American Educational Research Journal
Spring 2005, Vol. 42, No. 1, pp. 3–42

Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance

Russell W. Rumberger
University of California, Santa Barbara
Gregory J. Palardy
University of Georgia

This study investigated the relationships among several different indicators of high school performance: test scores, dropout rates, transfer rates, and attrition rates. Hierarchical linear models were used to analyze panel data from a sample of 14,199 students who took part in the National Education Longitudinal Survey of 1988. The results generally support the notion of an alternative as opposed to a common view of school effectiveness: Schools that are effective in promoting student learning (growth in achievement) are not necessarily effective in reducing dropout or transfer rates. In fact, after control for student inputs, high schools exhibit relatively little variability in dropout rates but considerable variation in transfer rates. In addition, characteristics of schools that contributed to performance in one area often did not contribute to performance in another. Given these findings, the authors suggest that, along with test scores, dropout and transfer rates should be used to judge school performance.

Keywords: dropouts, school effectiveness, student achievement.

School effectiveness has remained an important concern of researchers and policymakers for many years. Researchers have sought to explain why some schools are more effective than others, while policymakers have sought

---

Russell W. Rumberger is a Professor in the Gevirtz Graduate School of Education, University of California, Santa Barbara, CA 93106, e-mail *russ@education.ucsb.edu*. His areas of interest are the schooling of disadvantaged students, school dropouts, and school effectiveness.

Gregory J. Palardy is an Assistant Professor in the Department of Educational Psychology, University of Georgia, 325D Aderhold Hall, Athens, GA 30602-7143; e-mail *gpalardy@uga.edu*. His areas of interest are research methodology, school and teacher effectiveness, and issues related to educational opportunity.

---

ways to identify and improve the performance of low-achieving schools. In both the policy and research arenas, school effectiveness is most often assessed via student test scores. Test scores provide a direct measure of student learning, which is viewed as one of the most important outcomes of schooling. As a result, an increasing number of states have already developed accountability systems based on test scores. Recent federal legislation, the No Child Left Behind (NCLB) Act of 2001, requires that all states begin annual testing of students to measure "adequate yearly progress" of schools and districts in meeting state-defined standards (U.S. Department of Education, 2003).

Yet, school effectiveness can be assessed through other measures of student performance. At the high school level, school effectiveness can be measured via two related indicators: dropout rates, which indicate the percentage of students who quit school before completion, and graduation rates, which indicate the percentage of students who remain in school and earn a high school diploma.[1] NCLB requires that states eventually incorporate graduation rates into their annual report cards.[2]

One reason for using multiple indicators of school performance is that some schools may perform better on one type of outcome than another. This may be especially true if the resources and practices required to raise performance in one area are different from those required in another area. For example, if teacher resources are more critical in raising test scores, but other resources are more critical in keeping students from dropping out, then schools may have to choose where to focus scarce resources. Other contributors to school performance, such as academic climate, may be effective in improving performance in all areas.

Another reason for using multiple indicators of school performance measures is that the goal of improving one outcome may actually conflict with the goal of improving the other. One way to improve test scores is to increase the number of students who drop out or otherwise leave school, since school leavers generally have lower test scores than other students (Bowditch, 1993; Riehl, 1999; Rumberger & Larson, 1998). For example, both the New York City and Houston school systems have been investigated over the past year for "pushing out" their low-performing students into alternative, nondiploma programs or classifying them as transfers so they would not be counted as dropouts (Howard, 2004; Lewin & Medina, 2003).

To date, little research has investigated the relationships between multiple indicators of school performance. Although there is a large body of research on school effectiveness, most studies rely on a single indicator of school performance.[3] In the present study, in contrast, we examined the relationships among several complementary indicators of school performance at the high school level using the same national sample of schools. We addressed two basic questions: (a) Are schools that are effective in raising test scores also effective in reducing dropout and transfer rates? and (b) What school characteristics predict these alternative indicators of performance?

Research Literature

A vast body of research on school effectiveness has focused on understanding how various aspects of schools are related to differences in student outcomes. This research has assessed a number of student outcomes. The most frequent outcome is academic achievement as reflected in student test scores (e.g., Carbonaro & Gamoran, 2002; Coleman, Hoffer, & Kilgore, 1982; Gamoran, 1996; Lee & Bryk, 1989; Lee & Smith, 1993, 1995; Lee, Smith, & Croninger, 1997; Witte & Walsh, 1990), but a number of studies have examined other outcomes, including dropout rates (e.g., Bryk & Thum, 1989; Lee & Burkam, 2003; McNeal, 1997; Rumberger, 1995; Rumberger & Thomas, 2000), absenteeism (Bryk & Thum, 1989; Phillips, 1997), engagement (Lee & Smith, 1993), and social behavior (Lee & Smith, 1993). This research has also been guided by a number of theoretical perspectives and conceptual frameworks. In the following, we provide a brief review of these perspectives and frameworks before summarizing the empirical literature on school effectiveness based on alternative student outcomes.

Theoretical Perspectives

Two alternative perspectives on the relationship between school characteristics and student outcomes underlie most research on school effectiveness. The first and most widely held perspective is what we characterize as a *common* view of the schooling process according to which all aspects of school performance—test scores, attendance, and dropout—are influenced by a variety of school characteristics in similar ways. That is, the factors that influence student learning, such as a rigorous curriculum, high teacher expectations, and a strong academic climate, are also the ones that influence whether students remain in school. This perspective is consistent with the tenets of school effectiveness research (Purkey & Smith, 1985) as well as several existing theories of school dropout and departure that view student departure as largely voluntary and as the result of students becoming disengaged from school and from learning (Finn, 1989; Wehlage, Rutter, Smith, Lesko, & Fernandez, 1989).

The other perspective, which we characterize as a *differentiated* view, holds that different factors may influence different student outcomes. For example, dropout theories suggest that student departure is related to problems with not only student learning and academic engagement but social engagement as well (Finn, 1989; Wehlage et al., 1989). Promoting student social engagement may require different resources (e.g., counselors) and different policies (e.g., discipline policies) than those required to promote academic engagement and student learning. Given limited resources and expertise, schools that are effective in improving student academic engagement and learning may not necessarily be effective in improving student social engagement and completion. Both perspectives, however, view student withdrawal (dropout or transfer) as largely voluntary and the result of benign neglect.

The differentiated perspective is also consistent with the notion that schools may adopt explicit policies and school personnel may make conscious decisions to cause students to *involuntarily* withdraw from school, either as dropouts or as transfers (Bowditch, 1993; Fine, 1991; Riehl, 1999; Romo & Falbo, 1996). These policies and decisions may concern low grades, poor attendance, misbehavior, or being overage and can lead to suspensions, expulsions, or forced transfers. This perspective considers a school's own agency, rather than just that of the student, in producing dropouts and transfers. One metaphor that has been used to characterize this process is discharge: "Students *drop out* of school, schools *discharge* students" (Riehl, 1999, p. 231).

Both perspectives may be useful in explaining alternative aspects of school performance. School effectiveness research can help identify characteristics of schools that promote engagement, improve learning, and thereby reduce the likelihood of students prematurely withdrawing. School dropout research can help identify policies and practices of schools that cause students to leave school involuntarily as either dropouts or transfers. Indeed, as we demonstrate subsequently, existing research lends support for both perspectives.

Conceptual Frameworks

Although school effectiveness research has been guided by a number of conceptual frameworks (e.g., Frank, 1998; Shavelson, McDonnell, Oakes, & Carey, 1987; Tagiuri, 1968; Willms, 1992), the most common is based on an economic model of schooling (e.g., Hanushek, 1986, 1994; Levin, 1997) that focuses on two distinct aspects of schools: school *inputs* and school *processes* (policies and practices). School inputs include the characteristics of the student body, such as academic background and socioeconomic status (SES); structural features, such as type of school (public, Catholic, other private) and size; and school resources, such as teachers and textbooks. Many studies have attempted to identify the school inputs related to differences in school outputs (Glasman & Biniaminov, 1981; Hanushek, 1986; Teddlie & Reynolds, 2000). The well-known Coleman report, for example, which examined three types of school inputs—characteristics of the student body, characteristics of the teachers, and facilities and curriculum—showed that the extent of equality in the distribution of these inputs among schools was in the reverse order of their importance in affecting student achievement (see Coleman, 1990, for a summary and more recent commentary). That is, the input that mattered most, the characteristics of the student body, was the least equitably distributed among schools in the United States, whereas the input that mattered least, the facilities and curriculum, was the most equitably distributed; distribution of teachers and impact of teachers were intermediate.

The second aspect of schools concerns the processes and practices that take place within them, including leadership and decision-making practices (Hannaway & Carnoy, 1993), instructional practices (Rowan, Correnti, & Miller, 2002), and the overall academic and social climate of the school (Freiberg, 1999). Many studies of school effectiveness have sought to understand what school practices affect student achievement and the extent to which these practices explain how and why school inputs make a difference. For example, studies have demonstrated that Catholic schools are more effective than public schools because of their strong academic climate and the strong social relationships or social capital among parents and school personnel (Bryk, Lee, & Holland, 1993; Coleman & Hoffer, 1987; Coleman et al., 1982).

Empirical Research

A large body of empirical research has demonstrated that a number of specific school characteristics within these two domains can explain differences in school performance, particularly as measured via test scores and dropout rates. Yet, the research findings are far from consistent. In some cases, the impact of school characteristics on the same outcome varies across studies. In other cases, the impact of school characteristics varies across outcomes. In the sections to follow, we provide a brief summary of this research and highlight a few of the inconsistencies.

Student Characteristics

Research has demonstrated that a wide variety of individual student characteristics are related to student outcomes, including demographic characteristics, such as ethnicity and gender; family characteristics, such as SES and family structure; and academic characteristics, such as previous achievement and retention. These characteristics have been shown to relate to such student outcomes as engagement, achievement (test scores), and dropout (Bryk & Thum, 1989; Chubb & Moe, 1990; Goldschmidt & Wang, 1999; Johnson, Crosnoe, & Elder, 2001; Lee & Smith, 1995, 1999; McNeal, 1997; Rumberger, 1995; Rumberger & Thomas, 2000).

Student characteristics influence student achievement not only at an individual level but also at an aggregate or social level. That is, the social composition of students in a school (sometimes referred to as contextual effects) can influence student achievement apart from the effects of student characteristics at an individual level (Gamoran, 1992). Studies have shown that the social composition of schools predicts school engagement, achievement, and dropout rates even after control for the effects of individual background characteristics of students (Bryk & Thum, 1989; Chubb & Moe, 1990; Coleman et al., 1966; Goldschmidt & Wang, 1999; Jencks & Mayer, 1990; Lee & Burkam, 2003; Lee & Smith, 1999; Mayer, 1991; McNeal, 1997; Rumberger, 1995; Rumberger & Palardy, in press; Rumberger & Thomas, 2000). One measure of school composition, the mean SES of the student body, has generally shown a positive and significant effect on student achievement (e.g., Coleman et al., 1966; Lee & Bryk, 1989; Lee & Smith, 1999). However, its impact on dropout rates has been inconsistent, with some studies showing similar impacts (e.g., Goldschmidt & Wang, 1999; Mayer, 1991; Rumberger, 1995) and others revealing no significant impacts (e.g., Bryk & Thum, 1989; Lee & Burkam, 2003; McNeal, 1997; Rumberger & Thomas, 2000).

Structural Characteristics

Structural characteristics, such as school location (urban, suburban, rural), size, and type of control (public, private), are also associated with school performance. Although widespread achievement differences have been observed among schools on the basis of structural characteristics (e.g., between public and private schools), what remains unclear is whether structural characteristics themselves account for these differences or whether they are related to differences in student characteristics and school resources often associated with the structural features of schools. For example, several empirical studies have shown that average achievement levels are higher and dropout rates lower in private schools, in general, and Catholic schools, in particular, than in public schools, even after differences in student characteristics and resources have been taken into account (Bryk & Thum, 1989; Bryk et al., 1993; Chubb & Moe, 1990; Coleman & Hoffer, 1987; Coleman et al., 1982; Morgan & Sorensen, 1999; Rumberger & Thomas, 2000).

