DOCUMENT RESUME

ED 465 804 TM 034 191

AUTHOR	Seltzer, Michael; Choi, Kilchan; Thum, Yeow Meng
TITLE	Examining Relationships between Where Students Start and How Rapidly They Progress: Implications for Constructing Indicators That Help Illuminate the Distribution of Achievement within Schools. CSE Technical Report.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.; National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	CSE-TR-560
PUB DATE	2002-04-00
NOTE	38p.
CONTRACT	R305B60002
AVAILABLE FROM	Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California at Los Angeles, 300 Charles E. Young Dr. North, Los Angeles, CA 90095-1522. Tel: 310-266-1532. For full text: http://www.cse.ucla.edu/CRESST/Reports/TECH559.PDF.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Achievement Gains; Change; Elementary Education; Elementary School Students; Longitudinal Studies; *Status
IDENTIFIERS	Time Series Analysis

ABSTRACT

Attending to school mean rates of change and to differences in rates of change for various demographic groups is of central importance in monitoring school performance. This paper makes the case for the need to expand this focus by also considering the relationship between where students in a school start (their initial status), and how rapidly they progress. In particular, the paper explores several ways in which attending to initial status in analyses of student progress can help draw attention to possible concerns regarding the distribution of achievement within schools, and, it is hoped, help stimulate discussion among teachers and administrators at given school sites regarding these concerns. Key points are illustrated by fitting a series of growth models to the time series data for students in several schools in the Longitudinal Study of American Youth sample. (Contains 7 figures, 6 tables, and 32 references.) (Author/SLD)

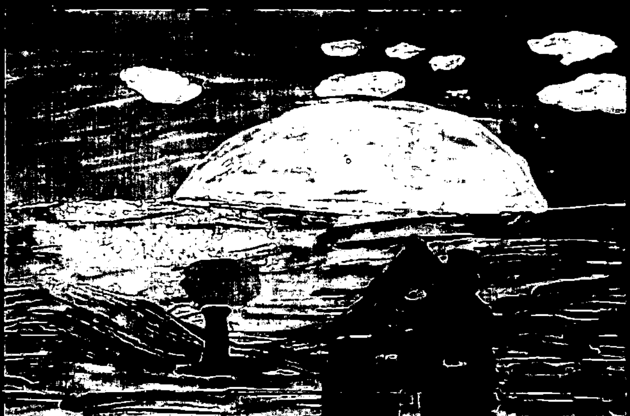
CRESST

National Center for Research on Evaluation, Standards, and Student Testing

Examining Relationships Between Where Students Start and How Rapidly They Progress: Implications for Constructing Indicators That Help Illuminate the Distribution of Achievement Within Schools

CSE Technical Report 560

Michael Seltzer, Kilchan Choi, and Yeow Meng Thum
CRESST/University of California, Los Angeles



UCLA Center for the Study of Evaluation

In Collaboration With:

University of Colorado at Boulder • Stanford University • The RAND Corporation
University of Southern California • Educational Testing Service
University of Pittsburgh • University of Cambridge

Examining Relationships Between Where Students Start and How Rapidly They Progress: Implications for Constructing Indicators That Help Illuminate the Distribution of Achievement Within Schools

CSE Technical Report 560

Michael Seltzer, Kilchan Choi, and Yeow Meng Thum
CRESST/University of California, Los Angeles

April 2002

Center for the Study of Evaluation
National Center for Research on Evaluation,
Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 3.3 Validity of Measures of Progress. Michael Seltzer, Project Director, CRESST/GSEIS/ University of California, Los Angeles.

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

Examining Relationships Between Where Students Start And How Rapidly They Progress: Implications For Conducting Analyses That Help Illuminate The Distribution of Achievement Within Schools

Michael Seltzer
CRESST/University of California, Los Angeles

Kilchan Choi
CRESST/University of California, Los Angeles

Yeow Meng Thum
CRESST/University of California, Los Angeles

Abstract

Attending to school mean rates of change and to differences in rates of change for various demographic groups is of central importance in monitoring school performance. In this paper, we argue for the need to expand this focus by also considering the relationship between where students in a school start (i.e., their initial status) and how rapidly they progress. In particular, we explore several ways in which attending to initial status in analyses of student progress can help draw attention to possible concerns regarding the distribution of achievement within schools, and, it is hoped, help stimulate discussion among teachers and administrators at given school sites regarding these concerns. To illustrate key points, we fit a series of growth models to the time series data for students in several schools in the Longitudinal Study of American Youth (LSAY) sample.

Introduction

A key approach to monitoring school or district performance involves using longitudinal data to study patterns of change in student achievement over a series of grades (see, for example, Willms, 1992, and Seltzer, Frank & Bryk, 1994). Thus, for example, interest might center on how rapidly students are progressing on average in particular curricular areas, or on the extent to which rates of progress differ across various demographic groups of students. As such, an important emphasis in the area of educational indicators entails computing and examining school

mean and district mean rates of change for cohorts of students across a series of grades.

Clearly much can be learned by moving beyond snapshots of student achievement at single points in time to analyses and summaries of student growth. To be sure, the notion of growth in knowledge and skills lies at the heart of definitions of learning and education (Bryk & Raudenbush, 1988; Willett, 1988, p. 346).

However, rather than focus exclusively on changes in student achievement over time (e.g., estimates of growth rates), we feel that it is also often important to consider levels of student achievement at the start of the time span one is studying (e.g., estimates of initial status). In this paper, we wish to explore several ways in which attending to the relationship between where students start and how rapidly they progress can help bring to light potentially important findings that might be masked if we limit our focus to computing school mean rates of change, or mean rates of change for various demographic groups. In particular, we will focus on how longitudinal analyses based on such an approach can help draw attention to possible concerns regarding the distribution of achievement within schools, and, it is hoped, help stimulate discussion among teachers and administrators at given school sites regarding these concerns.

To help illustrate various key points and ideas, we will fit a series of growth models to the time series data for students in several schools in the Longitudinal Study of American Youth (LSAY) sample (see Miller, Kimmel, Hoffer & Nelson, 1999). In these analyses we consider the value of examining correlation coefficients and regression slopes that capture how differences in initial status relate to subsequent differences in rates of change for the students in a school. Furthermore, in connection with comparing rates of change for various demographic groups, we emphasize the fact that comparisons that ignore differences in initial status can be highly misleading. For example, the size and direction of the difference in growth rates between girls and boys in a particular school may differ markedly, depending on whether we are focusing on girls and boys with relatively low initial status values, or on girls and boys with relatively high initial status values. As such, it becomes important to explore interactions between initial status and various demographic characteristics on rates of change. Throughout the analysis portions of our paper, we emphasize the use of graphical displays in helping to discern important patterns in the data, and to summarize and convey key results.

Moving Beyond Mean Rates of Change

Before turning to analyses of the LSAY data, we first wish to introduce certain key concepts. To help illustrate these concepts, suppose that we have measures of reading achievement across grades 2 through 5 for students in three different elementary schools. Suppose further that in the first school there is a positive relationship between where students start and how rapidly they progress. To help convey this pattern, Figure 1 displays the fitted reading achievement trajectories for 4 students. As can be seen, for the student with relatively low initial status, the slope of his trajectory is fairly flat (i.e., his rate of progress is very low). When we examine the set of 4 trajectories, we see that as initial status increases, the slopes of the trajectories increase (i.e., rates of change increase). Furthermore, we see that initial differences in student achievement (i.e., grade 2 levels of achievement) become magnified over time.

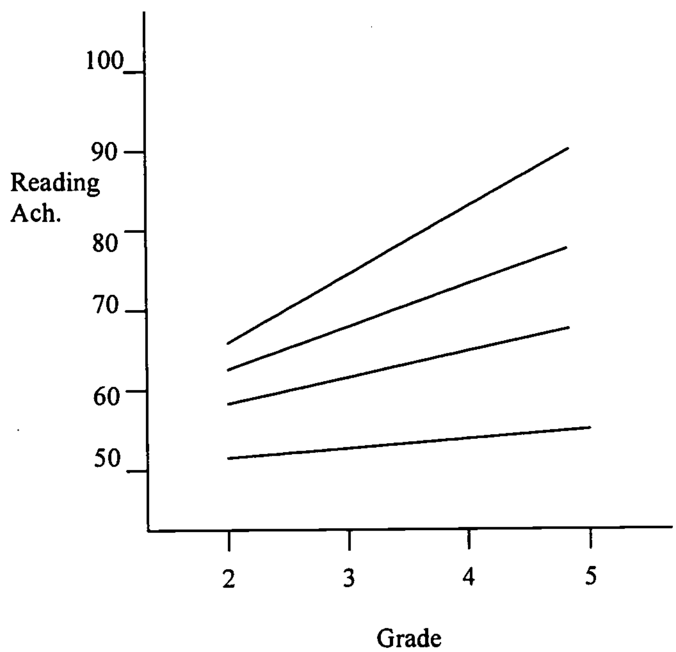


Figure 1. Positive Relationship between Initial Reading Achievement Levels and Rates of Change.

