years of schooling are then faced with making gains at the upper end of the test score scale, which consists of the more difficult reading comprehension exercises. Conversely, those children who were making smaller gains

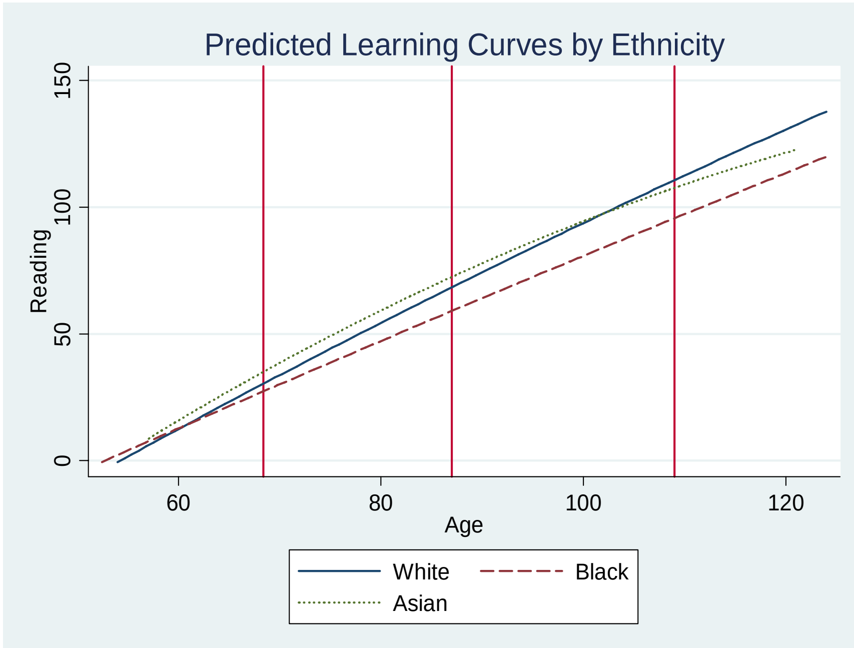

***Figure 3*. Predicted learning curves in reading by ethnicity for White, Black, and Asian students.**

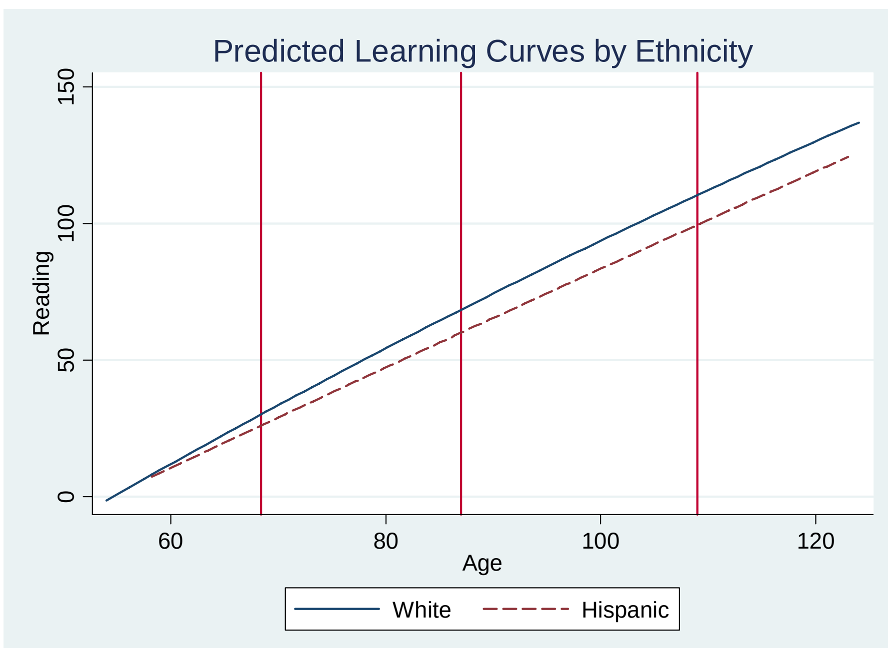

***Figure 4*. Predicted learning curves in reading by ethnicity for White and Hispanic students.**

during the first 2 years will still be making their gains at the lower and middle levels of the test score scale, which are characterized by the easier tasks. The fact that children are making their gains at different points on the test score scale reflecting different levels of complexity brings to mind the analogy of a bicycle race. That is, cyclists who are leading eventually come to a mountain, while their pursuers are still on the flat areas. While the leaders are coping with the difficult uphill terrain, the pursuers are gaining because they are coping with the flat easy terrain. The potential for the differential impact of language on the growth pattern of Asian children during the second 2 years of schooling is given further credence by a review of the location of their gains on the test score scale, as shown in Table 5.

About a quarter of the Asian children are making their gains in extrapolation and evaluation. Only the White non-Hispanics have such a large proportion of their children making gains at these more language-loaded tasks defining the upper ends of the test score scale. Proportionately more of the Asian and White children are climbing the mountain during the second 2 years of schooling.

Differential Growth Rates by Gender

Figure 5 presents the results of the multilevel analysis with interactions with the gender dummy variable. In this analysis the boys are the base group.

Inspection of the gender solution and the regression weight associated with ag87xgirl suggests that, while girls grew faster than boys in the first 2 years of school (2.06 score points more per month, versus the boys' 1.90 score points per month $\chi_{1df}^{\ 2}$ = 74.14, *p* = .00), there was no significant difference between their rates of growth during the second 2 years of schooling (1.61 points per month for boys and 