Yet other studies have revealed little or no private school advantage in either test scores or dropout rates (Alexander & Pallas, 1985; Carbanaro & Gamoran, 2002; Gamoran, 1996; Lee & Burkam, 2003; Willms, 1985). Moreover, it has been suggested that controlling for differences in demographic characteristics may still not adequately take into account fundamental and important differences (e.g., attitudes and motivation) among students in the two sectors (Witte, 1992, p. 389). Two studies of engagement also yielded inconsistent results: One revealed no differences between public and private schools in terms of levels of student engagement (Johnson et al., 2001), while another showed that engagement was significantly higher in private nonparochial schools than in public and private schools (Lee & Smith, 1995).

More recently, there has been considerable interest in another structural feature of schools: school size. While some studies have shown that large schools have significantly lower test scores and higher dropout rates than medium-sized or small schools (Lee & Burkam, 2003; Lee & Smith, 1997), other studies have revealed no significant effects of school size overall (e.g., McNeal, 1997; Phillips, 1997) or a significant impact only on lower SES schools (Rumberger, 1995).

School Resources

School resources consist of both fiscal resources and the material resources they can provide, such as teachers and textbooks. There is considerable debate in the research community about the extent to which school resources contribute to school effectiveness (Hanushek, 1989, 1994, 1997; Hedges, Laine, & Greenwald, 1994). And while there is a general consensus that teacher quality matters (Hanushek, 1986), the characteristics of teachers that contribute to school effectiveness, such as credentials and experience, are less clear (Darling-Hammond, Berry, & Thoreson, 2001; Goldhaber & Brewer, 2001; Wayne & Youngs, 2003). Beyond the quality of teachers, there is at least some evidence that the quantity of teachers—as measured by pupil/teacher ratios—has a positive and significant effect on student dropout rates (McNeal, 1997; Rumberger & Thomas, 2000).

School Processes

Despite all of the attention and controversy surrounding the previous factors associated with school effectiveness, it is the area of school processes that many people believe holds the most promise for understanding and improving school performance. Although most individual schools, or at least most public schools, have little control over school inputs (student characteristics, resources, and their structural features), they can and do have a fair amount of control over how they are organized and managed, the teaching practices they use, and the climate they create for student learning—features referred to as school processes.[4]

A number of school policies and practices have been shown to affect school performance. Some studies have revealed that school organizational practices, such as decision-making practices (including teachers' and parents' involvement in decision making), affect student achievement in middle and high schools (e.g., Lee & Smith, 1993, 1995; Lee et al., 1997; Morgan & Sorensen, 1999). Yet another study of middle schools showed that communal organization, including democratic governance, had no impact on either mathematics achievement or attendance (Phillips, 1997). Other studies have revealed that teachers' expectations and efficacy, as well as their instructional practices, affect student learning in high school (Carbonaro & Gamoran, 2002; Lee et al., 1997).[5] Still other studies have shown that an array of indicators related to the social and academic climate of schools—such as the number of advanced academic courses taken by students, the amount of homework done by students, and teachers' expectations and interest in students—influence a number of school performance indicators, including student achievement, engagement, and dropout (Bryk & Thum, 1989; Croninger & Lee, 2001; Gamoran, 1996; Lee & Burkam, 2003; Lee et al., 1997; Phillips, 1997; Rumberger, 1995).

In summary, some existing studies support the idea that a similar array of school characteristics—student composition, structure, resources, and processes—influence test scores, dropout rates, and other indicators of school performance in similar ways, supporting a common perspective of effective schools. However, a number of studies have shown that some characteristics have different effects on these alternative indicators of school performance, supporting a differential perspective of effective schools. That is, the characteristics that promote learning may not be the ones that promote lower dropout or transfer rates. As a result, schools that are effective in raising test scores may not be effective in retaining students. In fact, schools may improve their test scores by deliberately and systematically discharging the most difficult and lowest performing students. Yet, no study to date has directly compared the same schools on these alternative indicators of school performance to determine the extent to which they are related. This was the purpose of the present study.

Method

Data and Samples

This study was based on data from the National Education Longitudinal Study of 1988 (NELS). NELS, initiated in 1988, is a national longitudinal study of a representative sample of 25,000 eighth-grade students conducted by the National Center for Education Statistics (NCES). Base-year data were collected from questionnaires administered to students, their parents and teachers, and the principal of their school. Students were also given a series of achievement tests in English, math, science, and history/social studies. Follow-up data were collected in 1990, 1992, 1994, and, most recently, 2000 on a subset of the original sample (Carroll, 1996). Although the high school data are more than a decade old, they represent the largest and most comprehensive data set currently available in regard to American high schools.[6]

The present study was based on a sample of 14,199 students with valid questionnaire data from the 1988, 1990, and 1992 survey years; in 1990, these students attended a total of 912 high schools. Respondents without data on high school attended in 1990, those who did not have base year test scores, and those who did not attend a high school in 1990 with at least 5 students in the sample were not included.[7] The latter constraint was necessary to ensure reasonable reliability estimates of within-school parameter estimates. Students also had to have test data from at least 2 of the questionnaire years so that more reliable individual linear growth estimates could be specified.[8]

To adjust for nonrandom selection and retention in the longitudinal sample as well as nonrandom responses to the questionnaires and tests, NCES constructed student-level panel weights for members of the 1988/1990/1992 longitudinal cohort that can be used to generate national population estimates (Ingles et al., 1994). However, because NELS was a cohort study of students, NCES did not construct school-level weights for the high schools students attended in 1990. Some researchers who have used these data in conducting high school effectiveness studies have constructed their own school-level weights (e.g., Lee et al., 1997; Morgan & Sorensen, 1999), while other researchers have not (e.g., Carbanaro & Gamoran, 2002; Gamoran, 1996). Because of the limitations of the hierarchical linear modeling (HLM) software employed in this study, we were unable to use sample weights.[9] Nonetheless, we did investigate the characteristics of the student and school samples and found them to be representative of both students and schools nationally.[10]

Variables

This study was based on a conceptual framework that recognizes school performance as a multilevel phenomenon involving students as well as schools (see Figure 1).[11] Student outcomes are a function of the characteristics and experiences of individual students. School outputs represent the aggregated

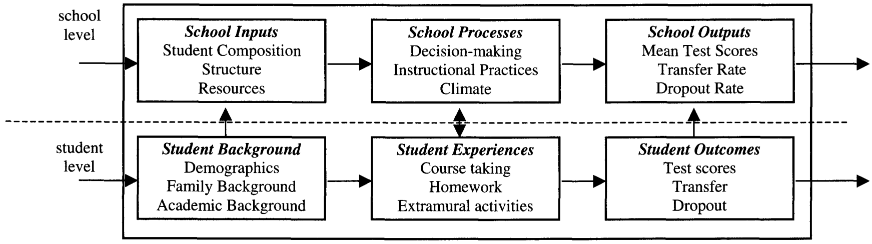

Figure 1. **Conceptual framework for analyzing school performance.**

performance of the students within a school and are a function of school inputs and school processes.

We focused on four measures or indicators of student performance at the high school level: achievement growth (or learning) over 4 years of high school, proportion of 10th-grade students who dropped out between Grades 10 and 12, proportion of 10th-grade students who transferred over the same period, and total proportion of 10th-grade students who left school before 12th grade (a sum of dropouts and transfers). These four measures were selected because they represent alternative yet important indicators of school effectiveness. Achievement growth and dropout rates have been used in past studies of school performance, albeit never in the same study.[12] Although transfer and attrition rates have rarely been used in past studies of school performance (for a recent exception, see Rumberger & Thomas, 2000), they may provide another indicator of how well schools are retaining their students, what some researchers have referred to as "holding power" (e.g., Balfanz & Legters, in press). Moreover, research indicates that mobile students are less likely to complete high school than students who remain in the same school (Rumberger & Larson, 1998; Swanson & Schneider, 1999; Teachman, Paasch, & Carver, 1996), and thus how well schools are retaining their students provides another indicator of school performance.

Achievement was measured with an arithmetic mean of scores on standardized achievement tests in four academic subjects (math, science, reading, and social science) administered in the spring semesters of 1988, 1990, and 1992, when most students were enrolled in Grades 8, 10, and 12, respectively. We used all four tests because they provide a more robust and comprehensive indicator of learning than a single test. And although studies have shown that the effects of school characteristics may differ across subject area tests (Gamoran, 1996; Ma, 2001), we created a single achievement composite measure to simplify comparisons with the other outcomes.[13] A similar practice is used by states to construct school accountability indicators.[14] In this study, subject area test scores from each year were calibrated to a common vertical scale via item response theory methods and then transformed to a t scale standardized on 10th-grade scores ($M = 50$, $SD = 10$).

Three other outcome measures were constructed. Student dropout indicates whether a student dropped out of school between 1990 and 1992, student transfer indicates whether a student transferred to a different school between 1990 and 1992, and student attrition indicates whether a student left school between 1990 and 1992.

Because our conceptual framework views school performance as a function of both student-level and school-level variables (see Figure 1), both types of variables were included in our analysis. The student-level variables included in this study were used to control for differences in the background characteristics of students before they entered high school (except for a variable measuring whether the student's family moved between the 10th and 12th grades). Adjusting for differences in student background characteristics generated more accurate estimates of a school's effectiveness by controlling for those factors over which most (public) schools have little control, that is, the characteristics of their student bodies. After controlling for those differences, we examined which school-level predictors explained the remaining variability in performance among schools. Detailed descriptions of all of the variables used in this study, together with descriptive statistics, are provided in Appendix A.

Student-Level Variables

We included several types of student background variables. The first type captures demographic characteristics of students, which were measured through a series of dummy variables associated with minority status (Asian, Black, Hispanic, or Native American). The second type captures family background characteristics, measured with three variables: SES, which is a composite measure developed by NCES that reflects parental education, income, and occupational status; nontraditional family, which identifies students who did not live with both parents; and family mobility, representing families who moved between the 10th and 12th grades. Two recent studies have shown that students from single-parent families and stepfamilies are more mobile than students from two-parent families, suggesting that nontraditional families are less stable and more prone to move or change schools (Astone & McLanahan, 1994; Rumberger & Larson, 1998).

The final type of student background variable was *academic and social background,* which was measured via several variables: self-reported grades, plans to finish college, misbehavior, retention between the first and eighth grades, and number of school changes between the first and eighth grades. These variables have been shown to predict student achievement, dropout, and mobility in previous studies (e.g., Lee et al., 1997; Rumberger, 1995; Rumberger & Thomas, 2000; Swanson & Schneider, 1999). Two additional variables were included in the achievement growth model: whether the student dropped out between 1990 and 1992 and whether the student transferred between 1990 and 1992.

School-Level Variables

On the basis of our conceptual framework, we constructed several types of school-level characteristics.[15] The first measured the student composition of schools. Nine student composition variables were constructed by aggregating individual values for the student sample to the school level: (a) mean SES of students in the school; (b) proportion of Black or Hispanic students in the school[16]; (c) proportion of students who came from nontraditional families (single-parent families and stepfamilies), since previous research has shown that such families involve less home supervision and school participation than two-parent families (Sui-Chu & Willms, 1996); (d) proportion of students who changed residences between 1990 and 1992; (e) mean grade point average of students in Grades 6–8; (f) mean amount of misbehavior among eighth-grade students; (g) proportion of students who were retained between Grades 1 and 8; and (h) proportion of students who planned to finish college. The first four variables captured demographic characteristics of students, while the last four captured students' academic background characteristics.[17]

The second category of school-level characteristics measured the *structural* characteristics of schools: whether the school was Catholic or another type of private school (as opposed to public), whether the school was located in an urban or rural area (vs. suburban), and the size of the school.[18] The third category measured school *resources:* student-teacher ratio; proportions of teachers with full credentials, credentials in the subjects they taught, a bachelor's degree in the subject they taught, and advanced degrees; and mean salary of teachers at the school. The final category measured school *processes* (practices and climate): quality of teaching and teacher support reported by the students[19]; teachers' efficacy and expectations for student achievement[20]; teachers' perceptions of school leadership, collaboration with other teachers, and control over school policy[21]; parent involvement[22]; academic climate, as measured by mean number of hours of homework reported by students and mean number of college-preparatory courses taken by students; social climate, as measured by proportion of students who reported feeling unsafe at the school and mean number of reports of classroom disruptions; and disciplinary climate, as measured by proportion of students who reported that the school's discipline policies were fair.