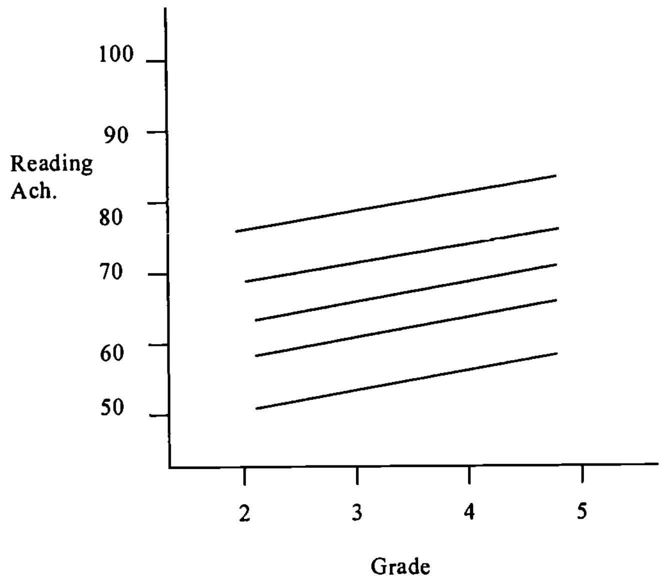


Figure 2. No Relationship between Initial Reading Achievement Levels and Rates of Change.

In school 2, initial status and rates of change are unrelated. Specifically, the four fitted trajectories displayed in Figure 2 reveal that students tend to progress at the same rate regardless of their initial status. Thus initial differences in achievement essentially hold steady over time.

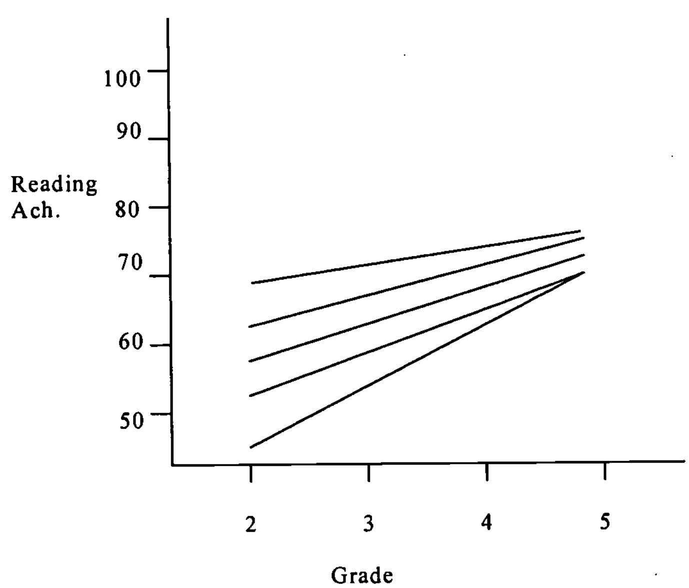


Figure 3. Negative Relationship between Initial Reading Achievement Levels and Rates of Change.

In some schools initial status and rates of change may be unrelated, but the underlying pattern of change may differ from that in school 2. For example, growth may be slow or rapid for students with low initial status, and slow or rapid for students with high initial status. Thus as initial status increases, the relationship between initial status and rates of change is non-systematic. This pattern will be illustrated in a later section of this article.

In school 3, the relationship between initial status and rates of change is negative. The four fitted trajectories in Figure 3 help convey that rates of change in this school are most rapid for those students with low initial status. As initial status increases, rates of progress decrease. Thus, in this particular school, initial differences in achievement tend to diminish over time.

An important implication is that while the average rates of change may be highly similar for two schools, the underlying relationship between initial status and rates of change within these schools may differ markedly. Thus, for example, in one school, initial differences in achievement may increase over time (Figure 1), while in the other, initial differences may decrease over time, i.e., those students who start low tend to catch up to those with relatively high initial levels of achievement (Figure 3).

An Illustrative Example Using Data from LSAY

To illustrate the above ideas, we now turn to analyses of the LSAY data. Note that the LSAY data set consists of over 50 cohorts of students in school districts throughout the U.S. Students in a given cohort attended the same middle school and then entered the same high school. In our paper, we focus on math achievement scores collected at the start of grades 7, 8, 9 and 10 for students in several different cohorts. Note that users of LSAY typically refer to a cohort of students in the sample as being nested within a particular school (e.g., school 308). We will use this terminology as well. But bear in mind that in general a given cohort was first nested within a particular middle school and then, subsequently, within a particular high school.

Growth modeling has become an increasingly popular tool for studying patterns of student change (see, e.g., Bryk & Raudenbush, 1987; Muthen & Khoo, 1998; Seltzer, Frank & Bryk, 1994). As will be seen, each of the growth models that we employ in this paper consists of two models: A level-1 or within-student model, and a level-2 or between-student

model. Within-student models enable us to capture key features of growth (e.g., initial status, rate of change) for each of the students in a sample. Between-student models enable us to estimate, for example, the mean rate of change for a group of students, assess the extent to which students vary in their rates of change, and identify important correlates of change.

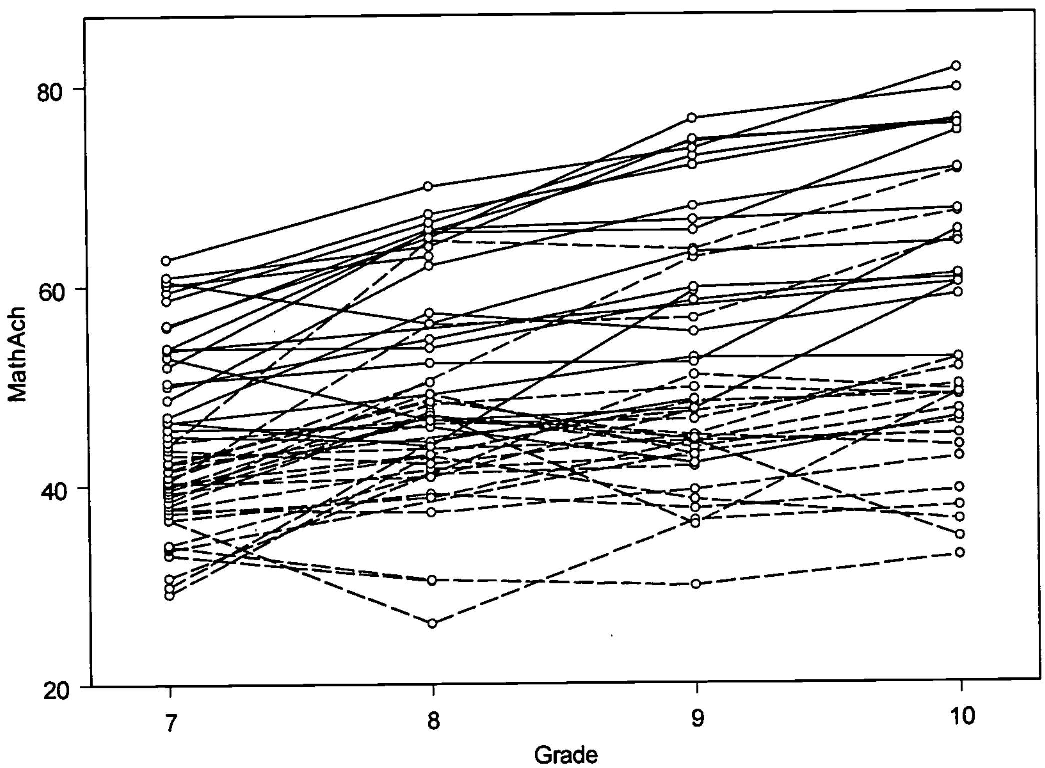


Figure 4. Observed math achievement trajectories for students in school 308. The dashed lines represent students whose grade 7 achievement scores lie below the mean initial status estimate for school 308 (44.72 points). The solid lines represent students whose grade 7 achievement scores lie above a value of 44.72.

We will first focus on estimating the mean rate of change and the correlation between initial status and rates of change for the sample of students in school 308. Figure 4 displays the series of math achievement scores for each of these students. As can be seen, the observed trajectories are roughly linear. In addition, we see that the trajectories for some students are fairly flat, while the trajectories for a number of other students are quite steep, indicating rapid rates of progress.

We now pose the following within-student (level-1) model for the time-series data for each of N=54 students in school 308:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(GRADE_{ti} - 7) + \varepsilon_{ti} \qquad \varepsilon_{ti} \sim N(0, \sigma^2), \qquad (1)$$

where Y_{ti} represents the math achievement score for student i at time t, and $GRADE_{ti}$ represents the grade that student i has entered at time t (i.e., *GRADE* takes on values ranging from 7-10). (Note that we are using a recent release of the LSAY data set, which contains achievement scale scores based on the application of IRT models discussed in Bock and Zimowski (1997).) The key parameters in this model are π_{0i} and π_{1i}. π_{1i} represents the growth rate (rate of change) for student i, and π_{0i} is an intercept. By virtue of centering *GRADE* around a value of 7, π_{0i} represents the expected math achievement score for student i at the start of grade 7 (i.e., initial status). (Note that centering provides a means of giving intercepts meaningful interpretations; see Endnote 1 and Bryk & Raudenbush, 1992, chpt. 6 for further details.) Finally, the ε_{ti} are residuals assumed normally distributed with mean 0 and variance σ^2.

What is the mean rate of change for students in school 308? Do rates of change for students in this school tend to vary systematically with initial status? To help address such questions, we specify the following between-student (level-2) model:

$$\begin{aligned} \pi_{0i} &= \beta_{00} + U_{0i} & U_{0i} &\sim N(0, \tau_{00}) \\ \pi_{1i} &= \beta_{10} + U_{1i} & U_{1i} &\sim N(0, \tau_{11}) \quad (Cov(U_{0i}, U_{1i}) = \tau_{01}), \ (2) \end{aligned}$$

where β_{00} and β_{11} represent, respectively, the means for initial status and rate of change for students in school 308. The U_{0i} and U_{1i} are level-2 residuals commonly termed random effects. U_{0i} captures the deviation of initial status for student i from the mean for initial status, and U_{1i} represents the deviation of the growth rate for person i from the mean growth rate. Note that the U_{0i} and U_{1i} are assumed normally distributed with variance τ_{00} and τ_{11}, respectively. Thus τ_{00} captures the amount of variation in initial status among students in school 308, and τ_{11} captures the amount of variation in growth rates. Furthermore, the covariance between initial status and rates of change for students in this school is captured by τ_{01} (i.e., $Cov(U_{0i}, U_{1i}) = \tau_{01}$). A large positive value for τ_{01} would imply that increases in initial status are accompanied by increases in rates

of change. A large negative value would indicate that as initial status increases, rates of change decrease.