1.58 points per month for girls, $\chi^2_{\ 1df}$ = 3.23, *p* = .07. The pattern was almost identical when the empirical Bayes estimates were contrasted for the two gender groups. For the first 2 years of schooling the girls' rate of gain was 2.27 points per month, while boys gained 2.14 points per month. During the second 2 years of schooling the comparable figures were 1.75 and 1.78 respectively. It is interesting to note that the empirical Bayes estimates of the differentials in rate of gain are very close to the fixed-effects estimate for the first 2 years of schooling and exactly the same for the second 2 years. This agreement stems from the fact that one would expect girls and boys to have very similar distributions across schools

Table 5

Percentage of Children Within Racial/Ethnic Groups Making Maximum Gains at Each of the Criterion-Referenced Points

Locus of maximum gain Grades 1–3	White	Black	Hispanic	Asian	Total
	%	%	%	%	%
Letter recognition	-	-	5.3	1.4	1.0
Beginning sounds	0.1	0.4	0.2	-	0.2
Ending sounds	0.4	2.2	1.1	0.5	0.8
Sight words	3.9	12.9	9.7	5.5	6.6
Comprehension of words in context	21.8	40.3	36.4	23.9	27.4
Literal inference	47.4	37.5	35.2	41.8	43.3
Extrapolation	22.8	6.1	11.0	24.0	17.9
Evaluation	3.7	0.5	1.1	2.9	2.7
Totals	100.0	100.0	100.0	100.0	100.0

Note. Groups labeled White and Black are non-Hispanic.

varying in performance. Figure 6 presents the empirical Bayes plots for the two gender groups.

Inspection of Figure 6 shows girls starting out at entry to kindergarten slightly ahead of boys, and this gap increases until spring first grade and then the gap is essentially the same if not slightly decreasing until spring third grade.

Parent Education and Differential Growth Rates

Figure 7 presents the multilevel fixed-effects estimates of slopes and changes in slopes for three different groups based on the working parent’s education level. If both parents were working, the parent with the highest education level was chosen. The three parental groups were:

$$\text{reading}_{ijk} \sim \mathrm{N}(XB, \Omega)$$

$$\text{reading}_{ijk} = \beta_{0ijk}\text{cons} + \beta_{1jk}\text{age87}_{ijk} + \beta_{2jk}\text{age2}_{ijk} + \beta_{3jk}\text{int2}_{ijk} + 5.672(0.446)\text{girl}_{jk} + 0.164(0.019)\text{ag87xgirl}_{ijk} + -0.195(0.027)\text{ag2girl}_{ijk} + -0.408(0.394)\text{int2xgirl}_{ijk}$$