Statistical Models

Because students in the NELS data sample are nested within schools, we used hierarchical linear models. HLM methods have been developed over the past 20 years to deal with issues specific to nested/multilevel data sets, including aggregation bias, misestimation of errors, and the unit of analysis problem (Raudenbush & Bryk, 2002). In this study, we used three types of hierarchical linear models. First, we used a three-level achievement growth model to estimate students' achievement growth or learning on a four-subject (math, reading, science, and history) achievement test composite over the 4-year period during which they were attending high school.[23] Second, we used a two-level multinomial logistic regression model to estimate school dropout and transfer rates. Third, we used a two-level binomial logistic regression model to estimate school attrition rates. All three models involved the same set of student-level and school-level independent variables. Detailed descriptions of the models are provided in Appendix C.

Several specific models were estimated for each model type. First, we estimated a null model with no predictors (with the exception of a Level 1 growth parameter in the growth model). This model estimated the overall between-school variance in each of the outcomes, providing a baseline for comparisons with later models. Second, we estimated a student model to control for the background characteristics of students in each school. By centering all of the continuous student-level predictors on the grand mean and leaving the remaining variables uncentered, the intercept term provides an estimate of the expected school outcome for the same average or "typical" student at each school.

Third, we estimated a school composition model that adjusted for the effects of the aggregated student characteristics, which previous research has shown to influence student outcomes above and beyond individual background characteristics (Gamoran, 1992; Willms, 1992). This model provides a better measure of a school's performance than the observed, average performance of students because it adjusts for differences in both the individual and compositional effects of students that schools enroll, which can account for a substantial amount of the observed differences in performance (Raudenbush & Willms, 1995).[24] For each educational outcome, we used the empirical Bayes estimate of each school's residual as a measure of the school's effectiveness. We divided each random effect by the standard deviation of the respective outcome from the null model to express school effectiveness in a standardized form, equivalent to the between-school effect size (ES) commonly used in school effectiveness studies, facilitating comparisons among outcome measures (e.g., Lee & Smith, 1997).

Finally, we estimated a series of additional school-level models to compare the effects of various school predictors on different school outcome measures. These models were estimated in steps, initially by entering similar types of predictor variables—structural, resource, and process—and then retaining the significant variables from the group before entering predictors from the next group. All of the predictor variables were converted to between-school effect sizes.[25]

Results

Alternative Measures of School Effectiveness

The first step in the analysis involved estimating a series of unconditional models without any student-level or school-level predictor variables to provide estimates of average school outcomes and the variability in these outcomes among schools. The achievement model estimated the average initial achievement of students in the 8th grade, before entering high school, and the average achievement growth between the 8th and 12th grades. Because the present study focused on what students learn during high school, we highlight the findings on achievement growth or learning. The other models estimated dropout, transfer, and attrition rates between Grades 10 and 12. The estimated parameters from the models are shown in the top section of Table 1.

Table 1
Parameter Estimates for Unconditional Model and Descriptive Statistics for Alternative School Outcomes

Type of measure	Achievement growth	Dropout	Transfer	Attrition
Parameter estimates for unconditional model				
Mean β	7.850**	−2.594**	−2.729**	−1.938**
Parameter variance				
Between students	8.056**			
Between schools	2.827**	0.477**	0.688**	0.412**
Percentage variance between schools	26.0			
Reliability				
Within students (Level 1)	.318			
Between students (Level 2)	.607	0.301	0.350	0.395
Descriptive statistics for school outputs				
Overall sample mean	7.85	0.07	0.06	0.13
Range of plausible values	4.56–11.14	0.02–0.22	0.01–0.25	0.04–0.34

Note. Parameter estimates for initial status are not shown. The range of plausible values represents the 95% range of expected school means, based on the assumption of normality for variance estimates and calculated according to the guidelines of Raudenbush and Bryk (2002, p. 71). The sample means and range of plausible values for dropout, transfer, and attrition rates were converted to proportions based on the formula $1/\{1+\exp[-(\beta)]\}$.
$^{**}p < .01$.

The mean values indicate that the average achievement in the four academic subjects increased by 7.85 points over the 4 years of high school. Between Grades 10 and 12, dropout rates averaged 7%, transfer rates averaged 6%, and attrition rates averaged 13%. However, these averages mask considerable variability among schools, as indicated by the significance tests for the estimated parameter variances (see Table 1). This variability can be illustrated by calculating a range of plausible values for each outcome under the assumption that they are normally distributed. These values, shown in the bottom section of Table 1, reveal considerable variation in all four outcomes among schools. For example, the average growth rate in achievement among high schools would be expected to range from a low of 4.56 points to a high of 11.14 points. In other words, in some high schools students learn more than twice as much as students in other high schools. Dropout, transfer, and attrition rates vary even more. For instance, dropout rates vary from a low of 2% to a high of 22%. These results suggest that where students attend high school has a great deal to do with how much they learn and their chances of graduating.

Yet, these differences in school outcomes are not simply a reflection of schools' effectiveness; they also reflect differences in the characteristics of students who attend them. To account for these differences, we estimated another model that controlled for differences in student background characteristics and for differences in the compositional effects of students. The results are displayed in Table 2. We interpret ESs in the same way as Lee and Smith (1997), considering those of 0.5 or above large, those between 0.3 and 0.5 moderate, those from 0.1 to 0.3 small, and those below 0.1 negligible.

The results reveal that a number of student-level and school-level compositional variables predict educational outcomes. In some cases, the same variables predict different educational outcomes in similar ways; in other cases, they do not. For example, after control for other student-level predictors, student SES had a small yet statistically significant effect on student learning (ES = 0.207); a large, significant effect on dropout (ES = –0.610); and a negli-

Table 2
Parameter Estimates for Student and Social Composition Model

Measure	Achievement growth	Dropout	Transfer	Attrition
Mean	7.850**	−2.594**	−2.729**	−1.938**
Student-level predictors				
SES	0.207**	−0.610**	0.039	−0.290**
Nontraditional family	−0.153**	0.569**	0.261**	0.469**
GPA in Grades 6–8	0.463**	−0.586**	−0.220**	−0.453**
Misbehavior in Grade 8	−0.016	0.311**	0.080†	0.236**
Retained in Grades 1–8	−0.612**	1.020**	0.129	0.733**
College aspirations in Grade 8	0.170**	−0.680**	−0.118	−0.458**
Asian	0.605**	−0.256	0.131	0.025
Black	−0.548**	−0.543**	−0.162	−0.402*
Hispanic	0.065	0.161	0.130	0.190
Native American	−0.017	1.028*	0.483†	1.012**
Moved during Grades 10–12	−0.162*	0.825**	1.627**	1.496**
Changed schools in Grades 1–8	0.076**	0.195**	0.228**	0.240**
Dropped out in Grades 10–12	−0.700**			
Transferred during Grades 10–12	−0.131			
School social composition				
Mean SES	0.253**	−0.257**	0.082	−0.049
Proportion moved	−0.007	0.125*	0.054	0.101*
Proportion from nontraditional families	−0.032	0.273**	0.251**	0.302**
Parameter variance between schools	1.820**	0.202**	0.669**	0.268**
Percentage of variance explained	35.62	57.68	2.69	34.81

Note. The model includes the same set of predictors for initial status. The mean coefficients for the dropout, transfer, and attrition models represent log odds. All predictors are expressed as school effect sizes, which were computed by first converting hierarchical linear modeling (HLM) coefficients to standard units and then dividing by the school-level standard deviation of the dependent variable estimated from the HLM null model. SES = socioeconomic level; GPA = grade point average.
†$p < .10$; *$p < .05$; **$p < .01$.

gible and insignificant effect on student transfer (ES = 0.039). Retention during Grades 1 through 8 had a large effect on student learning (ES = –0.612) and an even larger effect on dropout (ES = 1.020) but a small and insignificant effect on student transfer (ES = 0.129). Changing residences between Grades 10 and 12 had a small yet significant effect on student learning (ES = –0.162), a large effect on dropout (ES = 0.825), and a very large effect on student transfer (ES = 1.627), as one might expect.

After control for student-level predictors, a number of school composition variables also predicted educational outcomes. School mean SES had a small yet significant effect on student learning (ES = 0.253) and dropout (ES = –0.257) but a negligible and insignificant effect on student transfer (ES = 0.082). The percentage of students from nontraditional families had a nonsignificant effect on student learning (ES = –0.032) but small and significant effects on dropout (ES = 0.273) and transfer (ES = 0.251).

These student background and student composition predictors explained a sizeable amount of the variance in educational outcomes among schools, specifically 36% of the variance in student learning[26] and 58% of the variance in dropout rates among schools; however, they explained only 3% of the variance in transfer rates. Most of the explained variance was due to student-level rather than school-level predictors. Because all of the student-level predictors other than residential mobility measured characteristics of students before they entered high schools, these results suggest that high schools have quite a bit to do with how much students learn and whether they remain in their school rather than transferring but relatively little to do with whether they drop out of school. In other words, after control for student inputs, there is still considerable school-level variance to be explained in achievement growth and transfer rates but not dropout rates. These findings have important implications for federal and state policies that hold schools accountable for student learning and dropout rates but not whether they retain all of their students, a topic to which we return in the final section of the article.

The residuals from these models were then used to measure school effectiveness. Specifically, we used the residuals from the estimated models to measure how much better or worse each school was relative to its predicted performance after controlling for differences in student background characteristics and student composition. Descriptive statistics for these residuals are shown in Table 3.

Some residuals showed considerable variation among schools, suggesting that schools vary widely in their effectiveness, while other indicators did not. Schools varied widely, for example, in their effectiveness regarding student learning but varied little in their effectiveness regarding dropping out. This is consistent with the earlier discussion, which showed that student background and compositional characteristics explain much of the variation in dropout rates. Hence, after control for these differences, the proportion of remaining variability in the dropout rate that could potentially be attributed to school influences was less than for the other outcomes. School effectiveness in regard to transfer rates is more complicated: Few schools are very

Table 3
Analysis of Residuals From Student and Social Composition Model

Measure	Achievement growth	Dropout	Transfer	Attrition
Descriptive statistics				
Mean	0.00	0.00	0.00	0.00
Standard deviation	0.58	0.19	0.57	0.42
Maximum	2.73	−0.65	−0.98	−1.05
Minimum	−2.42	0.75	3.78	3.03
Percentage distribution[a]				
Standardized residual ≥ 0.5	19.2	0.2	13.7	8.3
−0.5 < standardized residual < 0.5	62.5	97.7	69.7	80.5
Standardized residual ≤ −0.5	18.3	2.1	16.6	11.2
Total	100.0	100.0	100.0	100.0
Correlations				
Achievement growth	—	−.036	−.013	−.033
Dropout		—	.082**	.689**
Transfer			—	.769**

Note. Estimated residuals from the student and social composition model shown in Table 2. To facilitate comparisons, the signs for the dropout, transfer, and attrition indicators are reversed.
[a]Standardized residuals were computed by dividing the residual by the school-level standard deviation of the dependent variable estimated from the hierarchical linear null model.
**$p < .01$.

effective in reducing transfer rates, but many more are ineffective in that they have much higher than predicted transfer rates.[27]

The variation in school effectiveness can be seen more clearly by classifying schools into three levels of effectiveness. We did so by first computing a standardized residual; that is, we divided the estimated residual for each school by the estimated standard deviation of the outcome based on the unconditional model shown in Table 1. Schools with a standardized residual greater than 0.5 can be considered effective schools, those with a standardized residual below 0.5 can be considered ineffective, and the remainder can be considered average. The resulting distribution, shown in Table 3, suggests that most schools are neither particularly effective nor particularly ineffective. If school outcomes are normally distributed, which in most cases they appear to be, then this finding is not unexpected. As a result, most schools are clustered in the middle and can be considered to be doing a reasonable job. However, some schools are doing better than expected, and some are doing worse. In terms of student learning, about 19% of the schools in our sample were effective in promoting student learning, and another 18% were ineffective. In terms of dropout rates, almost all schools were clustered in the middle category; only a few schools were effective or ineffective. In terms of transfer, about 14% were effective, and another 17% were ineffective. In regard to overall attrition rates, 8% were effective, and another 11% were ineffective.