All of the growth modeling results that we present in this paper were obtained using the software package WinBUGS (see Spiegelhalter, Thomas & Best, 2000). (Note that BUGS is a near acronym for 'Bayesian analysis using the Gibbs sampler'.) WinBUGS, which was developed by members of the MRC Biostatistics Unit in Cambridge, England, is a Windows-based program that enables one to fit a wide array of models, some of which may be difficult to estimate using conventional statistical tools. As discussed below, a variety of other software options are available for fitting the models presented in our paper. (A discussion of various implementation details regarding our use of WinBUGS, and of the estimation approach upon which WinBUGS is based, can be found in Appendix A.)

With respect to the mean initial status for students in school 308, we obtain an estimate of approximately 45 points (see Table 1). We also see that the resulting estimate for the mean rate of change is 3.52, which implies that student math achievement scores are, on average, increasing approximately 3.5 points per grade.

Next, note that the estimated covariance between initial status and rates of change is positive (i.e., 9.43), and that the 95% interval for this estimate contains only positive values. This provides strong evidence of a positive relationship between initial status and rates of change.

As Hays notes (1988, p. 555), a correlation coefficient is the standardized covariance between two variables. It can be computed by dividing the covariance between two variables (e.g., *Cov*(A, B)) by the square root of the product of their variances (e.g., $(Var(A) \times Var(B))^{1/2}$). Thus, the correlation between initial status and rates of change can be expressed as a function of the covariance between initial status and rates of change (τ_{10}), the variance in initial status (τ_{00}), and the variance in rates of change (τ_{11}): $\rho = \tau_{10} / (\tau_{00} \times \tau_{11})^{1/2}$. As can be seen in Table 1, we obtain an estimate for ρ of .50. Furthermore, we see that the 95% interval for ρ contains only positive values.

A positive relationship between where students start and how rapidly they progress can be discerned in Figure 4. Clearly the relationship is far from a perfect one. For example, some students with relatively high grade 7 achievement scores progress fairly slowly over time. In addition, several students with low grade 7 achievement scores progress quite rapidly. However, in general, we see that among students with relatively low grade 7 scores, many have trajectories that are fairly

flat, while among students who start out relatively high, many have trajectories that are substantially steeper. In connection with this, we see a fanning out of achievement scores over time.

Table 1: Comparing patterns of change for students in schools 308 and 143.[2]

	School 308 (N=54)		School 143 (N=54)	
	Estimate	95% Interval	Estimate	95% Interval
Fixed Effects				
Mean Initial Status (β_{00})	44.72	(42.28, 47.14)	44.60	(42.26, 46.92)
Mean Rate of Change (β_{10})	3.52	(2.69, 4.35)	3.33	(2.27, 4.40)
Variance Components				
Within-Person Error (σ^2)	13.07	(9.99, 17.38)	17.24	(12.58, 24.16)
Random Effects Variance for Initial Status (τ_{00})	70.04	(46.57, 108.30)	60.78	(38.87, 96.78)
Random Effects Variance for Rates of Change (τ_{11})	5.36	(3.02, 9.33)	7.23	(3.54, 14.26)
Random Effects Covariance (τ_{01})	9.43	(2.96, 17.68)	0.70	(-9.09, 9.07)
Corr. between Initial Status and Rates of Change ($\rho = \tau_{01} / \sqrt{\tau_{00} \times \tau_{11}}$)	0.50	(0.15, 0.74)	0.03	(-0.35, 0.44)

We now fit the model defined by Equations 1 and 2 to the data for students in school 143. In Table 1, we see that the estimate of the mean growth rate for students in this school is quite similar to the estimated mean rate for students in school 308. The mean initial status estimates for these schools are nearly identical.

An important difference in results is that while the estimate of the covariance between initial status and rates of change for school 143 is positive, it is substantially smaller than the covariance estimate for school 308, and the corresponding 95% interval comfortably includes a value of 0. Similarly, the estimated correlation between initial status and rates of change is extremely small (0.03), and the corresponding 95% interval contains a value of 0.

These results are consistent with an examination of the set of observed trajectories for the sample of students in school 143 (see Figure 5). As can be seen, there is extensive crisscrossing of trajectories. When we consider students with relatively low grade 7 scores and students with relatively high grade 7 scores, we see that for both groups, rates of change are rapid for some students but slow for others. Student progress is, in some sense, more ``fluid'' in school 143 than in school 308. By this, we mean that over time, an appreciable number of students with lower than average grade 7 scores catch up to or surpass students with higher than average grade 7 scores.

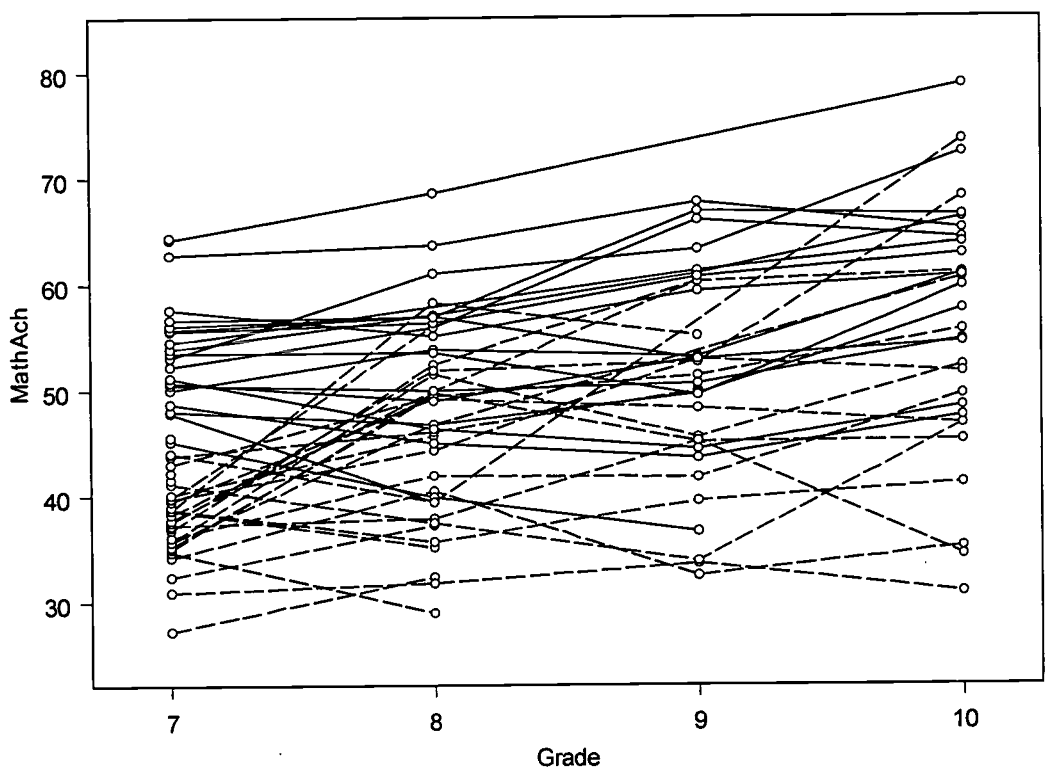


Figure 5. Observed math achievement trajectories for students in school 143. The dashed lines represent students whose grade 7 achievement scores lie below the mean initial status estimate for school 143 (44.60 points). The solid lines represent students whose grade 7 achievement scores lie above a value of 44.60.

14

Regressing Rates of Change on Initial Status

Correlation coefficients provide us with measures of the strength of linear association between two variables (e.g., variables A and B). In addition to information of this kind, we are also often interested in information concerning the expected amount of change in one variable (e.g., variable A) when the other increases 1 unit. This information would be captured by the slope relating changes in variable B to variable A, and estimating this slope would entail regressing variable A on B.

Analogously, we might ask: For students in school 308, how much of a change in rate of growth (π_{1i}) do we expect when initial status (π_{0i}) increases 1 unit? Addressing this question implies employing initial status as a predictor of rate of change. Thus we expand our between-student model as follows:

$$\begin{aligned} \pi_{0i} &= \beta_{00} + U_{0i} & U_{0i} &\sim N(0, \tau_{00}) \\ \pi_{1i} &= \beta_{10} + b(\pi_{0i} - \beta_{00}) + U_{1i} & U_{1i} &\sim N(0, \tau_{11}). \end{aligned} \tag{3}$$

A key parameter in this model is b, which captures the amount of change that we expect in π_{1i} when π_{0i} increases 1 unit. Note that regressing one parameter (π_{1i}) on another (π_{0i}) is termed a latent variable regression, and that coefficients such as b are termed latent variable regression coefficients. In addition to WinBUGS, a variety of other software packages can be used to fit such models, including HLM5 (Raudenbush, Bryk, Cheong & Congdon, 2000) and Mplus (Muthen & Muthen, 1998). We discuss various software possibilities in the final section of our paper.