$$\beta_{0ijk} = 61.136(0.440) + v_{0k} + u_{0jk} + e_{0ijk}$$

$$\beta_{1jk} = 1.897(0.018) + v_{1k} + u_{1jk}$$

$$\beta_{2jk} = -0.283(0.025) + v_{2k} + u_{2jk}$$

$$\beta_{3jk} = 6.923(0.288) + v_{3k} + u_{3jk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \\ v_{3k} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_v) \ : \ \Omega_v = \begin{bmatrix} 86.803(5.780) & & & \\ 2.676(0.204) & 0.100(0.008) & & \\ -3.345(0.294) & -0.127(0.011) & 0.209(0.019) & \\ 6.142(2.345) & 0.312(0.087) & -0.482(0.126) & 10.054(1.987) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \\ u_{3jk} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u) \ : \ \Omega_u = \begin{bmatrix} 391.595(8.389) & & & \\ 12.585(0.324) & 0.466(0.014) & & \\ -18.000(0.461) & -0.646(0.021) & 1.149(0.034) & \\ -42.489(5.551) & -1.142(0.276) & 0.208(0.292) & 98.521(6.868) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_e) \ : \ \Omega_e = \begin{bmatrix} 27.404(0.765) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 306298.600(39656 of 42432 cases in use)

***Figure 5*. Multilevel piecewise regression for gender.**

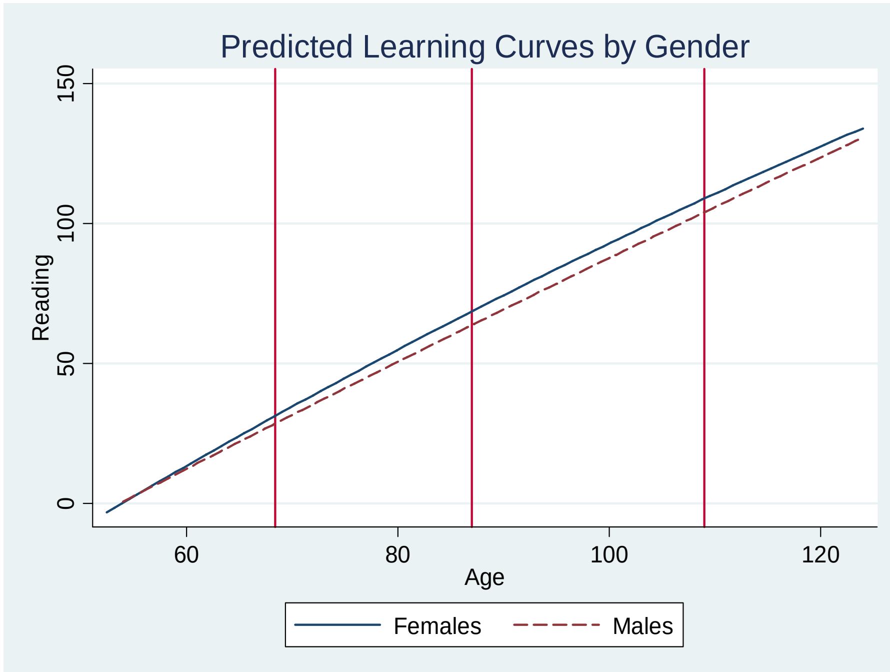

***Figure 6*. Predicted learning curves in reading by gender.**

(a) college degree or higher, (b) some college completed, and (c) high school degree or less. As before, the coefficients associated with the group interactions with age87 give the increment in the mean rate of gain for the children from each of the parental groups during the first 2 years. High school degree or less was the base group for all contrasts. Not surprisingly, the children of college-educated parents made the greatest gains per month, followed by those whose parents had some college, and then the base group made the smallest per-month gains. The per-month gains during the first 2 years of schooling for these three groups were 2.16, 2.00, and 1.81 for the base or high school group. These increments in rate of gain over the base group for children of college degreed parents and those children with parents with some college were all statistically significant (χ^2_{1df} = 176.84, p = .00; χ^2_{1df} = 70.59, p = .00) respectively. The children of college degreed parents gained significantly more than children with parents with some college (χ^2_{1df} = 38.64, p = .00) during the first 2 years of schooling. The rate of gain during the second 2 years of schooling shows the usual pattern of decline in the rate of growth, with slightly larger de-accelerations for children coming from families with better educated parents. The rate of gain for