Are Alternative Measures of School Performance Related?

Next, we examined the correlations among these alternative measures of school performance. These correlations, shown in Table 3, suggest that alternative measures are generally unrelated to each other. For example, school effectiveness as measured by student learning was not related to school effectiveness in terms of dropout, transfer, or overall attrition rates. School effectiveness in regard to dropout rates was slightly and positively correlated with school effectiveness in regard to transfer rates, suggesting that schools that are effective in reducing dropout rates are somewhat effective in reducing transfer rates. Of course, school effectiveness in terms of dropout and transfer was highly correlated with school effectiveness in terms of attrition rates, given that the attrition outcome was the sum of the dropout and transfer outcomes.

The relationships among these alternative measures of school effectiveness are illustrated in the scatterplots presented in Figure 2. The top panel shows the much greater variability in school effectiveness in terms of student learning than in terms of dropout rates. It also shows that there is no definite pattern in the relationship between these two measures. Schools are represented in all four quadrants, meaning that some schools are effective according to both measures of school performance, some schools are ineffective according to both measures, and some schools are effective according to one measure but not the other. The bottom panel shows much greater variability among schools in both student learning and student transfer and shows no consistent pattern in the relationship between these two alternative measures.

What Factors Account for Differences in School Performance?

The next part of our study attempted to explain differences in school performance. We examined three types of factors: structural characteristics of schools, school resources, and school policies and practices. Because structural and resource factors are both considered input characteristics, which school site personnel may have little power to influence, we examined these factors together in a single model. We then investigated process variables related to the internal policies and practices of schools.

Structural and Resource Factors

Estimated parameters from the structural and resource models are shown in Table 4. The results indicate that, after control for other input characteristics, both Catholic (ES = 0.082) and public magnet (ES = 0.075) schools had a negligible, positive effect on student learning relative to comprehensive public schools, while non-Catholic private schools had a small (ES = 0.131) effect. In addition, while Catholic schools had lower dropout rates (ES = –0.275) than comprehensive public schools, they had higher transfer rates (ES = 0.239) and overall attrition rates (ES = 0.210) as well.

School size was also related to school performance, but not in consistent ways. Despite the current interest in small schools, our results do not show that small schools are more effective than mid-sized (600 to 1,200 students)

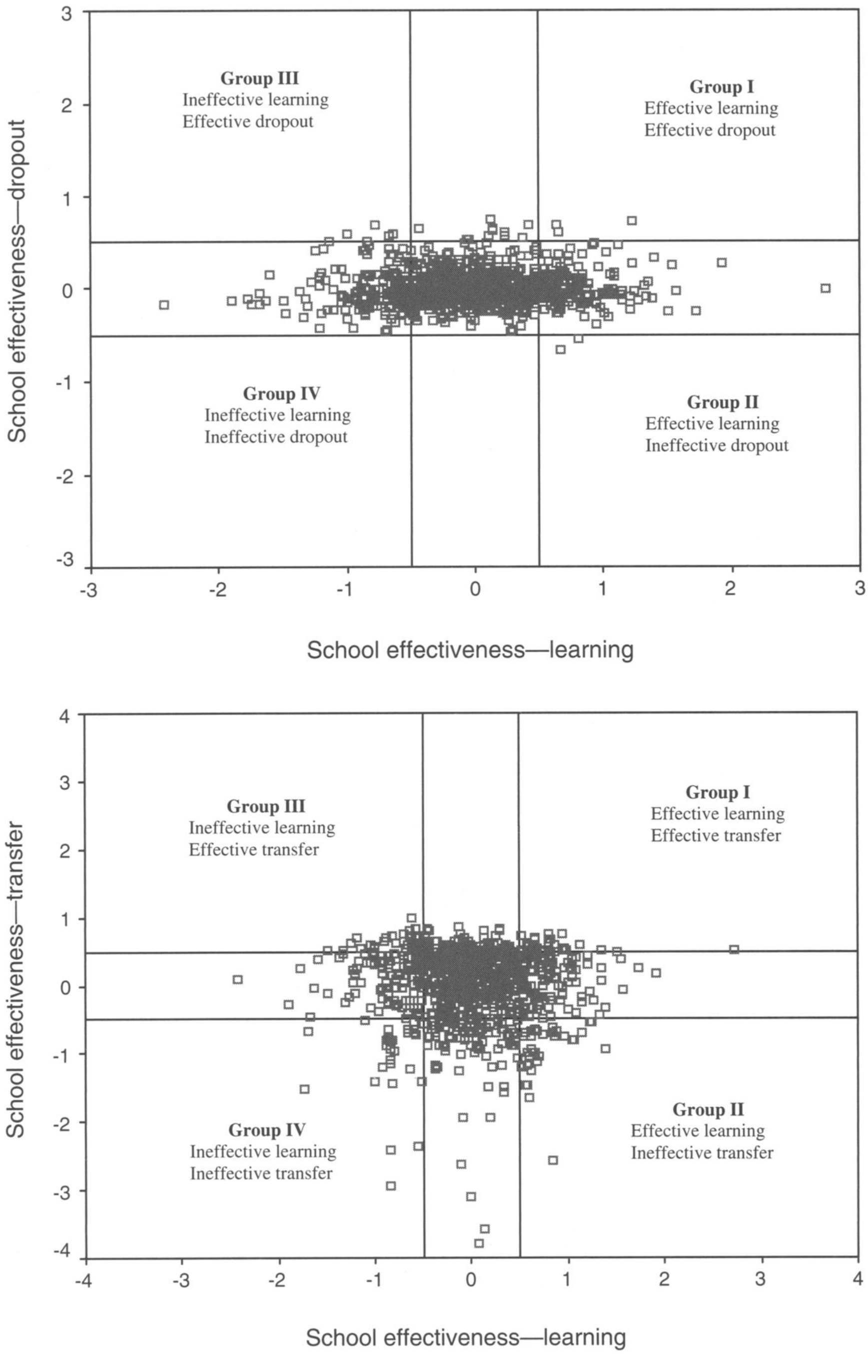

Figure 2. **Scatterplots of alternative measures of school effectiveness (expressed as standardized residuals).**

Table 4
Parameter Estimates: Structure and Resource Models

Measure	Achievement growth	Dropout	Transfer	Attrition
Composition				
Mean SES	0.175**	−0.229*	−0.011	−0.121
Proportion moved	−0.016	0.106†	0.075	0.108*
Proportion from nontraditional families	−0.029	0.245**	0.261**	0.292**
Structure				
Catholic	0.082†	−0.275*	0.239**	0.210**
Private	0.131†	−0.188	0.059	0.080
Magnet	0.075†	−0.024	−0.094†	−0.077†
Small school	0.007	−0.078	0.029	−0.027
Large school	0.110**	0.263**	−0.112†	0.061
Extra-large school	0.085†	0.174*	−0.002	0.027
Urban	0.052	−0.016	0.174**	0.120*
Rural	0.042	−0.012	−0.063	−0.028
Resources				
Mean teacher salary	0.063	−0.132†	−0.139†	−0.164**
Student-teacher ratio	−0.053†	−0.068	0.152*	0.106†
Parameter variance between schools	1.738**	0.127	0.580**	0.241**
Percentage variance between schools	38.52	73.39	15.63	41.46

Note. The model includes the same set of student-level predictors shown in Table 2 and the same set of student-level and school-level predictors for initial status. All predictors are expressed as school effect sizes, which were computed by first converting hierarchical linear modeling (HLM) coefficients to standard units and then dividing by the school-level standard deviation of the dependent variable estimated from the HLM null model. SES = socioeconomic status.
†$p < .10$; *$p < .05$; **$p < .01$.

schools in regard to any of the performance measures we examined. However, large schools were more effective (ES = 0.110) than mid-sized schools in promoting student learning; on the other hand, they had higher dropout rates (ES = 0.263) than mid-sized schools. Similar but smaller effects were found for extra-large schools (more than 1,800 students). These results suggest that the optimal school size for student learning may not be the optimal size for student graduation. They also suggest that one reason large schools may have higher learning rates is their higher dropout rates among students who are generally low achieving.

Our results also suggest that different resource variables have significant effects on different performance measures. Schools with higher teacher salaries had lower dropout rates (ES = −0.132), lower transfer rates (ES = −0.139), and lower attrition rates (ES = −0.164). Schools with higher student-teacher ratios had higher transfer (ES = 0.152) and attrition (ES = 0.106) rates. However, the ESs for all of these variables were small. Interestingly, none of the variables related to teacher credentials had significant effects on any of the performance outcomes.[28]

Altogether, the structural and resource models explained little additional variance in learning relative to the compositional model. However, they did explain additional variance in the dropout, transfer, and attrition models.

Process Factors

The final models introduced a series of variables measuring school policies, practices, and climate. The estimates from the final model, including the significant process variables, are shown in Table 5. As in the previous models, the effects of school factors varied across the outcome measures. Two measures of academic climate—mean time students spend on homework (ES = 0.128) and mean number of advanced courses taken by students (ES = 0.100)—had positive effects on student learning, while the latter measure had a large negative impact on dropout (ES = –0.595) and a moderate impact on attrition rates (ES = –0.290). Two measures of social climate—proportion of students who feel unsafe and student reports of class disruptions—had positive effects on school dropout (ESs = 0.105 and 0.133, respectively) and attrition rates (ESs = 0.104 and 0.152, respectively), and the former also had a negative impact on student learning (ES = –0.109). Schools where more students feel the discipline is fair had lower attrition rates (ES = –0.102).

Two measures of teacher attitudes—efficacy and expectations—also contributed to school effectiveness, but not necessarily as expected. Schools where teachers reported higher efficacy in their ability to teach did not have higher learning rates or lower dropout rates than other schools, but they did have higher transfer (ES = –0.125) and attrition (ES = 0.120) rates—an unexpected result. At the same time, schools where teachers had high expectations for student learning had higher learning rates (ES = 0.102) and lower dropout rates (ES = –0.153) than other schools. Finally, two measures of school organization affected dropout rates but not other measures of school performance: Schools where teachers reported strong principal leadership had higher dropout rates (ES = 0.110), net of other factors, while schools where teachers reported greater control over curriculum and discipline policy had lower dropout rates (ES = –0.156), suggesting that teachers and principals may differ in how they handle students at risk of dropping out.

The final (process) model explained a small amount of additional school-level variance in most outcomes but a sizeable amount in the dropout models (see the bottom portion of Table 5). Altogether, the final model explained 43% of the variance in achievement growth, 99% of the variance in dropout rates, 19% of the variance in transfer rates, and 49% of the variance in attrition rates.