Before moving on, we wish to comment briefly on several other features of the above between-student model. First, as in the case of the previous between-student model (Equation 2), β_{00} represents school mean initial status. Secondly, note that in the model for rates of change, we have centered π_{0i} around β_{00}. This helps give β_{10} a more useful interpretation. As we will see below, β_{10} represents the expected growth rate for a student whose initial status value is equal to the mean initial status value for school 308. Finally, in contrast to the between-student model defined in Equation 2, τ_{11} now represents the amount of variation in rates of change that remains after we take into account differences in initial status (see Endnote 3).

Table 2: Comparing initial status/rate of change slopes for schools 308 and 143.

	School 308 (N=54)		School 143 (N=54)	
	Estimate	95% Interval	Estimate	95% Interval
Fixed Effects				
Mean Initial Status (β_{00})	44.73	(42.34, 47.15)	44.62	(42.28, 46.94)
Mean Rate of Change (β_{10})	3.49	(2.66, 4.32)	3.32	(2.23, 4.44)
Initial status / Rate of Change Slope (b)	.174	(0.067, 0.290)	0.010	(-0.147, 0.179)
Variance Components				
Within-Person Error (σ^2)	13.44	(10.23, 17.87)	17.07	(12.42, 24.04)
Random Effects Variance for Initial Status (τ_{00})	67.27	(44.61, 104.40)	60.54	(38.95, 95.91)
Random Effects Variance for Rates of Change (τ_{11})	3.06	(1.22, 6.65)	7.69	(3.77, 14.90)

When we fit the growth model defined by Equations 1 and 3 to the data for school 308, we obtain an estimate of 0.174 for b (see Table 2). As can be seen, the corresponding 95% interval contains only positive values. The estimates for b, β_{01} and β_{00} in Table 2 can be used to compute expected rates of change for students with various initial status values. Thus we have the following equation:

$$E(\pi_{1i}) = 3.49 + .174\,(\pi_{0i} - 44.73), \tag{4}$$

where $E(\pi_{1i})$ denotes an expected rate of change. Note that for students with initial status values equal to the mean initial status value (i.e., $(\pi_{0i} - 44.73) = 0$), the expected rate of change is: 3.49 + .174(0) = 3.49. However, for students with initial status values 10 points above the mean initial status value (i.e., $(\pi_{0i} - 44.73) = 10$), the expected rate of change is appreciably faster: 3.49 + .174(10) = 5.23. Conversely, for students with initial status values 10 points below the mean initial status value (i.e., $(\pi_{0i} - 44.73) = -10$), the expected rate of change is substantially slower: 3.49 + .174(-10) = 1.75 (see Figure 6).

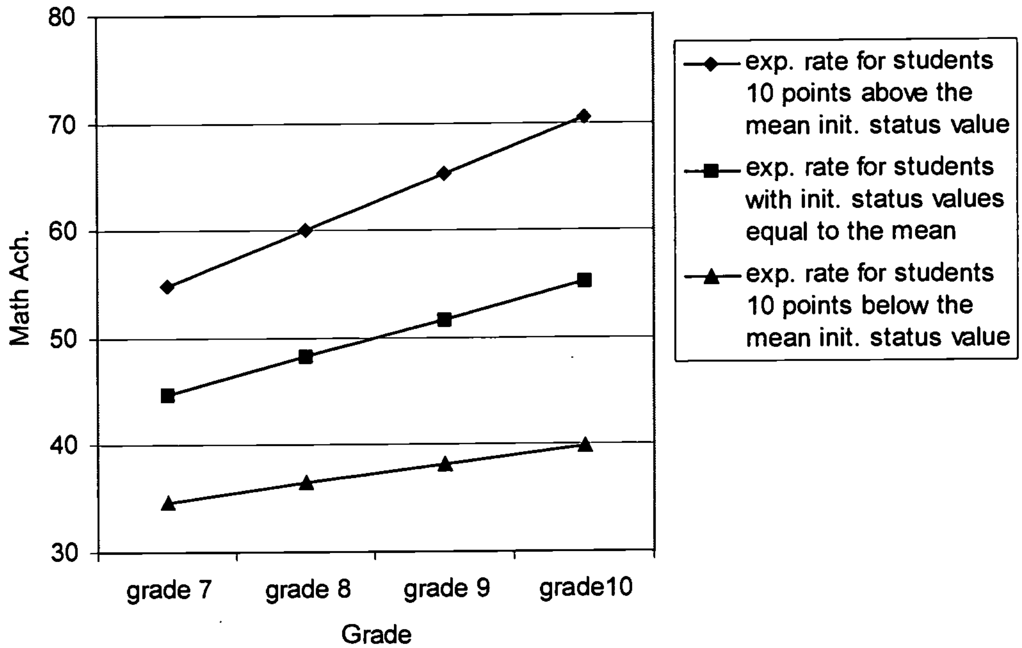


Figure 6. Expected growth trajectories for students in school 306 based on the results in Table 2.

In contrast, for school 143, we obtain an estimate for b of 0.010, and it can be seen that the resulting 95% interval comfortably includes a value of 0 (see Table 2). For students with initial status values equal to the mean initial status value for school 143 (i.e., $(\pi_{0i} - 44.62) = 0$), the expected growth rate is: $3.32 + .010(0) = 3.32$. For students with initial status values 10 points above the mean initial status value (i.e., $(\pi_{0i} - 44.62) = 10$), note that the expected growth rate is only slightly faster than the mean rate: $3.32 + .010(10) = 3.42$. Likewise, the expected growth rate for students with initial status values 10 points below the mean initial status value is extremely similar to the mean rate: $3.32 + .010(-10) = 3.22$.

As can be seen, analyses of this kind help draw our attention to possible concerns regarding the distribution of achievement within a school. For example, in a school such as 308, though students are, on average, progressing at a rate of approximately 3.5 points per year, we see that rates of rates of progress among students with relatively low initial levels of achievement are rather minimal. In such settings, the gap in

achievement between students who start low and those who start high can widen substantially over time. In bringing to light patterns of this kind, such analyses can help stimulate discussion among a school's teachers and administrators regarding important areas of concern and possible courses of action (e.g., programmatic changes that could help promote more rapid rates of change among students with low initial levels of achievement). We discuss the use of various summaries and analyses involving the relationship between initial status and rates of change in more detail in the final section of our paper.

Comparing Mean Rates of Change for Different Demographic Groups

We now turn to the issue of comparing rates of progress for different demographic groups of students. For illustrative purposes, we focus on differences in rates of change between girls and boys in school 142 in the LSAY sample. For these analyses, we will work with the same level-1 model as in the above sections (see Equation 1). At level 2, we pose the following model:

$$\pi_{0i} = \beta_{00} + \beta_{01}\, GENDER_i + U_{0i} \quad U_{0i} \sim N(0, \tau_{00})$$
$$\pi_{1i} = \beta_{10} + \beta_{11}\, GENDER_i + U_{1i} \quad U_{1i} \sim N(0, \tau_{11})\ (Cov(U_{0i}, U_{1i}) = \tau_{01}), \quad (5)$$

where $GENDER_i$ takes on a value of 0 if student i is a boy and a value of 1 if student i is a girl. By virtue of this coding scheme, β_{00} represents the expected initial status for boys in school 142, and β_{01} captures the expected difference in initial status between girls and boys. Similarly, β_{10} represents the expected rate of change in math achievement for boys, and β_{11} captures the expected difference in rates of change between girls and boys. In this model, τ_{00} and τ_{11} represent, respectively, the variance in initial status and the variance in growth rates within each demographic group (i.e., male students and female students). Furthermore, $Cov(U_{0i}, U_{1i}) = \tau_{01}$ now represents the covariance between initial status and rates of change within each group.

The resulting estimate for β_{01} indicates that initial status is, on average, approximately 2.8 points higher for the girls in our sample, though the lower boundary of the 95% interval for β_{01} includes a value of 0 (see Table 3). In contrast, the estimate for β_{11} (-1.63) suggests that rates of change for girls are, on average, appreciably lower than rates of change for boys. As can be seen, the 95% interval for β_{11} excludes a value of 0, though just barely so. Thus, while the expected rate of change for boys is

5.22 points per year, the expected rate for girls is 5.22 − 1.63 = 3.59 points per year.

Table 3: Comparing rates of change for girls and boys in school 142 (24 girls, 36 boys).

	Estimates	95% Interval
Fixed Effects		
Model for Initial Status (π_{0i}):		
Boys (β_{00})	50.78	(47.33, 54.21)
Girls/Boys Contrast (β_{01})	2.82	(-2.71, 8.27)
Model for Rates of Change (π_{1i}):		
Boys (β_{10})	5.22	(4.18, 6.27)
Girls/Boys Contrast (β_{11})	-1.63	(-3.27, -0.01)
Variance Components		
Within-Person Error (σ^2)	18.69	(14.55, 24.27)
Random Effects Variance for Initial Status (τ_{00})	94.58	(64.00, 144.10)
Random Effects Variance for Rates of Change (τ_{11})	5.10	(2.72, 9.18)
Covariance between Initial Status and Rates of Change (τ_{01})	11.06	(3.86, 20.22)
Correlation between Initial Status and Rates of Change ($\rho = \tau_{01} / \sqrt{\tau_{00} \times \tau_{11}}$)	0.52	(0.18, 0.76)

It is important to note that our comparison of growth rates for girls and boys does not take into account the fact that math achievement at the start of grade 7 is somewhat higher for the girls in our sample. If differences in initial status were inconsequential in terms of how rapidly students progress, this would not be a concern. However, the results that we obtain for τ_{01} and the correlation coefficient (ρ) point to a fairly strong

positive relationship between initial status and rates of change for boys and for girls. Note, in particular, that the resulting estimate for ρ is 0.52.