$$\text{reading}_{ijk} \sim \text{N}(XB,\ \Omega)$$

$$\begin{aligned}\text{reading}_{ijk} = {} & \beta_{0ijk}\text{cons} + \beta_{1jk}\text{age87}_{ijk} + \beta_{2jk}\text{age2}_{ijk} + \beta_{3jk}\text{int2}_{ijk} + 6.409(0.536)\text{smcol}_{jk} + 12.607(0.643)\text{cold}_{jk} + 0.347(0.026)\text{ag87xcold}_{ijk} + 0.194(0.023)\text{ag87xsmcol}_{ijk} + \\ & -0.347(0.040)\text{ag2xcold}_{ijk} + -0.161(0.035)\text{ag2xsmcol}_{ijk} + 1.470(0.489)\text{int2xcold}_{ijk} + 0.277(0.500)\text{int2xsmcol}_{ijk}\end{aligned}$$

$$\beta_{0ijk} = 58.074(0.467) + v_{0k} + u_{0jk} + e_{0ijk}$$

$$\beta_{1jk} = 1.811(0.019) + v_{1k} + u_{1jk}$$

$$\beta_{2jk} = -0.223(0.029) + v_{2k} + u_{2jk}$$

$$\beta_{3jk} = 6.247(0.357) + v_{3k} + u_{3jk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \\ v_{3k} \end{bmatrix} \sim \text{N}(0,\ \Omega_v)\ :\ \Omega_v = \begin{bmatrix} 55.728(4.453) & & & \\ 1.766(0.164) & 0.072(0.007) & & \\ -2.472(0.251) & -0.100(0.010) & 0.180(0.018) & \\ 0.143(2.045) & 0.158(0.081) & -0.306(0.124) & 9.287(1.958) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \\ u_{3jk} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 387.463(8.322) & & & \\ 12.417(0.322) & 0.460(0.014) & & \\ -17.931(0.458) & -0.639(0.021) & 1.141(0.034) & \\ -43.147(5.668) & -1.242(0.277) & 0.230(0.297) & 101.719(6.940) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 27.415(0.761) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 296591.400(38490 of 42432 cases in use)

***Figure 7*. Multilevel model estimates for parents; education**

the second 2 years of schooling for children who had at least one parent with at least a college degree was 1.59 score points per month. This is the same rate of gain as that for the children whose parents had a high school degree or less. Children who had at least one parent with some college had a rate of gain during the second 2 years of 1.61 points per month. None of the contrasts between these three groups' rates of gain during the second 2 years were significant.

Table 6 summarizes the fixed-effects and empirical Bayes estimates of the rates of gain for the first 2 years of schooling and then the second 2 years of schooling. The empirical Bayes estimates that include school-level performance information show the gaps between the three parental groups increasing both before and after spring first grade.

Table 6

Fixed and Model Based Empirical Bayes Estimates of Rates of Gain by Parental Education Groups for Grades K-1 Versus Grades 1-3

	Slope K-1		Slope 1st-3rd	
Group	Fixed	EB	Fixed	EB
College degree	2.16	2.46	1.59	1.83
Some college	2.01	2.23	1.61	1.80
High school	1.81	1.99	1.58	1.71