Discussion

The present study examined the relationships among different measures of school performance. In particular, it incorporated measures other than test scores. Virtually all accountability systems rely on the use of test scores to

Table 5
Parameter Estimates for Final Model

Measure	Achievement growth	Dropout	Transfer	Attrition
Composition				
Mean SES	0.009	0.146	0.088	0.133
Proportion moved	0.000	0.080	0.077	0.096†
Proportion from nontraditional families	0.017	0.125†	0.250**	0.229**
Structure				
Catholic	0.044	−0.128	0.227**	0.279**
Private	0.070	−0.091	0.091	0.156*
Small school	0.000	−0.054	0.039	−0.006
Large school	0.124**	0.227**	−0.113†	0.040
Extra-large school	0.105*	0.145*	−0.027	0.025
Magnet	0.061	−0.003	−0.082	−0.058
Urban	0.038	−0.026	0.174**	0.112†
Rural	0.031	0.021	−0.060	−0.007
Resources				
Student-teacher ratio	−0.048	−0.050	0.158*	0.089
Mean teacher salaries	0.046	−0.076	−0.116†	−0.127*
Process				
Mean hours spent on homework	0.128*	0.059	−0.109	−0.003
Mean number of advanced courses	0.100*	−0.595**	0.060	−0.290**
Proportion of students who feel unsafe	−0.109**	0.105†	0.069	0.104†
Class disruptions	−0.051	0.133†	0.123	0.152*
Fair discipline	0.012	−0.003	−0.102	−0.102†
Teacher efficacy	0.029	0.077	0.125†	0.120**
Teacher expectations	0.102*	−0.153†	−0.043	−0.084
Principal leadership	−0.048	0.110*	0.009	0.073
Teacher control	−0.051	−0.156*	0.018	−0.089
Parameter variance between schools	1.619**	0.007**	0.558**	0.208**
Percentage variance between schools	42.73	98.53	18.83	49.48

Note. The model includes the same set of student-level predictors shown in Table 2 and the same set of student-level and school-level predictors for initial status. All predictors are expressed as school effect sizes, which were computed by first converting hierarchical linear modeling (HLM) coefficients to standard units and then dividing by the school-level standard deviation of the dependent variable estimated from the HLM null model. SES = socioeconomic status.
†$p < .10$; *$p < .05$; **$p < .01$.

evaluate the performance of schools. Yet, such a singular focus may ignore other important outcomes of schooling.[29] Moreover, policies designed to improve test score performance could hurt performance in other areas. This is particularly true at the high school level, where ensuring that students graduate may be at least as important as ensuring that students improve their test scores.[30]

The results of this study generally support the notion of an alternative view of school effectiveness: Schools that are effective in promoting student learning (growth in achievement) are not necessarily effective in reducing dropout and transfer rates. Even though conventional notions of school effectiveness suggest that characteristics of effective schools—such as committed staff and competent leaders—should improve all student outcomes (Purkey & Smith, 1985), existing theories of dropout suggest that different factors influence student engagement and dropout, in part because they emphasize the social as well as the academic side of schooling (Finn, 1989; Wehlage et al., 1989). Moreover, institutional perspectives and empirical studies have shown that schools may improve learning outcomes by discharging the lowest performing students (Bowditch, 1993; Riehl, 1999). In fact, recent state and federal pressure for school accountability, including rigorous high school exit examinations, has prompted increased accounts of schools actively discharging low-performing students (e.g., Howard, 2004; Lewin & Medina, 2003). Our finding that large schools not only have higher dropout rates, but also higher learning rates among the students who remain, may be a reflection of such activities.

Several other important findings emerged from this study. First, we found that schools have relatively small effects on student learning in comparison with student background characteristics, which supports one of the original conclusions of the landmark Coleman report (Coleman et al., 1966). In the present study, about 26% of the overall variability in student learning was attributable to students' high schools.[31] Nonetheless, students who attend highly effective (i.e., 95th percentile) schools can still expect to learn twice as much as students who attend highly ineffective (5th percentile) schools.[32] At the same time, even if differences among schools were completely eliminated, there would still be considerable variability in achievement growth among students. This provides an upper limit on how much inequality in student learning could be expected to be eliminated through school reforms alone as opposed to reforms in social policy that address inequalities in student and family circumstances that contribute to student learning.

Second, we found that differences in the background characteristics of students account for more than half of the variability in school dropout rates, but only about a third of the variability in student learning and attrition rates and almost none of the variability in transfer rates. One interpretation of this finding is that schools have relatively little impact on dropout rates relative to other student outcomes. In a recent study involving the use of the same data set and focusing on high school dropout rates in 499 local labor markets, the authors reached such a conclusion about school effects on dropout rates: "After controlling for individual-level socio-demographic and academic variables, we found that dropout rates do not vary significantly across labor markets" (Warren & Lee, 2003, p. 124). Yet, in our study, we still found significant variation in school dropout rates after controlling for student background characteristics and the social composition of schools. We also found that school policies and practices accounted for almost 25% of the remaining variability in school dropout rates, far more than any other outcome. Thus, while there is little variability across schools in dropout rates, schools seem to have considerable control in regard to improving these rates.

At the same time, little policy attention has been focused on the phenomenon of student transfer, even though there is more variability among schools in transfer rates than in dropout rates. Yet, while there is considerable variability in transfer rates among schools, we were unable to identify features of schools that could explain much of that variability. Similarly, we were unable to identify features of schools that could explain much of the remaining variability in student learning and transfer rates.

Third, we found that different school factors were associated with different measures of school performance. Only two factors had significant effects on all four performance measures, and both represented aspects of student composition: proportion of students from nontraditional families and mean achievement of students in the eighth grade. The effects of other factors varied across the performance outcomes. For example, higher teacher salaries were associated with lower dropout, transfer, and attrition rates but were not associated with high achievement growth. In general, resource variables had stronger effects on transfer rates than achievement growth or dropout rates, which may suggest that such factors influence students' and parents' decisions to change schools even if they show no apparent effects on student learning. Parental involvement had a positive effect on achievement growth, a negative effect on dropout rates, and no effect on transfer and attrition rates. Class disruptions and students' perceptions of fair discipline were associated with lower transfer and attrition rates but had no significant effects on achievement growth or dropout rates.

Some of our findings challenge widespread notions about effective schools. Although a variety of evidence from the 1980s and early 1990s showed that Catholic schools are more effective than public schools in raising student achievement and reducing dropout rates (e.g., Bryk et al., 1993; Coleman & Hoffer, 1987; Coleman et al., 1982; Evans & Schwab, 1995; Neal, 1997; Sander & Krautman, 1995), we found that Catholic schools had negligible effects on student learning after controlling for the effects of student background characteristics and school resources. The finding that Catholic schools are no more effective in promoting learning than public comprehensive schools has been reported in other recent studies involving the same NELS data and subject-specific measures of student learning (e.g., Gamoran, 1996; Lee & Smith, 1997). We did find that Catholic schools had significantly lower dropout rates, but they had higher transfer and attrition rates as well. This suggests that Catholic schools are less effective in retaining the students who initially chose to attend them.

The results of this study also challenge some widely held beliefs about school size. Despite the recent interest in small schools (Gewertz, 2000), we found that small high schools were neither more effective in promoting student achievement nor more effective in reducing dropout and transfer rates than mid-sized high schools. In fact, larger high schools—those with more than 1,200 students—had significantly higher levels of student achievement growth than mid-sized schools after control for differences in student background characteristics, although the magnitude of the difference was small. However, larger schools also had higher dropout and transfer rates, a finding confirmed in another recent study based on a smaller sample of high schools (Lee & Burkam, 2003). This suggests that while larger schools are more effective in promoting learning among students who are more stable and generally less at risk, they are less effective than mid-sized schools in retaining students who are more at risk of dropping out or transferring.[33]

The results of this study have several implications for education policy. One is that school accountability systems that rely solely on test scores provide an incomplete indication of school performance. High schools that are effective in promoting student learning are not necessarily more effective in reducing dropout rates or attrition rates. In fact, high schools have relatively little influence on dropout rates in comparison with other student outcomes. Furthermore, even evaluating high schools on the basis of both achievement growth and dropout rates may be insufficient, because transfer or mobility rates are ignored. Although mobility rates are not generally considered relevant indicators of school performance, we found that they are less influenced by student and family background factors, including residential mobility, than dropout rates. This suggests that schools should be accountable for all of the students who enter, not only the ones who remain.

Another implication is that policies designed to promote school performance in one area may not promote performance or can even lead to worse performance in another area. This certainly appears to be the case with school size. We found that larger schools are more effective in promoting student learning than mid-sized or smaller schools, but they are also less likely to keep students enrolled. The policy implication is that one size does not fit all—more stable students may benefit from attending larger schools, while more at-risk students may benefit from attending mid-sized schools.

Finally, the results of this study have implications for school choice policies. To the extent that support for school choice is based on the belief that private schools are more effective than public schools (Levin, 1998), the present study does not support school choice. Although Catholic schools (but not other private schools) had significantly lower dropout rates than public schools, they also had significantly higher transfer and attrition rates.

In short, the results of the present study suggest that test scores alone are insufficient for measuring school performance. Failure to use other complementary measures of school performance in addition to test scores could lead to erroneous conclusions about which schools are effective and what characteristics promote school effectiveness.

Notes

An earlier version of this article was presented at the annual meeting of the American Educational Research Association, New Orleans, April 2002. We would like to thank the editor and four anonymous reviewers for their helpful comments and suggestions.

[1]Although these two measures are related, their computation requires different types of data (see Kaufman, in press). According to the most recent school quality indicators compiled by *Education Week,* 49 states were using data other than test scores to evaluate schools in November 2003, although the specific type of information (e.g., graduation rates) was not indicated (Quality Counts, 2004).

[2]NCLB allows states considerable discretion over what specific measure of high school completion is used and how they meet their annual performance objectives (see Orfield, Losen, Wald, & Swanson, 2004).

[3]Some exceptions are Bryk and Thum (1989), who examined attendance and dropout; Crain, Mahard, and Narot (1982), who examined test scores and 16 attitude variables; Lee and Smith (1995), who examined engagement and test scores; and Phillips (1997), who examined attendance and test scores.

[4]Some researchers have referred to them as "type B effects" since, when statistical adjustments are made for the effects of school inputs (type A effects), they provide a better and more appropriate basis for comparing the performance of schools (Raudenbush & Willms, 1995; Willms, 1992; Willms & Raudenbush, 1989).

[5]Research has also demonstrated that teacher efficacy is shaped by school characteristics, including the characteristics of the student body (Lee, Dedrick, & Smith, 1991; Newmann, Rutter, & Smith, 1989).

[6]NCES has begun a new national longitudinal survey of American high schools and their students known as the Education Longitudinal Study of 2002 (see http://nces.ed.gov/surveys/els2002/). However, this study is following a cohort of 10th-grade students, so it will not be as useful as NELS in examining student achievement over the entire 4 years of high school.

[7]We also deleted one additional school that did not have any students who were still enrolled in 1992, a requirement for estimating the hierarchical multinomial regression model used in the present analysis.

[8]Requiring 2 years of test data instead of 3 not only increased the overall sample size and the within-school sample size (mean of 16 students per school) for our study but also made the sample more representative, since students who transferred or dropped out of 10th grade were less likely to be tested in 12th grade (see Morgan & Sorensen, 1999, Table 1) and more likely to have minority backgrounds and to come from low-income households. Approximately one quarter of our sample had only two test scores.

[9]In two-level linear models, HLM allows weighting at both levels. In three-level linear models, HLM allows weighting only at Level 1. In two-level nonlinear models, weights are not allowed. Our two-level models were nonlinear. In our three-level growth model, Level 1 was within-person growth, while Level 2 was the individual level of analysis. Therefore, it would not have been appropriate to include individual weights at Level 1 of our three-level growth model.

[10]The student sample appeared to be representative of high school students in 1990 on the basis of the sample means of several variables (see Appendix A). First, the mean SES value for the students in the sample was 0.01, almost the same as the normalized value of zero when the composite measure was created via a weighted sample of all 10th-grade students (Ingles et al., 1994). Similarly, the mean 10th-grade composite test score in the sample was 50.1 (not shown), almost the same as the normalized value of 50.0 when the variable was created with a weighted sample of all 10th-grade students. Finally, the dropout rate was 6.8%, similar to the weighted national estimate of 5.6% for 10th-grade students using the same data (U.S. Department of Education, 1999, Table B9). On the basis of comparisons between the school sample (see Appendix A) and characteristics of all public and private schools in the United States from the 1987–1988 Schools and Staffing Survey (the one conducted closest to the time of NELS), the school sample also appeared to be representative. First, 15% of the sample schools were private, similar to the 11% national figure (U.S. Department of Education, 1990, Table 1). Second, 31% of the schools were located in urban areas and 30% in rural areas, as compared with corresponding national percentages of 32% and

35% (U.S. Department of Education, 1996, Table 1.1). Third, the mean student-teacher ratio was 16.4, similar to the national mean of 17.1 (Table 1.2). Finally, the mean proportion of minority students was .24, similar to the national proportion of .28 (Table 1.3).

[11]It can also be conceptualized at more than two levels, including classrooms and districts. See Willms (1992) for a more detailed discussion.

[12]Carbonaro (1998) and Swanson and Schneider (1999) examined predictors of achievement growth and dropout at the student level.

[13]In previous work involving the same data, we estimated a second-order confirmatory factor analysis measurement model wherein achievement factors were estimated from the four academic subject tests at each of the three time points (Palardy, 2003). The results showed that the factor loadings for each academic subject were highly similar. As a consequence, the latent variables had nearly equivalent values as the equally weighted composite scores used in this study.