Analogous to using ANCOVA models to compare groups of interest adjusting for initial pretest differences, we now attempt to obtain an estimate of the difference in growth rates between girls and boys adjusting for differences in initial status. To accomplish this, we pose the following between-student model:

$$\begin{aligned} \pi_{0i} &= \beta_{00} + \beta_{01}(GENDER_i - GENDER_{.}) + U_{0i} \quad & U_{0i} \sim N(0, \tau_{00}) \\ \pi_{1i} &= \beta_{10} + \beta_{11}\,GENDER_i + b(\pi_{0i} - \beta_{00}) + U_{1i} \quad & U_{1i} \sim N(0, \tau_{11}). \quad (6) \end{aligned}$$

In contrast to Equation 5, π_{0i} appears as a covariate in our model for growth rates. The parameter of primary interest in this model is β_{11}, which represents an expected difference in growth rates between girls and boys that is adjusted for differences in initial status. Put differently, β_{11} represents an expected difference in growth rates *holding constant initial status*. The parameter b is a regression coefficient that relates differences in initial status to rates of change. Note that this model essentially assumes that the slope relating initial status to rates of change is equivalent for girls and boys. As will be seen in a later section of our paper, this assumption appears to be extremely reasonable in the case of this school. (Note that the centerings employed in the above between-student model are intended to give the parameters β_{00} and β_{10} more useful interpretations. See Endnote 4 for details. Also see Endnote 4 for a discussion of certain statistical advantages connected with employing initial status as a covariate.)

In Table 4, we see that the resulting estimate for b is 0.153 and that the corresponding 95% interval contains only positive values. This suggests a strong, positive relationship between initial status and subsequent rates of progress for boys and for girls. When we take into account the fact that boys, on average, have lower initial status values than girls, we see that the resulting expected difference in growth rates between girls and boys is appreciably larger than the unadjusted difference (i.e., -2.06 versus -1.63), and that the upper boundary of the resulting 95% interval lies well below a value of 0. This result implies that for girls and boys with similar levels of achievement at the start of grade 7 (i.e., holding constant initial status), the expected rate of growth in achievement for boys is approximately 2 points per year faster than the expected rate for girls. This difference in rates translates into a difference

in expected achievement scores after 3 years of schooling (i.e., at the start of grade 10) of approximately 6 points.

Table 4: Comparing rates of change for girls and boys in school 142 adjusting for differences in initial status (24 girls, 36 boys).

	Estimate	95% Interval
Fixed Effects		
Model for Initial Status(π_{0i}):		
Mean Initial Status(β_{00})	51.91	(49.28, 54.54)
Girls/Boys Contrast(β_{01})	2.80	(-2.57, 8.12)
Model for Rates of Change(π_{1i}):		
Boys(β_{10})	5.37	(4.37, 6.41)
Girls/Boys Contrast(β_{11})	-2.06	(-3.67, -0.45)
Initial Status / Rate of Change Slope (b)	0.153	(0.066, 0.249)
Variance Components		
Within-Person Error(σ^2)	19.39	(15.07, 25.15)
Random Effects Variance for Initial Status(τ_{00})	90.27	(60.84, 137.80)
Random Effects Variance for Rates of Change(τ_{11})	2.37	(0.83, 6.07)

Thus, as in the case of the analyses in the previous section, our attention is again drawn to issues concerning the distribution of achievement within schools. What processes or factors likely underlie this pattern of results for girls and boys? What might be done to try to promote more rapid rates of progress among girls in this school? These are some of the questions that the analyses in this section encourage us to consider. Before moving on to the next section, we wish to point out that the samples of girls and boys in school 142 are extremely similar in terms of such potentially important intake characteristics as home resources and educational aspirations. Thus the differences in rates of change that we see cannot be accounted for by these factors.

Overall Comparisons Can Be Misleading: Examining Interactions Between Initial Status and Demographic Characteristics

While comparing rates of change for various demographic groups of interest can be extremely useful, such comparisons can potentially be misleading, even when we have taken into account differences in initial status. To help illustrate this point, we will focus on patterns of growth for students in school 302 in the LSAY sample.

Table 5: Comparing rates of change for girls and boys in school 302 adjusting for differences in initial status (38 girls, 41 boys).

	Estimate	95% Interval
Fixed Effects		
Model for Initial Status (π_{0i}):		
Mean Initial Status (β_{00})	52.54	(50.42, 54.67)
Girls/Boys Contrast (β_{01})	-0.62	(-4.88, 3.66)
Model for Rates of Change (π_{1i}):		
Boys (β_{10})	4.28	(3.65, 4.91)
Girls/Boys Contrast (β_{11})	0.20	(-0.73, 1.12)
Initial Status / Rate of Change Slope (b)	0.100	(0.048, 0.157)
Variance Components		
Within-Person Error (σ^2)	13.20	(10.88, 16.07)
Random Effects Variance for Initial Status (τ_{00})	82.34	(59.05, 117.50)
Random Effects Variance for Rates of Change (τ_{11})	0.72	(0.24, 1.98)

Initial status values and rates of change differ very little for girls and boys in this school. As can be seen in Table 5, initial status values for girls are, on average, approximately .60 points lower for girls than boys, and rates of change, holding constant initial status, are slightly faster for

girls (0.20). (Note that the unadjusted difference in rates slightly favors girls as well.)

In the case of the analysis presented in Table 5, a key assumption is that the slope relating initial status to rates of change (b) is equivalent for girls and boys. This is analogous to the assumption of parallel within-group slopes in classic ANCOVA analyses. An implication of this assumption is that the expected difference in rates of change between girls and boys is 0.20 regardless of whether we are considering boys and girls with relatively low initial achievement values, or whether we are considering boys and girls with high initial values. As will become clear, this assumption is extremely reasonable in the case of school 142, but highly questionable in the case of school 302.

We now pose a model that allows for the possibility that the slope relating initial status to rates of change may differ for girls and boys. That is, we include an interaction term (i.e., $GENDER_i \times (\pi_{0i} - \beta_{00})$) in our model for rates of change:

$$\begin{aligned}
\pi_{0i} &= \beta_{00} + \beta_{01}(GENDER_i - GENDER.) + U_{0i} \quad U_{0i} \sim N(0, \tau_{00}) \\
\pi_{1i} &= \beta_{10} + \beta_{11}\, GENDER_i + b_1(\pi_{0i} - \beta_{00}) + b_2 (GENDER_i \times (\pi_{0i} - \beta_{00})) \\
&\quad + U_{1i} \quad For\ Boys: U_{1i} \sim N(0, \tau_{11B})\ ;\ For\ Girls: U_{1i} \sim N(0, \tau_{11G}) \quad (7)
\end{aligned}$$

where β_{11} and b_1 represent, respectively, the main effects of gender and initial status on rates of change, and where b_2 captures the interaction between initial status and gender. Note also that τ_{11G} represents the variance in growth rates for girls that remains after taking into account initial status, and τ_{11B} represents the remaining variance in growth rates for boys.

Let's now unpack the above between-student model. Based on the equation for growth rates in this model, the expected rate of change for boys ($GENDER_i = 0$) is:

$$E(\pi_{1i} \mid GENDER_i = 0) = \beta_{10} + b_1(\pi_{0i} - \beta_{00}) \qquad (8)$$

and the expected rate of change for girls ($GENDER_I = 1$) is:

$$E(\pi_{1i} \mid GENDER_i = 1) = (\beta_{10} + \beta_{11}) + (b_1 + b_2)(\pi_{0i} - \beta_{00}) \qquad (9)$$

As can be seen in Equation 8, b_1 is the slope capturing the relationship between initial status and rates of change for boys. In Equation 9, we see that the initial status / rate of change slope for girls is equal to $b_1 + b_2$.

Thus b_2 captures the difference between the initial status / rate of change slopes for girls and boys.

Table 6: Examining the interaction between initial status and gender on rates of change in school 302 (38 girls, 41 boys).

	Estimate	95% Interval
Fixed Effects		
Model for Initial Status(π_{0i}):		
Mean Initial Stuatus(β_{00})	52.54	(50.41, 54.66)
Girls/Boys Contrast(β_{01})	-0.61	(-4.89, 3.70)
Model for Rates of Change(π_{1i}):		
Boys(β_{10})	4.26	(3.59, 4.93)
Girls/Boys Contrast(β_{11})	0.20	(-0.76, 1.15)
Initial Status / Rate of Change Slope (b_1)	0.142	(0.080, 0.214)
Interaction between Gender and Initial Status (b_2)	-0.121	(-0.226, -0.018)
Variance Components		
Within-Person Error(σ^2)	13.12	(10.86, 15.96)
Random Effects Variance for Initial Status(τ_{00})	82.38	(59.15, 117.10)
Random Effects Variance for Rates of Change		
Boys (τ_{11B})	0.37	(0.11, 1.72)
Girls(τ_{11G})	1.23	(0.49, 3.02)

In Table 6, we see that the resulting estimate of b_1 is 0.142, and that the lower boundary of the corresponding 95% interval is well above a value of 0. This suggests a strong positive relationship between initial status and rates of change for boys. That is, for boys, differences in initial status appear to be very consequential with respect to subsequent rates of change. In contrast, we obtain a negative estimate for b_2 (i.e., -0.121). It

can also be seen that the corresponding 95% interval for b_2 contains only negative values. This suggests that the initial status / rate of change slopes for girls and boys differ substantially. In particular, summing the point estimates for b_1 and b_2 we obtain a value of 0.021. Thus for girls, differences in initial status appear to be inconsequential with respect to subsequent rates of change. (Note that in fitting the above interaction model to the data for students in school 142, we obtain a point estimate for b_2 that is extremely close to 0. This suggests that in school 142, initial status / rate of change slopes appear to be similar for girls and boys.)