Figure 8 shows the plots of the empirical Bayes estimates for the three groups of children. Clearly the plots of the empirical Bayes estimates approximate a fan spread effect (Campbell & Erlebacher, 1970), with the gaps increasing with age and schooling. While the rate of gain de-accelerates for all groups during the second 2 years of schooling, the gap between the three groups continues to increase even though they are gaining at quite different points on the test score scale. That is, the group of children with at least one parent with a college degree or higher are showing greater rates of gain even though they are learning more complex tasks. Less than half (46.0%) of the children from families with a high school degreed parent or lesser levels of education were making their gains in reading levels defined as beginning literacy (literal inference) or the two highest levels (extrapolation or evaluation). The corresponding percentages for children with parents with some college or a college degreed parent were 66.4 % and 84.1% respectively. It would seem that children of college-educated parents are typically progressing through much more complex material at a faster rate than children from a family with a parent having a high school degree or less and who are progressing through mostly preliteracy tasks. The increasing gaps between children from differing parent education groups shown here adds additional support for the hypothesis that language is a bigger factor in the learning of literacy tasks after first grade, and the home educational support system is likely to play a large role in the language development necessary for expanding reading comprehension skills.

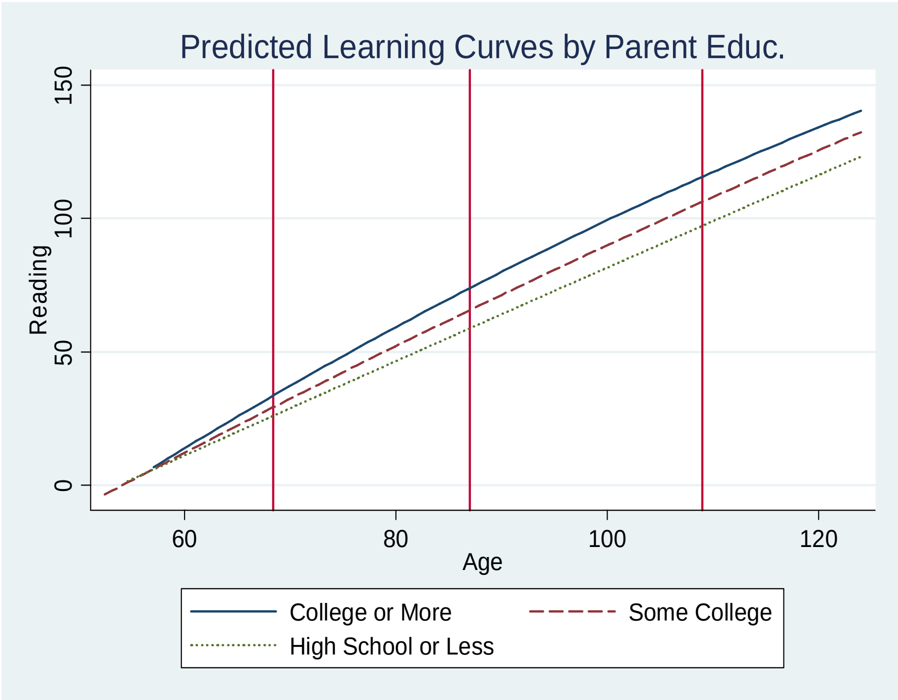

***Figure 8*. Predicted learning curves in reading by parents' education**

What does this say about the deficit versus lag controversy in children's reading development? It is the author's opinion that the de-acceleration observed here is a function of the increased emphasis on the language component in the reading task. Therefore, it would seem that children who lag will continue to lag unless there is a significant improvement in their language development. Under this hypothesis, one might expect English language learners to show some spurts in their reading skills as they improve their language skills. Those that don't will appear to fit a deficit model.

Summary and Conclusions

This manuscript addressed concerns about the potential for differential gains in reading taking place during the first 2 years of formal schooling (kindergarten-first grade) versus the next 2 years of schooling (first-third grade). A multilevel piecewise regression with a nodeat spring first grade was used in order to define separate regressions for the two time periods. Empirical Bayes estimates of individual growth curves were estimated, and smoothed estimates were plotted for selected demographic groups. The findings can be summarized as follows:

1. On average, there was a significant de-acceleration in the rates of gain when children went from primarily learning phonics-based tasks (first 2 years) to more language based reading for comprehension (second 2 years). Children who have made rapid gains during the first 2 years of schooling are more likely to be making gains at the upper end of the test score scale, in the neighborhood of the more difficult comprehension tasks during the second 2 years. This encounter with reading for understanding tasks and their potential requirement for new or more advanced language skills is likely contributing to their de-acceleration.