[14]For example, California constructs an academic performance index for every school that is a weighted composite of test scores in a number of subject areas (see http://www.cde.ca.gov/ta/ac/ap/apidescription.asp).

[15]Many of the school-level variables were based on mean values of the within-school samples of students (mean $N = 16$) and their teachers. Although the mean within-school sample size of 16 is sufficient to compute school mean characteristics, we do not know representative each sample is. However, because the overall student sample and school sample appear to be representative of the overall population of students and schools in the United States, as we point out in Footnote 9, we believe that the within-school samples are representative as well. The teacher data came from the first of the two teachers per student who were surveyed in 1990, following the lead of Lee et al. (1997). These data, too, may not be representative of all teachers in the school but are probably representative of the teachers who taught the sampled students.

[16]Two earlier studies showed that schools with high minority enrollments had lower student performance (Bryk & Thum, 1989; Rumberger, 1995).

[17]The academic background variables were constructed to account for three different peer effects of students: academic achievement, achievement motivation, and social behavior (see Kahlenberg, 2001, pp. 48–58).

[18]We coded size as a series of dummy variables; 601–1,200 students represented the omitted category, because recent research conducted by Lee and Smith (1997) showed that this school size produced the highest achievement gains in math and reading.

[19]These measures were informed, in part, by the work of Lee and Smith (1999), Phillips (1997), and Rumberger and Thomas (2000). See Appendix B for a list of the specific variables used to construct these and other composite measures.

[20]Originally we constructed a single measure, teachers' responsibility for learning, based on the work of Lee et al. (1997). However, on the basis of a recent review of the theoretical and empirical literature on teacher efficacy (Tschannen-Moran, Hoy, & Hoy, 1998) and an exploratory factor analysis, we instead created two separate variables, one that measured teachers' efficacy in regard to their ability to teach (sometimes referred to as specific efficacy) and one that measured teachers' beliefs in students' ability to learn (sometimes referred to as general efficacy or locus of control). We treated these variables as process (i.e., endogenous) variables because, as pointed out in Footnote 5, research has demonstrated that they are shaped by school characteristics.

[21]These measures were based on the work of Lee et al. (1991), Miller and Rowan (2003), Newmann et al. (1989), and Rowan, Raudenbush, and Kang (1991).

[22]This measure was based on the work of Morgan and Sorensen (1999).

[23]As a means of producing accurate parameter estimates, all predictor variables were used to estimate both initial achievement (intercept) and achievement growth (slope). Only the estimates for the growth terms are reported here.

[24]Controlling for student background characteristics at the school level captures both the individual effects and the contextual effects of student background. Although students and parents may choose a school because of the positive effects of school context, both types of effects are generally beyond the control of most (public) schools (Raudenbush & Willms, 1995).

[25]Predictor variables of continuous measures were first multiplied by their own standard deviation and then divided by the school-level standard deviation of the outcome measure from the unconditional model. Predictor variables for dichotomous measures were simply divided by the school-level standard deviation. Because our models included both student-level predictors and school-level predictors, it might seem more appropriate to convert student-level predictors to between-student effect sizes by dividing by the student-level standard deviation. However, the nonlinear HLM program used in this analysis does not estimate student-level standard deviations. Thus, to facilitate comparisons among predictors, we elected to use school-level standard deviations as common metrics and refer to them as school effect sizes. This was the same procedure used by Lee and Smith (1997) and Lee et al. (1997).

[26]Not surprisingly, 81% of the variation in initial student achievement was explained by student background characteristics, which suggests a robust student background model.

[27]In other words, the residuals for the transfer model were not normally distributed. According to Raudenbush and Bryk (2002, p. 274), the failure of the normality assumption does not bias estimates of fixed effects but can affect hypothesis testing. On the basis of the recommendations of Raudenbush and Bryk (2002, p. 276), we used robust standard errors to determine significance levels in the nonlinear models.

[28]There are two possible reasons for our lack of significant findings for measures of teacher characteristics: (a) Our within-school samples of students and teachers were small, and (b) we aggregated test scores and teacher characteristics across subject areas. Existing research suggests that the impact of teacher credentials differs across high school subject areas (Darling-Hammond et al., 2001; Goldhaber & Brewer, 2000, 2001).

[29]Cuban (2003), for example, argued that the fundamental purpose of public schools is to "transform children into civic-minded, independent-thinking, and socially responsible adults committed to both the common good and engaged in productive work" (p. 6). Neither this study nor any accountability system of which we are aware has measured such outcomes.

[30]These two goals are becoming increasingly related as states institute high school exit examinations.

[31]Raudenbush and Bryk (2002) found that 58% of the variance in student learning during Grades 1 to 3 was between schools (p. 240), suggesting that elementary schools may have more of an effect than high schools on student learning.

[32]Interestingly, Jencks and Brown (1975) reached a similar conclusion in their study of 98 Project Talent high schools in the early 1960s (p. 292). They too pointed out, however, that schools have relatively little effect on overall differences in student achievement.

[33]While Lee and her colleagues (Lee & Burkam, 2003; Lee & Smith, 1999) found that mid-sized schools had higher achievement and lower dropout rates than either larger or small schools, we did not. Their studies were based on two different samples of students and schools, while our study was based on a single sample. In particular, their study showing higher achievement in mid-sized schools included only students who remained in the same school in Grades 10 and 12 (excluding dropouts and transfers) and examined only achievement in mathematics and reading. The current study, although based on the same NELS data set, involved a larger sample of students and schools (14,199 students in 912 schools vs. 9,812 students in 789 schools), a more representative sample (including students who dropped out or transferred after 10th grade), and an achievement outcome consisting of four academic subject area tests (which other studies have shown are differentially related to school factors; see, e.g., Gamoran, 1996). To what extent these differences contributed to the differences in the findings remains unclear.

References

Alexander, K. L., & Pallas, A. (1985). School sector and cognitive performance: When is a little a little? *Sociology of Education, 58,* 115–128.

Astone, N. M., & McLanahan, S. S. (1994). Family structure, residential mobility, and school dropout: A research note. *Demography, 31,* 575–584.

Balfanz, R., & Legters, N. (in press). How many central city high schools have a severe dropout problem, where are they located, and who attends them? Initial estimates using the Common Core of Data. In G. Orfield (Ed.), *Dropouts in America: Confronting the crisis.* Cambridge, MA: Harvard Educational Publishing.

Bowditch, C. (1993). Getting rid of troublemakers: High school disciplinary procedures and the production of dropouts. *Social Problems, 40,* 493–509.

Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good.* Cambridge, MA: Harvard University Press.

Bryk, A. S., & Thum, Y. M. (1989). The effects of high school organization on dropping out: An exploratory investigation. *American Educational Research Journal, 26,* 353–383.

Carbonaro, W. J. (1998). A little help from my friend's parents: Intergenerational closure and educational outcomes. *Sociology of Education, 71,* 295–313.

Carbonaro, W. J., & Gamoran, A. (2002). The production of achievement inequality in high school English. *American Educational Research Journal, 39,* 801–827.

Carroll, D. (1996). *National Education Longitudinal Study (NELS:88/94): Methodology report.* Washington, DC: U.S. Government Printing Office.

Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, and America's schools.* Washington, DC: Brookings Institution.

Coleman, J. S. (1990). *Equality and achievement in education.* San Francisco: Westview Press.

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, F., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Government Printing Office.

Coleman, J. S., & Hoffer, T. (1987). *Public and private high schools: The impact of communities.* New York: Basic Books.

Coleman, J. S., Hoffer, T., & Kilgore, S. B. (1982). *High school achievement: Public, Catholic, and private schools compared.* New York: Basic Books.

Crain, R. L., Mahard, R. E., & Narot, R. E. (1982). *Making desegregation work: How schools create social climates.* Cambridge, MA: Ballinger.

Croninger, R. G., & Lee, V. E. (2001). Social capital and dropping out of high school: Benefits to at-risk students of teachers' support and guidance. *Teachers College Record, 103,* 548–581.

Cuban, L. (2003). *Why is it so hard to get good schools?* New York: Teachers College Press.

Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis, 23,* 57–77.

Evans, W. N., & Schwab, R. M. (1995). Finishing high school and starting college: Do Catholic schools make a difference? *Quarterly Journal of Economics, 110,* 941–974.

Fine, M. (1991). *Framing dropouts: Notes on the politics of an urban public high school.* Albany: State University of New York Press.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59,* 117–142.

Frank, K. A. (1998). Quantitative methods for studying social context in multilevels and through interpersonal relations. In P. D. Pearson & A. Iran-Nejad (Eds.), *Review of research in education* (Vol. 23, pp. 171–216). Washington, DC: American Educational Research Association.

Freiberg, H. J. (Ed.). (1999). *School climate: Measuring, improving, and sustaining healthy learning environments.* Philadelphia: Falmer Press.

Gamoran, A. (1992). Social factors in education. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (pp. 1222–1229). New York: Macmillan.

Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. *Educational Evaluation and Policy Analysis, 18,* 1–18.

Gewertz, C. (2000, September 13). Gates Foundation awards $56 million for small schools. *Education Week,* p. 14.

Glasman, N. S., & Biniaminov, I. (1981). Input-output analyses of schools. *Review of Educational Research, 51,* 509–539.

Goldhaber, D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22,* 129–145.

Goldhaber, D. D., & Brewer, D. J. (2001). Evaluating the evidence on teacher certification: A rejoinder. *Educational Evaluation and Policy Analysis, 23,* 79–86.

Goldschmidt, P., & Wang, J. (1999). When can schools affect dropout behavior? A longitudinal multilevel analysis. *American Educational Research Journal, 36,* 715–738.

Hannaway, J., & Carnoy, M. (Eds.). (1993). *Decentralization and school improvement: Can we fulfill the promise?* San Francisco: Jossey-Bass.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24,* 1141–1177.

Hanushek, E. A. (1989). The impact of differential expenditures on school performance. *Educational Researcher, 18*(4), 45–62.

Hanushek, E. A. (1994). Production functions in education. In T. Husen & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (pp. 4059–4069). New York: Pergamon Press.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19,* 141–164.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher, 23*(3), 5–14.

Howard, J. (Executive Producer). (2004, August 25). *60 minutes* [television broadcast]. New York: CBS News.

Ingles, S. J., Dowd, K. L., Baldridge, J. D., Stipe, J. L., Bartot, V. H., & Frankel, M. R. (1994). *National Education Longitudinal Study of 1988, second follow-up: Student component data file user's manual.* Washington, DC: U.S. Department of Education.

Jencks, C. S., & Brown, M. D. (1975). Effects of high schools on their students. *Harvard Educational Review, 45,* 273–324.

Jencks, C., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. In L. Lynn Jr., & M. G. H. McGeary (Eds.), *Inner-city poverty in the United States* (pp. 111–186.). Washington, DC: National Academy Press.

Johnson, M. K., Crosnoe, R., & Elder, G. H., Jr. (2001). Students' attachment and academic engagement: The role of race and ethnicity. *Sociology of Education, 74,* 318–340.

Kahlenberg, R. D. (2001). *All together now: Creating middle-class schools through public school choice.* Washington, DC: Brookings Institution.

Kaufman, P. (in press). The national dropout data collection system: Assessing consistency. In G. Orfield (Ed.), *Dropouts in America: Confronting the crisis.* Cambridge, MA: Harvard Educational Publishing.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education, 62,* 172–192.

Lee, V. E., & Burkam, D. T. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal, 40,* 353–393.

Lee, V. E., Dedrick, R. F., & Smith, J. B. (1991). The effect of social organization of schools on teachers' efficacy and satisfaction. *Sociology of Education, 64,* 190–208.

Lee, V. E., & Smith, J. B. (1993). Effects of school restructuring on the achievement and engagement of middle-grade students. *Sociology of Education, 66,* 164–187.

Lee, V. E., & Smith, J. B. (1995). Effects of high school restructuring and size on gains in achievement and engagement for early secondary school students. *Sociology of Education, 68,* 241–279.

Lee, V. E., & Smith, J. B. (1997). High school size: Which works best for whom? *Educational Evaluation and Policy Analysis, 19,* 205–227.

Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. *American Educational Research Journal, 36,* 907–945.

Lee, V. E., Smith, J. B., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education, 70,* 128–150.