Analogous to ANCOVA analyses, such differences in initial status / rate of change slopes have important implications for drawing conclusions concerning expected differences in rates of change between girls and boys. To help grasp the results of this analysis, we consider the expected rates of change for girls and boys whose initial status values are 12 points below the mean initial status estimate for school 302 $((\pi_{0i} - 52.54) = -12)$, whose initial status values are equal to the mean initial status estimate $((\pi_{0i} - 52.54) = 0)$, and whose initial status values are 12 points above the mean initial status estimate $((\pi_{0i} - 52.54) = 12)$.

We first substitute the estimates for β_{10}, b_1 and β_{00} into Equation 8:

$$E(\pi_{1i} \mid GENDER_i = 0) = 4.26 + 0.142\,(\pi_{0i} - 52.54). \tag{10}$$

We then substitute the estimates for β_{10}, β_{11}, b_1, b_2 and β_{00} into Equation 9:

$$E(\pi_{1i} \mid GENDER_i = 1) = (4.26 + .20) + (0.142 - 0.121)\,(\pi_{0i} - 52.54). \tag{11}$$

Equation 11 simplifies as follows:

$$E(\pi_{1i} \mid GENDER_i = 1) = 4.46 + 0.021\,(\pi_{0i} - 52.54) \tag{12}$$

Based on Equation 10, the expected rate for a boy whose initial status value is 12 points below the mean initial status estimate is: $4.26 + 0.142\,(-12) = 4.26 - 1.70 = 2.56$. For a girl whose initial status value is 12 points below the mean initial status estimate, the expected rate based on Equation 12 is: $4.46 + 0.021\,(-12) = 4.46 - 0.25 = 4.21$. Thus among students with initial status values that are 12 points below the grand mean, the expected rate of growth is appreciably higher for girls than boys.

Based on Equations 10 and 12, the expected rate of change for a boy whose initial status value is equal to the mean initial status estimate is

((π_{0i} − 52.54) = 0) is 4.26, while for girls, the expected rate is 4.46. In contrast to students whose initial status values are 12 points below the mean initial status estimate, we see that for students with initial status values equal to 52.54, the expected rates of change for boys and girls are quite similar.

Turning to students whose initial status values are 12 points above the mean initial status estimate, the expected rate of change for boys is 4.26 + 0.142 (12) = 4.26 + 1.70 = 5.96. For girls, the expected rate is:
4.46 + 0.021 (12) = 4.46 + 0.25 = 4.71. Thus for students with initial status values that are 12 points above the mean initial status estimate, the expected rate of change is markedly higher for boys than girls.

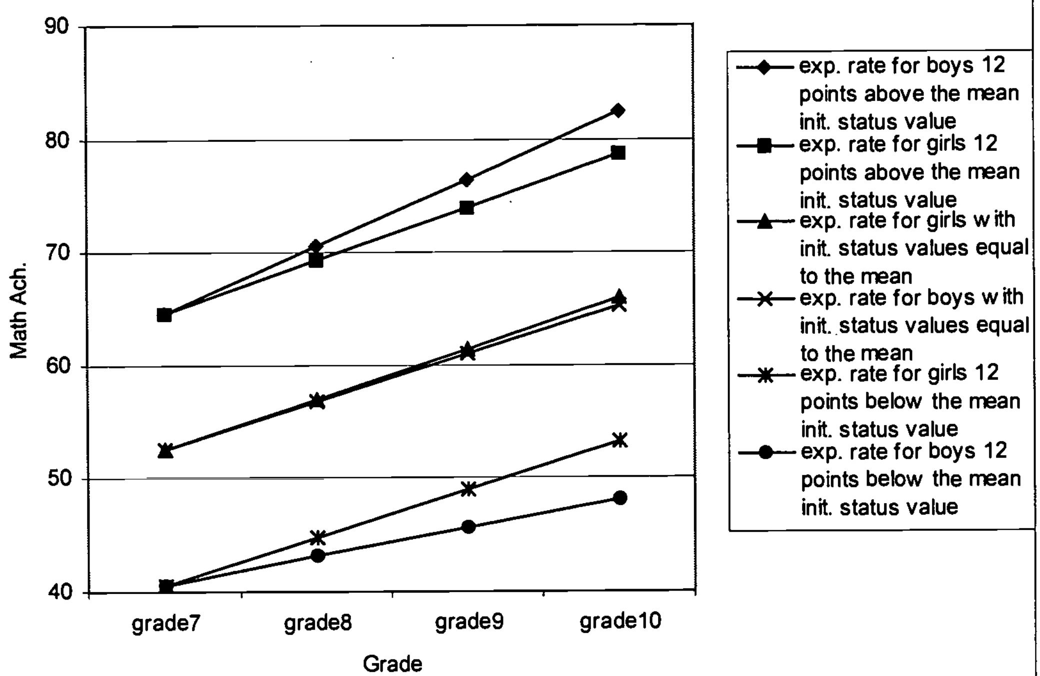


Figure 7. Expected growth trajectories for girls and boys in school 302 based on the initial status x gender interaction analysis. The mean initial status for school 302 is 52.54

Note that for the three values of initial status that we are considering, as initial status increases, the expected rate of change for girls increases somewhat from a value of 4.21 to a value of 4.71. This slight increase is connected to the fact that the estimate of the initial status / rate

of change slope for girls takes on a small positive value (i.e., 0.021). For boys, however, we see that as initial status increases, the expected rate of change for boys increases substantially from a value of 2.56 to a value of 5.96. These expected trajectories are displayed in Figure 7.

An implication of the above analyses is that estimates of overall differences in rates of change for different demographic groups may be misleading. In particular, they may mask the fact that the size and direction of the difference may vary markedly depending upon whether we are focusing on students with relatively low or high initial status values. (Note that analyses of this kind prove to be extremely valuable in longitudinal studies of interventions. In particular, one can investigate interactions between initial status and type of treatment on rates of change. See Endnote 5.)

Discussion

Attending to mean rates of change, and overall differences in rates of change for various demographic groups is of central importance in monitoring the performance of schools and districts. In this paper, we have argued for the need to expand this focus by also considering the relationship between initial status and rates of change. In particular, we have focused on various summaries and analyses that can help draw our attention to possible concerns regarding the distribution of achievement within a school (e.g., why does initial status appear to be very consequential with respect to subsequent rates of progress?; among students with high initial status values, why are rates of progress appreciably more rapid for boys than girls?).

We saw that estimates of correlations between initial status and rates of change, and estimates of initial status / rate of change slopes, provide us with two key measures of within-school relationships between initial status and rates of change. We also focused on the potential value of comparing growth rates for various demographic groups of interest adjusting for differences in initial status. Note that we feel it is important to consider both unadjusted and adjusted differences in rates. Each of these measures provides us with useful information. Regardless of whether one group of students differs from another in terms of initial status, it is important to know whether these groups tend to be progressing at fairly similar rates or not. Unadjusted differences provide us with information of this kind. However, if initial status in a school of interest appears to be very consequential with respect to subsequent rates

of achievement, and if the groups that we wish to compare differ appreciably in terms of initial status, then it is also valuable to ask: What is the expected difference in rates of change holding constant initial status?

An extremely important point, however, is that comparisons of growth rates can be misleading regardless of whether we adjust for differences in initial status or not. Specifically, as we saw above, the size and direction of expected differences in rates of change between demographic groups of interest can vary substantially across various initial levels of achievement. In the case of our example, we saw that among students with relatively low initial status, rates of change were appreciably faster for girls than boys. However, among students with relatively high initial status, rates of progress tended to be more rapid for boys than girls.

Our hope is that the kinds of summaries and analyses illustrated above can, in conjunction with various measures of school performance (see, e.g., Bryk, Thum, Easton & Luppescu, 1998; Sanders & Horn, 1994; Thum, 2001; Willms, 1992), help stimulate fruitful discussion among school personnel regarding possible areas of concern. For example, suppose that there is evidence of a strong positive relationship between initial status and rates of change in a particular school. Drawing on their extensive experience and contextual knowledge, it would be extremely valuable for the teachers and administrators in this school to discuss the factors likely underlying this relationship. For example, to what extent might this be due to the school's policies regarding tracking? To what extent might this be due to policies regarding the number and kinds of mathematics courses that students are required to take? In these discussions, it would be important to consider those factors that are likely contributing to differences in achievement among students observed at the start of the series of grades under consideration (e.g., the nature and extent of previous academic difficulties; differences in the quality of prior instruction received by students). The particular factors underlying the strong relationship between initial status and rates of change in this school may be quite clear, or it may be the case that further study and discussion are necessary.

Note that with the help of research personnel, various hunches could be assessed more formally through additional growth modeling analyses if the necessary data are available. To continue with the above example, suppose we wish to explore whether the strong positive relationship between initial status and rates of change for students in this

school (i.e., a large, positive estimate for b) is in large part due to differences in the quality of pre-algebra instruction received by students. If this is so, fitting a model in which rates of change are modeled as a function of initial status *and* measures of this explanatory factor would result in a substantially smaller estimate for b.

Employing initial status as a predictor of rates of change can also help broaden the kinds of questions that we are able to address in studies of school effects based on analyses of large-scale longitudinal data sets (e.g., NELS, LSAY). Specifically, through the use of three-level growth models, we can systematically model differences in the relationship between initial status and rates of change across schools. At level 1, as above, we model the time-series observations for each student as a function of grade. At level 2, we model rates of change as a function of initial status for each of the schools in a sample. This might be termed a within-school model, since it enables us to capture the relationship between initial status and rates of change within each school in a sample. To study factors underlying the variability in initial status / rate of change slopes across schools, we then treat these slopes as outcomes in a level-3 (between-school) model. This enables us to examine how differences in various school policies, practices and intake characteristics relate to differences in the magnitude of initial status / rate of change slopes. Through applications of three-level growth models of this kind one can attempt to identify, for example, those school factors that appear to eventuate in high mean rates of progress, and in fairly weak relationships between initial status and rates of change (Choi & Seltzer, in prep).