2. There were differential de-accelerations by racial/ethnic groups, with Asian children in particular showing proportionately greater de-accelerations during the first-to-third grade time period. The data suggest that this result may reflect the fact that Asian children were among those groups who had progressed at a relatively fast rate through the phonics-based material during the first 2 years, and therefore they had to make their gains during the second 2 years on the much more complex comprehension tasks requiring a larger language component.

3. Hispanic children showed proportionately less de-acceleration than did the Asian children. However, the Asian children were proportionately more likely than the Hispanic children to be making their gains at the upper levels of the test score scale in the neighborhood of the more difficult reading comprehension tasks during the second 2 years of schooling. These tasks at the upper end of the test score distribution are more language loaded.

4. Male and female growth curves differed only in the first 2 years, with girls gaining slightly more than boys on the more phonics-based material. There was no difference

in their growth rates during the second 2 years as both groups moved towards reading comprehension.

5. When growth curves were plotted separately for children differing in the level of parental education, a typical fan spread effect was noticed. That is, the predicted growth curves show increasing gaps between the groups as age and years of schooling increases. This increasing gap occurs even though children from college-educated families were proportionately much more likely to be making their gains at the upper end of the test score distribution, characterized by the more difficult comprehension tasks. It would seem that having a strong home educational support system helps to overcome the de-acceleration that may be partly the result of the point on the test score scale where the gains are taking place. Having educated parents is likely to encourage both formal and informal emphasis on early language development.

6. On the whole, the results suggest that if the schools put a greater emphasis on vocabulary development, the transition from phonics to reading for understanding might be a much smoother transition and would probably lead to less de-acceleration during the second 2 years of formal schooling. Children who are still making their gains in the lower level phonics-based material between first and third grade will likely show even greater de-acceleration in growth when their skills reach the point where they are beginning to do reading comprehension, unless their language capabilities are substantially increased.

References

Anderson, R. C, Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Champaign-Urbana, IL: Center for Study of Reading.

Campbell, D. T., & Erlebacher, A. E. (1970). How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Helmuth (Ed.), *Compensatory education: A national debate. Vol. 3. Disadvantaged child* (pp. 185–210). New York: Brunner-Mazel.

Cleary, T. A., Linn, R. L., & Rock, D. A. (1968). An exploratory study of programmed tests. *Educational and Psychological Measurement*, *28*, 345–360.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Fowler, A., et al. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 85*, 1–18.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: a longitudinal individual growth curves analysis. *Journal of Educational Psychology*, *88*, 3–17.

Goldstein, H. (1998). MLwiN [Computer software]. Retrieved September 25, 2007 from http://www.cmm.bristol.ac.uk/MLwiN

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

National Center for Education Statistics. (2004). *User's manual for the ECLS-K third grade. Public-use data file and electronic code book. Early Childhood Longitudinal Study, Kindergarten Class of 1998-99* (NCES 2004-001). Jessup, MD: ED Pubs.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington DC: U.S. Government Printing Office.

Rock, D. A., & Pollack, J. M. (2002). *A model based approach to measuring cognitive growth in pre-reading and reading skills during the kindergarten year*. (ETS Research Rep. No. RR-02-18). Princeton, NJ: ETS.

Seber, G. A. F., & Wild, C. J. (1989) *Nonlinear regression*. New York: John Wiley & Sons.

Snow, C. E., Burns, S., & Griffin, P. (Eds.). (1998). *National Research Council: Preventing reading difficulties in young children*. Washington DC: National Academy Press.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading difficulties: A regression based test of the phonological-core variable difference model. *Journal of Educational Psychology*, *86*, 24–53.

Stata Corporation. (2005). *Stata graphics reference manual, Release 9*. College Station, TX: Stata Press.

Wald, A. (1943). Tests of statistical hypothesis concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society, 54*, 426–482.