Levin, H. M. (1997). Raising school productivity: An x-efficiency approach. *Economics of Education Review, 16,* 303–311.

Levin, H. M. (1998). Educational vouchers: Effectiveness, choice, and costs. *Journal of Policy Analysis and Management, 17,* 373–392.

Lewin, T., & Medina, J. (2003, July 31). To cut failure rate, schools shed students. *New York Times,* p. A1.

Ma, X. (2001). Stability of school academic performance across subject areas. *Journal of Educational Measurement, 38,* 1–18.

Mayer, S. (1991). How much does a high school's racial and socioeconomic mix affect graduation and teenage fertility rates? In C. Jencks & P. Peterson (Eds.), *The urban underclass* (pp. 321–341). Washington, DC: Brookings Institution.

McNeal, R. B. (1997). High school dropouts: A closer examination of school effects. *Social Science Quarterly, 78,* 209–222.

Miller, R. J., & Rowan, B. (2003). Sources and consequences of organic management in elementary and secondary schools. In W. Hoy & C. Miskel (Eds.), *Studies in leading and organizing schools* (pp. 51–89). Greenwich, CT: Information Age.

Morgan, S. L., & Sorensen, A. B. (1999). Parental networks, social closure, and mathematics learning: A test of Coleman's social capital explanation of school effects. *American Sociological Review, 64,* 661–681.

Neal, D. (1997). The effects of Catholic secondary schooling on educational achievement. *Journal of Labor Economics, 15,* 98–123.

Newmann, F. M., Rutter, R. A., & Smith, M. S. (1989). Organizational factors that affect school sense of efficacy, community, and expectations. *Sociology of Education, 62,* 221–238.

Orfield, G., Losen, D., Wald, J., & Swanson, C. (2004). *Losing our future: How minority youth are being left behind by the graduation rate crisis.* Cambridge, MA: Harvard University.

Palardy, G. J. (2003). *A comparison of hierarchical linear and structural equation growth models and their application in school effectiveness research.* Unpublished doctoral dissertation, University of California, Santa Barbara.

Phillips, M. (1997). What makes schools effective? A comparison of the relationships of communitarian climate and academic climate to mathematics achievement and attendance during middle school. *American Educational Research Journal, 34,* 633–662.

Purkey, S. C., & Smith, M. S. (1985). School reform: The district policy implications of the effective schools literature. *Elementary School Journal, 85,* 354–389.

Quality Counts 2004. (2004, January 8). *Education Week.* Retrieved August 30, 2004, from http://www.edweek.org/sreports/qc04

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics, 20,* 307–335.

Riehl, C. (1999). Labeling and letting go: An organizational analysis of how high school students are discharged as dropouts. In A. M. Pallas (Ed), *Research in sociology of education and socialization* (pp. 231–268). New York: JAI Press.

Romo, H. D., & Falbo, T. (1996). *Latino high school graduation: Defying the odds.* Austin: University of Texas Press.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools. *Teachers College Record, 104,* 1525–1567.

Rowan, B., Raudenbush, S. W., & Kang, S. J. (1991). Organizational design in high schools: A multilevel analysis. *American Journal of Education, 99,* 238–266.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal, 32,* 583–625.

Rumberger, R. W., & Larson, K. A. (1998). Student mobility and the increased risk of high school dropout. *American Journal of Education, 107,* 1–35.

Rumberger, R. W., & Palardy, G. J. (in press). Does segregation (still) matter? The impact of student composition on academic achievement in high school. *Teachers College Record.*

Rumberger, R. W., & Thomas, S. L. (2000). The distribution of dropout and turnover rates among urban and suburban high schools. *Sociology of Education, 73,* 39–67.

Sander, W., & Krautmann, A. C. (1995). Catholic schools, dropout rates and educational attainment. *Economic Inquiry, 33,* 217–233.

Shavelson, R., McDonnell, L., Oakes, J., & Carey, N. (1987). *Indicator systems for monitoring mathematics and science education.* Santa Monica, CA: RAND.

Sui-Chu, E. H., & Willms, J. D. (1996). Effects of parental involvement on eighth-grade achievement. *Sociology of Education, 69,* 126–141.

Swanson, C. B., & Schneider, B. (1999). Students on the move: Residential and educational mobility in America's schools. *Sociology of Education, 72,* 54–67.

Tagiuri, R. (1968). The concept of organizational climate. In R. Taguiri & G. H. Litwin (Eds.), *Organizational climate: Exploration of a concept* (pp. 1–32). Boston: Graduate School of Business Administration, Harvard University.

Teachman, J. D., Paasch, K., & Carver, K. (1996). School capital and dropping out of school. *Journal of Marriage and the Family, 58,* 773–783.

Teddlie, C., & Reynolds, D. (Eds.). (2000). *The international handbook of school effectiveness research.* New York: Falmer Press.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68,* 202–248.

U.S. Department of Education. (1990). *Schools and staffing in the United States: Comparisons of public and private schools, 1987–88.* Washington, DC: Author.

U.S. Department of Education. (1996). *Schools and staffing in the United States: A statistical profile, 1993–94.* Washington, DC: Author.

U.S. Department of Education. (1999). *Dropout rates in the United States: 1997.* Washington, DC: Author.

U.S. Department of Education. (2003). *No Child Left Behind.* Retrieved June 23, 2003, from http://www.nochildleftbehind.gov/next/overview/index.html

Warren, J. R., & Lee, J. C. (2003). The impact of adolescent employment on high school dropout: Differences by individual and labor-market characteristics. *Social Science Research, 32,* 98–128.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73,* 89–122.

Wehlage, G. G., Rutter, R. A., Smith, G. A., Lesko, N., & Fernandez, R. R. (1989). *Reducing the risk: Schools as communities of support.* New York: Falmer Press.

Willms, J. D. (1985). Catholic-school effects on academic achievement: New evidence from the High School and Beyond follow-up study. *Sociology of Education, 59,* 98–114.

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

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

Witte, J. F. (1992). Private school versus public school achievement: Are there findings that should affect the educational choice debate? *Economics of Education Review, 11,* 371–394.

Witte, J. F., & Walsh, D. J. (1990). A systematic test of the effective schools models. *Educational Evaluation and Policy Analysis, 12,* 188–212.

Manuscript received June 14, 2004
Accepted August 10, 2004

APPENDIX A

Variable Descriptive Statistics and Labels

Variable name	M	SD	Min.	Max.	Description (NELS variables)
Measurement variables (Level 1 of learning outcome; N = 39,241)					
Test composite	50.07	9.27	24.86	78.73	Mean of math, reading, scence,and history test scores
Time	0.46	0.40	0.00	1.00	Time (0 = 8th; .5 = 10th; 1 = 12th)
Student variables (Level 2 of learning outcome, Level 1 of categorical outcomes; N = 14,199)					
Dropped out in Grades 10–12	0.07	0.25	0.00	1.00	Dropped out at any time (F2DOSTAT = 3 or 5)
Transferred in Grades 10–12	0.06	0.24	0.00	1.00	Transferred schools between 10th and 12th grades (F2F1SCFG = 1)
Left in Grades 10–12	0.13	0.34	0.00	1.00	Transferred or dropped out between 10th and 12th grades
Asian	0.06	0.24	0.00	1.00	(F1RACE = 1)
Black	0.09	0.29	0.00	1.00	(F1RACE = 2)
Hispanic	0.11	0.31	0.00	1.00	(F1RACE = 3)
Native American	0.01	0.10	0.00	1.00	(F1RACE = 5)
SES	0.04	0.81	−2.95	2.76	Composite of family income, parents' educational and occupational prestige (F1SES)
Nontraditional family	0.33	0.47	0.00	1.00	Does not live with both birth parents (F1S92A or F1S92D ≠ 1)
GPA in Grades 6–8	3.01	0.72	0.50	4.00	GPA composite (BYGRAD)
Misbehavior in Grade 8	−0.11	0.90	−0.65	3.51	Factor score (BYS55A, BYS55E, BYS55F)
College aspirations in Grade 8	0.71	0.46	0.00	1.00	Planned to finish college (BYS45 = 5 or 6)
Retained in Grades 1–8	0.16	0.37	0.00	1.00	Ever held back a grade in school as reported by student (BYS74 = 2) or parent (BYP44 = 1)
Changed schools in Grades 1–8	1.02	1.40	0.00	5.00	Number of times student changed schools between the 1st and 8th grades (BYP40)
Moved during Grades 10–12	0.15	0.35	0.00	1.00	Family moved to a new home in past 2 years (F2S96A = 1 or F2D80A = 1)
School variables (Level 3 of learning outcome, Level 2 of categorical outcomes; N = 912)					
Composition					
Mean SES	0.01	0.52	−1.33	1.54	Mean SES of students (F1SES)
Proportion minority	0.24	0.29	0.00	1.00	Proportion of Black and Hispanic students (mean of F1RACE = 2 or 3)

(continued)

APPENDIX A *(Continued)*

Variable Descriptive Statistics and Labels

Variable name	*M*	*SD*	Min.	Max.	Description (NELS variables)
Proportion from nontraditional	0.34	0.17	0.00	1.00	Proportion of students who do not live in household with birth mother and father (F1S92A = 2 and/or F1S92A = 2)
Proportion moved	0.15	0.10	0.00	0.63	Proportion of students who changed residences
Mean 8th-grade misbehavior	−0.08	−0.65	1.19	0.31	Mean factor score (BYS55A, BYS55E, BYS55F)
Mean GPA	2.98	0.30	1.97	3.79	Mean GPA Grades 6–8 (BYGRADS)
Proportion retained Grades 1–8	0.16	0.12	0.00	0.71	Proportion of students retained in Grades 1–8 reported by parent or student (BYS74, BYP44)
Proportion college aspirations	0.64	0.18	0.00	1.00	Proportion of students who plan to finish college (BYS45 = 5 or 6)
Structure					
Catholic	0.07	0.25	0.00	1.00	(G10CTRL1 = 2)
Private	0.08	0.27	0.00	1.00	(G10CTRL1 = 3–5)
Magnet	0.05	0.22	0.05	0.22	Public magnet school (F1C4AB = 1)
Small	0.23	0.42	0.00	1.00	(F1C2 = 1–600)
Large	0.28	0.45	0.00	1.00	(F1C2 = 1,201–1,800)
Extra large	0.13	0.33	0.00	1.00	(F1C2 = 1,801+)
Urban	0.31	0.46	0.00	1.00	School located in urban setting (F1URBAN = 1)
Rural	0.30	0.46	0.00	1.00	School located in rural setting (F1URBAN = 2)
Resources					
Student-teacher ratio	16.39	5.71	5.26	69.23	(F1SCENRL/F1C35)
Mean salary	29.18	4.79	1.60	4.15	Mean teacher salary (F1C42A + F1C42B)/2,000
Proportion of teachers with advanced degrees	0.54	0.21	0.03	2.00	Proportion of teachers with advanced degrees [(F1C44C + F1C44D)/F1C35]
Subject certified	0.85	0.27	0.00	1.00	Proportion of teachers certified to teach in their teaching subject area (depending on their teach ing area, one of F1T3_8A-D = 1)
Teacher experience	0.84	0.24	0.00	1.00	Proportion of teachers with 4 or more years of secondary-level teaching experience (F1T3_4B = 2–9)
Standard credential	0.83	0.28	0.00	1.00	Proportion of teachers with a standard teaching credential (F1T3_7 = 1)

Variable name	*M*	*SD*	Min.	Max.	Description (NELS variables)
BA in teaching area	0.39	0.19	0.00	1.00	Proportion of teachers with at least a BA in their primary teaching subject (depending on their teaching area, one of F1T310B2-E2 = 1)
Highly qualified teacher	0.33	0.19	0.00	1.00	BA in subject area, 4+ years of experience, and standard credential
Process					
Discipline fair	0.63	0.15	0.18	1.00	Proportion of students who agree discipline is fair at the school (F1S7D = 1 or 2)
Homework time	4.61	2.05	1.06	14.00	Mean number of hours spent on homework per week (F1S36A2)
Teaching quality[a]	0.00	0.39	−1.54	1.14	Standardized principal component
Parent involvement	2.42	0.21	1.50	3.60	Proportion of parents who agree or strongly agree that they have adequate say in school policy (F2P42M = 1 or 2)
Academic track	0.34	0.22	0.00	1.00	Proportion of students in academic track (F1HSPROG = 2)
Class disruptions[a]	0.00	0.36	−1.05	1.66	Standardized principal component
Unsafe	0.08	0.08	0.00	0.47	Proportion of students who report they feel unsafe at school (F1S7M = 1 or 2)
NAEP composite	13.76	2.27	6.00	27.74	Number of NAEP units in math, science, English, and social science earned in high school (F2ral1_C + al2_C + geo_C, tri_C + pre_C + cal_C + bio_C + che_C + phy_C + soc_C + his C)
Teacher sense of community[a]	−0.01	0.69	−2.40	2.01	Standardized principal component
Teacher sense of control[a]	−0.01	0.72	−2.15	2.46	Standardized principal component
Teacher curriculum coordination[a]	−0.01	0.60	−2.22	2.38	Standardized principal component
Principal leadership[a]	−0.01	0.74	−2.65	1.86	Standardized principal component
Teacher expectations	−0.01	0.65	−2.43	2.15	Standardized principal component
Teacher efficacy[a]	0.01	0.56	−3.15	2.01	Standardized principal component

Note. NAEP = National Assessment of Educational Progress.
[a]See Appendix B for detailed description.