In our analyses, we employed growth models in which change in math achievement scores was modeled as a linear function of grade (see Equation 1). For each of the cohorts whose time-series data we analyzed (e.g., the time-series data for students in schools 308, 143, 142 and 302), various exploratory analyses and formal tests that we conducted pointed to this being a reasonable and parsimonious representation of growth. Note, however, that the various kinds of analyses presented above can be extended to settings in which patterns of change are nonlinear. For example, consider a school in which growth in student achievement tends to accelerate over time. Suppose further that acceleration is more pronounced for those students with relatively high initial levels of achievement. Patterns of this kind can be detected, summarized and studied more formally by using the modeling framework illustrated above to fit models in which student differences in acceleration are modeled as a function of initial status.

As Rogosa (1988) notes, the magnitude of the correlation between initial status and rate of change will typically depend upon the point in time (e.g., grade) chosen as the starting point of a time series. Consider, for example, the pattern of change in school 308. The correlation between initial status (status at the start of grade 7) and rate of change is positive. As such, in Figure 4 we see that students tend to be diverging in their achievement levels over time. As a result, the correlation between status at a later point in time (e.g., status at the start of grade 9) and rate of change will be appreciably larger than the correlation between status at the start of grade 7 and rate of change. The implication of this for design and analysis in longitudinal investigations is that it is extremely important that one choose substantively sensible starting points and end points, i.e., starting points and end points that will enable one to address the key questions motivating a study.

A wide variety of software options are available for fitting the kinds of two-level models presented above. All results reported in this paper were obtained using the Windows version of BUGS (i.e., WinBUGS). Annotated copies of our programs are available upon request. WinBUGS opens up a variety of modeling possibilities that have proved to be extremely valuable in our growth-modeling work (e.g., see Endnote 6). One can also estimate all of the above two-level models using current versions of such structural equation modeling programs as Mplus, LISREL, EQS and AMOS. In addition, all but the very last of the two-level models presented in this paper can be estimated using HLM5. Note, however, that we will very likely see versions of HLM in the near future that enable one to estimate growth models containing interactions between initial status and various demographic characteristics such as gender (see Endnote 7).

In short, a number of software options exist for formulating and fitting models that enable us to employ initial status as a predictor of change. This, in turn, creates a variety of opportunities for conducting analyses that can help illuminate important features of the distribution of student progress within schools.

Endnotes

1. Based on the model for student growth specified in Equation 1, the expected score for student i at the start of grade 7 (i.e., $GRADE_{ti} = 7$) is: $\pi_{0i} + \pi_{1i}(7 - 7) = \pi_{0i}$. Had we not centered *GRADE*, then π_{0i} would have

represented the expected achievement score for student i at the start of grade 0, which clearly is not very useful for our purposes.

2. The parameter estimates and 95% intervals that we report in our tables are posterior medians and .025 and .975 quantiles obtained via WinBUGS. These can be viewed as Bayesian analogues of point estimates and confidence intervals. See Appendix A for further details.

3. Because π_{0i} is now included as a predictor of π_{1i}, the residuals U_{0i} and U_{1i} are assumed to be uncorrelated, i.e., $Cov(U_{0i}, U_{1i}) = 0$.

4. In the equation in which initial status is modeled as a function of *GENDER*, we see that *GENDER* is now centered around its grand mean. As a result, β_{01} still retains its meaning as the expected difference in initial status between girls and boys. However, β_{00} now takes on a more useful interpretation for our purposes, i.e., β_{00} represents the grand mean initial status value for students in school 142 (see, e.g., Bryk & Raudenbush (1992) for a discussion of centering level-2 predictors). In the case of classic ANCOVA models, covariates such as pretest scores (e.g., X_i) are typically deviated from their grand means ($X_i - X.$) (see, e.g., Reichardt, 1979). Similarly, in the model for growth rates in Equation 6, our covariate (π_{0i}) is deviated from its grand mean (i.e., $\pi_{0i} - \beta_{00}$). As a result, β_{10} now represents the expected rate of change for a male student whose initial status value is equal to the grand mean (i.e., $\pi_{0i} - \beta_{00} = 0$). We can also refer to β_{10} as an adjusted rate of change for boys. Similarly, $\beta_{10} + \beta_{11}$ represents the expected rate of change for a female student whose initial status value is equal to the grand mean, i.e., it is an adjusted rate of change for girls. These adjusted rates are analogous to adjusted posttest means in ANCOVA settings.

Note, however, that our model differs from ANCOVA models in a crucially important way. In ANCOVA analyses, we often employ student pretest scores as covariates (e.g., X_i). Such scores contain measurement error. A potential problem is that measurement error contained in X_i can attenuate estimates of the regression coefficient for X_i. This, in turn, can result in underadjustments for initial group differences (see, e.g., Reichardt, 1979). One way of overcoming such problems is to employ latent variables as covariates. Thus, rather than employing observed (e.g., Y) grade 7 math achievement scores as a covariate, we employ the latent variable π_{0i} – a parameter capturing initial status – as our covariate. For discussions of the advantages of employing initial status rather than

observed achievement at time 1 as a covariate in studies of change, see Khoo (2001) and Raudenbush, Bryk, Cheong & Congdon (2000).

5. Consider, for example, a longitudinal study of the effectiveness of two remedial reading programs (e.g., Programs A and B). It may be the case that among students with extreme reading difficulties, rates of progress are, on average, more rapid for students in Program A, whereas among students with milder difficulties, rates of progress tend to be more rapid for students in Program B. We term phenomena of this kind Initial Status × Treatment interactions. See Muthen & Curran (1997) and Seltzer, Choi & Thum (2001) for illustrative examples.

6. Some of the modeling capabilities of WinBUGS that we have found to be particularly useful include the following. First, under normality assumptions, results for parameters of interest are potentially vulnerable to outliers. While one can use WinBUGS to conduct analyses under normality assumptions, one can also use WinBUGS to conduct analyses under heavy-tailed distributional assumptions, which has the effect of downweighting outlying observations. In the context of growth modeling applications, the term outlier refers to outlying time-series observations (e.g., a time-series observation for a student that is unusually low given the overall trend in that student's data), and to outlying individuals (e.g., a student whose rate of change is unusually slow in relation to other individuals in that student's school) (see Seltzer & Choi, in press). Secondly, it is possible in WinBUGS, as well as in HLM, to model differences in school mean rates of change as a function of differences in school mean initial status. Third, in situations where rate of change or acceleration is related to initial status, it is often valuable to consider two or more initial status values of substantive interest and ask: For students who start at these different levels, how much of a difference in achievement scores do we expect to see at various subsequent points in time (e.g., at the start of grade 10)? Estimates and intervals for contrasts of this kind can be obtained easily via WinBUGS. This turns out to be enormously useful in settings in which growth is nonlinear. Fourth, WinBUGS can be used to fit three-level growth models in which we can explore how differences in various school policies and characteristics relate to differences in the magnitude of initial status / rate of change slopes.

7. This prediction is based on the fact that the estimation strategy used by the HLM program to estimate latent variable regressions (see Raudenbush & Sampson, 1999) can in principle be extended to settings in which one wishes to specify interactions between initial status and various dichotomous predictors. Once the necessary modifications are made to the HLM program, it will be possible to conduct analyses of this kind.

Appendix A

Recently developed estimation tools such as the Gibbs sampler make it possible to obtain estimates and intervals for parameters of interest in a wide-range of complex modeling settings (see, e.g., Carlin & Louis, 1996; Gelfand, Hills, Racine-Poon & Smith, 1990; Gelman, Carlin, Stern & Rubin, 1995; Seltzer, 1993; Tanner, 1996). In technical parlance, these iterative techniques provide a means of simulating the marginal posterior distributions of parameters of interest. Thus, for example, the Gibbs sampler could be used to simulate the marginal posterior distribution of the mean growth rate (i.e., β_{10} in Equation 2) for school 308. This distribution, which we will denote $p(\beta_{10}|\mathbf{y})$, would provide us with a summary of the plausibility of different possible values for β_{10} based on the data at hand, and, if available, any relevant prior information. (Note that in our analyses of the LSAY data, we employ "diffuse priors"', i.e., priors that allow the data to dominate our inferences.) Note further that the mode, median and mean of $p(\beta_{10}|\mathbf{y})$ would provide us with various point estimates for β_{11}, and the .025 and .975 quantiles of this distribution would provide us with the Bayesian analogue of a confidence interval.

WinBUGS, which is freely available via the Web, provides a relatively easy means of implementing the Gibbs sampler in a wide array of modeling settings. (Note that versions of BUGS for a variety of other platforms are freely available as well.) In essence, WinBUGS generates large samples of values for parameters of interest (e.g., β_{10}), which enable us to approximate the marginal posterior distributions of these parameters with high degrees of accuracy. Thus, for example, the mean and median of the sample of values generated for β_{10} would provide us with extremely accurate estimates of the mean and median of $p(\beta_{10}|\mathbf{y})$. Similarly, the .025 and .975 quantiles of this sample of values would provide us with highly accurate estimates of the .025 and .975 quantiles of $p(\beta_{10}|\mathbf{y})$.