APPENDIX B

Principal Component Descriptions, Path Loadings, and Variance Explained

Factor and items label	Item description	Item loading
Teaching quality[a]		
F1S7G	The teaching is good at school	.702
F1S7H	Teachers are interested in students	.789
F1S7I	When student works hard teachers praise effort	.681
F1S7J	In class, often feel put down by teachers (reverse coded)	.591
F1S7L	Most teachers listen to student	.747
F1S66	Teachers expect student to succeed at school	.697
Variance explained (%)		49.5
Teacher efficacy[b]		
F1T4_5A	Can get through to most difficult students	.666
F1T4_5B	Teacher responsible for keeping students from dropping out	.622
F1T4_5C	Change approach if students not doing well	.626
F1T4_5D	Different methods can affect achievement	.730
F1T4_5E	Little I can do to ensure high achievement (reverse coded)	.610
F1T4_5F	Teacher making difference in students' lives	.595
F1T4_11F	Create lessons students will enjoy learning	.516
Variance explained (%)		39.3
Teacher expectations[b]		
F1T4_1D	Success/failure due to factors beyond me	.569
F1T4_1E	Student misbehavior interferes with teaching	.678
F1T4_1I	Students not capable of learning material	.631
F1T4_2J	Feel waste of time to do best at teaching	.648
F1T4_2N	Students' attitudes reduce academic success	.755
Variance explained (%)		43.4
Class disruptions[a]		
F1S7F	Other students often disrupt class	.673
F1S7K	Often feel put down by students in class	.501
F1S7N	Disruptions impede student's learning	.712
F1S7O	Misbehaving students often get away with it	.696
Variance explained (%)		42.4
Teacher community[b]		
F1T4_1B	Can count on staff members to help out	.741
F1T4_1C	Colleagues share beliefs about mission	.720

Factor and items label	Item description	Item loading
F1T4_2E	Great deal of cooperative effort among staff	.846
F1T4_2F	Broad agreement among faculty about mission	.771
F1T4_2H	School seems like a big family	.753
Variance explained (%)		58.9
Teacher control[c]		
F1T4_9A	Teachers' influence over discipline policy	.767
F1T4_9B	Teachers' influence over in-service programs	.714
F1T4_9C	Influence in grouping students by ability	.774
F1T4_9D	Influence over establishing curriculum	.748
Variance explained (%)		56.4
Teacher curriculum coordination[b]		
F1T4_1A	Coordinate course content with department teachers	.783
F1T4_1N	Coordinate content with teachers outside department	.666
F1T4_2P	Familiar with content taught by department teachers	.657
Variance explained (%)		49.6
Principal leadership[b]		
F1T4_1F	Principal poor at getting resources (reverse coded)	.711
F1T4_1G	Principal deals with outside pressures	.669
F1T4_1H	Principal makes plans and carries them out	.798
F1T4_1J	Goals/priorities for the school are clear	.709
F1T4_1L	Staff members recognized for job well done	.644
F1T4_1O	Principal knows what kind of school he wants	.848
F1T4_1P	Administration knows problems faced by staff	.756
F1T4_2I	Principal lets staff know what is expected	.861
F1T4_2K	Principal is interested in innovation	.778
F1T4_2M	Principal consults staff before decisions	.723
Variance explained (%)		56.6

Note. All variables were coded on 4–6-point Likert-type scales. Factor loadings were computed with all cases from the 1990 sample of NELS with valid F1sch_id data (N = 19,392).
[a] 4-point Likert scale (1 = *strongly agree,* 4 = *strongly disagree*).
[b] 6-point Likert scale (1= *strongly disagree,* 6 = *strongly agree*).
[c] 5-point Likert scale (1 = *no influence,* 5 = *great deal of influence*).

APPENDIX C

Descriptions of Hierarchical Linear Models

Three types of hierarchical linear models were estimated in this study: (a) a three-level achievement growth model, (b) a two-level dichotomous attrition model, and (c) a two-level multinomial attrition model. Here we briefly describe the three models.

Three-Level Achievement Growth Model

The first model estimated growth in a composite measure of student achievement between the 8th and 12th grades. The first step was to specify a Level 1 growth model as follows:

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}a_{tij} + e_{tij},$$

where Y_{tij} is the composite theta score at time t of person i attending high school j; π_{0ij} is the intercept parameter, true theta score of person i upon entering high school j; π_{1ij} is the slope parameter, theta growth rate for person i in high school j; a_{tij} represents time t for person i in high school j; and e_{tij} is the random effect for person i in school j. For this study, time was specified as three values (0, .5, 1) corresponding to 8th-, 10th-, and 12th-grade achievement values. As a result, the growth parameter, π_{1ij}, represents the estimated growth rate or total amount of learning for a student between Grades 8 and 12, or the entire 4 years of high school.

The next step was to specify a Level 2 or student model. An unconditional model, without any student-level predictors, was used to estimate mean values for both initial achievement and achievement growth rates for the entire sample of students as well as the variance in these parameters at the individual (Level 2) and school (Level 3) levels. The variance estimates were used to compute the proportion of variance in these parameters that existed at the individual and school levels (known as the intraclass correlation). A set of student-level predictors was then introduced to control for the effects of individual background characteristics on achievement growth:

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}X_{1ij} + \cdots + \beta_{0pj}X_{pij} + r_{0ij}$$

and

$$\pi_{1ij} = \beta_{10j} + \beta_{11j}X_{1ij} + \cdots + \beta_{1pj}X_{pij} + r_{1ij},$$

where β_{00j} and β_{10j} are the mean initial and mean test score growth rates, respectively, in school j; β_{01j} through β_{0qj} and β_{11j} through β_{1qj} are the estimated effects of the student-level predictors on initial achievement and growth rates, respectively, within each school; and r_{0ij} and r_{1ij} are the error terms. (As a means of producing unbiased estimates, the same set of predictors was used as controls in the initial achievement outcome [π_{0ij}] model.) All of the continuous measures were centered on the grand mean for the entire sample of students, while all of the other (dummy) variables were not centered. As a result, the intercept term, β_{10j}, represents the adjusted mean achievement growth for each school or the expected achievement for a student, absent any of the dummy variables, who had average characteristics on all other variables. This is a very useful feature of HLM in that it allows one to determine to what extent the observed differences in school performance can be attributed to differences in students as opposed to differences in features of the schools themselves.

The Level 3 model, an intercept-and-slopes-as-outcomes model, was then estimated to examine whether between-school variance in mean growth rates in test scores could be explained by measured school characteristics:

$$\beta_{00j} = \gamma_{000} + \gamma_{001} W_{1j} + \cdots + \gamma_{00q} W_{qj} + u_{00j}$$

$$\beta_{0pj} = \gamma_{0p0}$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101} W_{1j} + \cdots + \gamma_{10q} W_{qj} + u_{10j}$$

$$\beta_{1pj} = \gamma_{1p0},$$

where γ_{000} and γ_{100} are the overall or grand mean of student initial test scores and test score growth, respectively; γ_{001} to γ_{00q} and γ_{101} to γ_{10q} are the estimated effects of the school-level variables on the mean initial test scores and test score growth, respectively; and u_{00j} and u_{10j} are the residuals at school j. In all of the remaining equations, the intercept terms were "fixed" ($u_{pj} = 0$), so the effect of the within-school predictors was constrained to be the same for all schools (Raudenbush & Bryk, 2002, pp. 25–26).

Two-Level Dichotomous Attrition Model

The next model was a two-level attrition model in which the dependent variable was a dichotomous variable indicating whether students left (transferred or dropped out) their 10th-grade high school over the subsequent 2 years. In the case of dichotomous outcomes, it is necessary to specify both a Level 1 (within-school) sampling model and a Level 1 structural model (Raudenbush & Bryk, chap. 10). For binary student outcomes, the Level 1 sampling model is a Bernoulli model,

$$\mathrm{Prob}(Y_{ij} = 1 \mid \beta_j) = \Phi_{ij},$$

and the conditional Level 1 structural model is

$$\log[\Phi_{ij}/(1 - \Phi_{ij})] = \eta_{ij} = \beta_{0j} + \beta_{1j} X_{1j} + \cdots + \beta_{pj} X_{pj},$$

where the left-hand term serves as a link function (Raudenbush & Bryk, pp. 293–294). Again, the continuous predictors were centered on the grand mean to produce adjusted school means. The Level 2 (between-school) structural model is similar to the achievement growth model:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_{1j} + \cdots + \gamma_{0q} W_{qj} + u_{0j}$$

$$\beta_{pj} = \gamma_{p0}.$$

With an unconditional model, the expected probability of a student ij leaving school for a given random school effect, u_{0j}, is

$$E(Y_{ij} \mid u_{0j}) = 1/\{1 + \exp[-(\gamma_{00} + u_{0j})]\},$$

which also serves as the school attrition rate.

Two-Level Multinomial Attrition Model

The third model was a multinomial model with three outcomes (1 = student dropped out between Grades 10 and 12, 2 = student transferred between Grades 10 and 12, and 3 = student remained in same school between Grades 10 and 12). In this case, there are three probabilities,

$$\mathrm{Prob}(Y_{ij} = 1 | \beta_j) = \Phi_{1ij},$$
$$\mathrm{Prob}(Y_{ij} = 2 | \beta_j) = \Phi_{2ij},$$

and

$$\mathrm{Prob}(Y_{ij} = 3 | \beta_j) = \Phi_{3ij} = 1 - \Phi_{1ij} - \Phi_{2ij},$$

with a multinomial sampling model. To specify all three possible outcomes, it is necessary to estimate only two probabilities, with Category 3 serving as a reference category. Accordingly, two Level 1 structural models were estimated:

$$\log[\Phi_{1ij}/(1 - \Phi_{3ij})] = \eta_{1ij} = \beta_{0j(1)} + \beta_{1j(1)}X_{1j} + \cdots + \beta_{pj(1)}\, X_{pj}$$

and

$$\log[\Phi_{2ij}/(1 - \Phi_{3ij})] = \eta_{2ij} = \beta_{0j(2)} + \beta_{1j(2)}X_{1j} + \cdots + \beta_{pj(2)}\, X_{pj},$$

along with two sets of Level 2 structural models:

$$\beta_{0j(1)} = \gamma_{00(1)} + \gamma_{01(1)}\, W_{1j} + \cdots + \gamma_{0q(1)}\, W_{qj} + u_{0j(1)}$$
$$\beta_{pj(1)} = \gamma_{p0(1)}$$
$$\beta_{0j(2)} = \gamma_{00(2)} + \gamma_{01(2)}\, W_{1j} + \cdots + \gamma_{0q(2)}\, W_{qj} + u_{0j(2)}$$
$$\beta_{pj(2)} = \gamma_{p0(2)}.$$

All of the models were estimated with the same set of predictor variables and centering with one exception: The three-level achievement growth model included student-level predictors for dropout and transfer, so the predicted learning rates were for stable students.