We ran our WinBUGS analyses on a 700mhz Pentium III computer. For each of our analyses, the WinBUGS program required less than 5 seconds of CPU to generate samples of 10,000 values (i.e., to complete

10,000 iterations of the Gibbs sampler). Depending on the particular analysis, various available diagnostics indicated that samples ranging from approximately 15,000 to 30,000 values were sufficient for simulating the marginal posterior distributions of the parameters in our models with very high degrees of accuracy. Given that in our applications WinBUGS required very little CPU time to complete large numbers of iterations, we based our final results on values generated over 60,000 iterations, thus ensuring extremely high degrees of accuracy.

In applications of the Gibbs sampler, values generated for a given parameter over successive iterations will tend to be correlated to some extent. (Note that this helps explain why fairly large samples of values are needed to simulate marginal posteriors of interest with high degrees of accuracy. In short, the effective sample size of a set *m* correlated values will be smaller than *m*.) One needs to be alert, however, to situations in which degrees of autocorrelation are extremely high. The technical phrase for this in the literature is "poor mixing"'. To help detect any possible convergence or mixing problems, for each analysis we ran WinBUGS twice using different starting values for the variance components in our models, and using different seeds. We then compared results (e.g., posterior medians and .025 and .975 quantiles) based on the output from each run, and inspected various diagnostic plots and statistics (i.e., trace plots, autocorrelation function (ACF) plots, and Raftery-Lewis statistics). These procedures gave no indication of any convergence or mixing problems. On the contrary, mixing appeared to be quite good in our applications, and convergence was rapid. While various diagnostics are available in WinBUGS (e.g., trace plots, ACF plots), the programs CODA (Best, Cowles & Vines, 1995) and BOA (Smith, 2000) provide users with a comprehensive set of diagnostic tools. CODA can be downloaded from the BUGS website, and BOA can be downloaded from the Web as well.

The resulting marginal posterior medians and .025 and .975 quantiles that we obtained via WinBUGS constitute the parameter estimates and 95% intervals that we report in our Tables. For checking purposes, we re-ran the analyses reported in Tables 1-5 using HLM. The results that we obtained via HLM for key fixed effects (e.g., β_{10} in Equation 2; β_{11} in Equation 6) and latent variable regression coefficients (e.g., b in Equations 3 and 6) are extremely similar to those that we obtained using WinBUGS. Finally, we re-ran the analysis reported in Table 6 using Mplus (Muthen & Muthen, 1998). The Mplus results for key latent variable regression parameters (i.e., b_1 and b_2 in Equation 7) are extremely close to those reported in Table 6.

References

Best, N., Cowles, M. & Vines, K. (1995). CODA: Convergence diagnosis and output analysis software for Gibbs sampling output, Version 0.30. MRC Biostatistics Unit, Cambridge University.

Bock, R.D. & Zimowski, M. (1997). Multiple group IRT. In W. van der Linden & R. Hambleton (Eds.), *Handbook of Modern Item Response Theory* (pp. 33-49). New York: Springer Verlag.

Bryk, A. S. & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101, 147-158.

Bryk, A. S. & Raudenbush, S. W. (1988). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. *American Journal of Education*, 97, 65-108.

Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage Press.

Bryk, A., Thum, Y., Easton, J. & Luppescu, S. (1998). Assessing school academic productivity: The case of Chicago school reform. *Social Psychology of Education*, 2, 103-142.

Carlin, B. & Louis, T. (1996). *Bayes and empirical Bayes methods for data analysis*. London: Chapman and Hall.

Choi, K. & Seltzer, M. (in prep). Addressing Questions Concerning Equity in Longitudinal Studies of School Effectiveness: Modeling Heterogeneity in Relationships between Initial Status and Rates of Change. CSE Technical Report. Los Angeles: Center for Research on Evaluation, Standards, and Student Peformance, UCLA.

Gelfand, A., Hills, S., Racine-Poon, A. & Smith, A. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association, 85*, 972-985.

Gelman, A., Carlin, J., Stern, H. & Rubin, D. (1995). *Bayesian data analysis*. London: Chapman & Hall.

Hays, W. (1988). *Statistics* (4th ed.). New York: Holt, Rinehart & Winston.

Khoo, S.T. (2001). Assessing program effects in the presence of treatment-baseline interactions: A latent curve approach. *Psychological Methods*, 6, 234 -257.

Miller, J., Kimmel, L., Hoffer, T. & Nelson, C. (1999). *Longitudinal Study of American Youth: User's Manual*. Chicago: International Center for the Advancement of Scientific Literacy, Academy of Sciences.

Muthen, B. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins & J. Horn (Eds.), *Best*

methods for the analysis of change: Recent advances, unanswered questions, future directions (pp. 1-17). Washington D.C.: American Psychology Association.

Muthen, B. & Curran, P. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power. *Psychological Methods*, 2, 371-402.

Muthen, B. & Khoo, S.T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences*, 10, 73-102.

Muthen, L. & Muthen, B. (1998). *Mplus user's guide*. Los Angeles: Muthen and Muthen.

Raudenbush, S., Bryk, A., Cheong, Y. & Congdon, R. (2000). *HLM5: Hierarchical linear and non-linear modeling*. Homewood, IL: Scientific Software International.

Raudenbush, S. & Sampson, R. (1999). Assessing direct and indirect associations in multilevel designs with latent variables *Sociological Methods and Research*, 28, 123-153.

Reichardt, C. (1979). The statistical analysis of data from non-equivalent group designs. In T.D. Cook & D.T. Campbell (Eds.), *Quasi-Experimentation: Design and Analysis Issues for Field Settings* (pp. 147-205). Boston: Houghton-Mifflin.

Rogosa, D. (1988). Myths about longitudinal research. In K.W. Schaie, R.T. Campbell, W. Meredith & S.C. Rawlings (Eds.), *Methodological Issues in Aging Research* (pp. 171-209). New York: Springer.

Sanders, W. & Horn, S. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation*, 9, 299-311.

Seltzer, M. (1993). Sensitivity analysis for fixed effects in the hierarchical model: A Gibbs sampling approach. *Journal of Educational Statistics, 18*, 207-235.

Seltzer, M. & Choi, K. (in press). Sensitivity analysis for multilevel models. In N. Duan & S. Reise (Eds.), *Multilevel modeling: Methodological advances, issues, and applications*. Mahwah, NJ: Lawrence Earlbaum.

Seltzer, M., Choi, K. & Thum, Y. (2001). *Latent variable modeling in the hierarchical modeling framework: Exploring initial status × treatment interactions in longitudinal studies of interventions*. CSE Technical Report. Los Angeles: Center for Research on Evaluation, Standards, and Student Testing, UCLA.

Seltzer, M., Frank, K. & Bryk, A. (1994). The metric matters: The sensitivity of conclusions about growth in student achievement to choice of metric. *Educational Evaluation and Policy Analysis*, 16, 41-49.

Smith, B. (2000). *Bayesian output analysis program (BOA) version 0.50.* The University of Iowa College of Public Health.

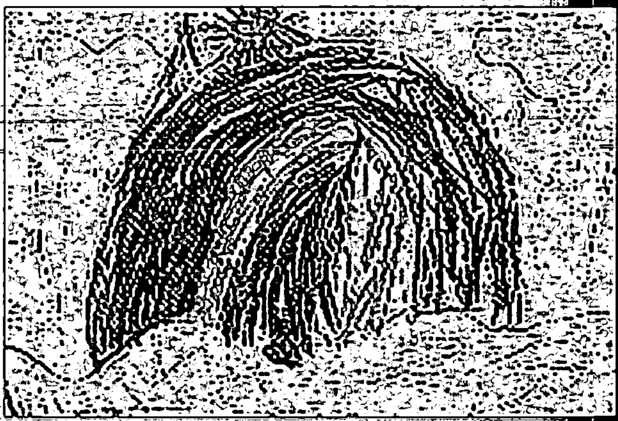
Spiegelhalter, D., Thomas, A. & Best, N. (2000). Win*BUGS, Version 1.3, User Manual.* MRC Biostatistics Unit, Cambridge University.

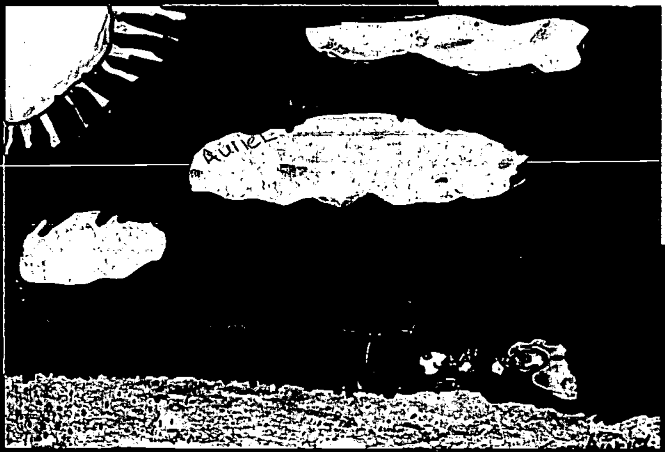
Tanner, M. (1996). *Tools for statistical inference* (3rd ed.). New York: Springer-Verlag.

Thum, Y. M. (2001). Measuring progress toward a goal: Estimating teacher productivity using a multivariate multilevel model for value-added analysis. Under review.

Willett, J. (1988). Questions and answers in the measurement of change. In E.Z. Rothkopf (Ed.), *Review of Research in Education*, 15, (pp. 345-422). Washington, D.C.: American Educational Research Association.

Willms, J.D. (1992). *Monitoring school performance: A guide for educators.* Washington, D.C.: The Falmer Press.





NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").