ACT Research Report Series 2008-1

Benefits of Additional High School Course Work and Improved Course Performance in Preparing Students for College

Richard Sawyer

ACT

August 2008

For additional copies write:
ACT Research Report Series
P.O. Box 168
Iowa City, Iowa 52243-0168

Benefits of Additional High School Course Work and Improved Course Performance in Preparing Students for College

Richard Sawyer

Abstract

For more than two decades, education authorities have warned that the reading, writing, mathematics, and science skills of America's young adults are insufficient to maintain its economic strength. Some indicators suggest that the average educational achievement of high school students in the U.S. is mediocre in comparison to that of other industrialized countries, and that it is not improving significantly. Moreover, there are wide differences in educational achievement among demographic groups.

Authorities have recommended many different strategies for improving the educational achievement of high school students. One such strategy is to encourage them to take more rigorous college-preparatory courses and to earn higher grades in these courses. We studied the effectiveness of this strategy using data from students who took ACT's EXPLORE test in eighth grade, the PLAN test in tenth grade, and the ACT in eleventh/twelfth grade. The outcome variables in the study were students' ACT scores in English, Mathematics, Reading, and Science. The predictor variables were students' background characteristics, their previous educational achievement (as measured by their EXPLORE scores), the high school they attended, their course work, their course grades, and variables related to the context in which they took the ACT.

We constructed models for predicting the outcome variables from the predictor variables. We then used the models to estimate the proportion of students who, under various scenarios of enhanced preparation, would have ACT score levels indicating that they were adequately prepared to take typical first-year college courses. The principal results are as follows:

- Students' background characteristics, EXPLORE scores, high school attended, high school course work, and high school grades are all related to ACT scores, but EXPLORE scores are by far the most strongly related.
- Improving EXPLORE scores is likely to be more effective in improving ACT scores than other forms of enhanced preparation.
- Taking more standard or advanced courses in high school and earning higher grades is more beneficial to students who have high EXPLORE scores to begin with.
- There is significant variation in high schools' average ACT scores, even after accounting for differences in their students' characteristics. The benefit of additional standard course work, advanced/honors course work, and higher grades also varies significantly among high schools.

The report concludes with a discussion of implications of the results and recommendations for additional research.

Benefits of Additional High School Course Work and Improved Course Performance in Preparing Students for College[1]

This report investigates the effectiveness of taking additional courses and earning higher grades for improving high school students' academic preparation for college. It is based on data from students who took ACT's EXPLORE test in eighth grade, the PLAN test in tenth grade, and the ACT in eleventh/twelfth grade. The outcome variables in the study are students' ACT scores in English, Mathematics, Reading, and Science.

Importance of Educational Achievement

A substantial body of research concludes that education is vitally important to the economic prosperity of societies and individuals. At the social level, increasing educational achievement makes labor more productive, and more competitive with that in other countries (Bernanke, 2007). At the individual level, postsecondary education is required to obtain and hold jobs that pay enough to maintain a high standard of living (U.S. Department of Labor, 2007; Council on Competitiveness, 2007). Therefore, it is important that young people be adequately prepared to continue their education after graduating from high school.

Educational achievement also has implications beyond maintaining economic prosperity. For example, Lochner (2004) found that, even after controlling for a variety of other socioeconomic characteristics, parents' and children's education levels were negatively correlated with children's participation in crime. Literacy and numeracy skills are also important for people to participate fully in society (de León, 2002; Steen, 1999). Citizens need these skills to keep informed of current events, to understand and critically analyze leaders' qualifications and proposals, and to participate in civic and social organizations. As changes in

[1] I would like to thank Jeff Allen, Julie Noble, Richard Phelps, and Dan Vitale for their very helpful comments and suggestions.

communications technologies have increased the volume and choices of information, the skills required to use them effectively have also increased.

Educational Achievement in the U.S.

The U.S. does not excel among other industrialized nations in the educational achievement of its young people. The PISA 2003 survey compared the mathematics, reading, and science skills of U.S. fifteen-year-olds to those of fifteen-year-olds in other countries belonging to the Organisation for Economic Co-operation and Development. The U.S. average scores in mathematics and science were statistically significantly lower than the corresponding average scores of other industrialized nations (Lemke et al., 2004). In mathematics, the U.S. ranked 24th out of 29; in science, the U.S. ranked 20th. (The U.S. performed relatively better in reading, in which it ranked 15th.) In the follow-up PISA 2006 survey (Organisation for Economic Co-operation and Development, 2007), U.S. average scores in mathematics and science were still statistically significantly lower than the corresponding average scores of other industrialized countries.

In contrast, results from the 2003 Trends in Mathematics and Science Study (TIMSS) (TIMSS and PIRLS International Study Center, 2004), showed that fourth-grade and eighth-grade students in the U.S. had statistically significantly higher average scores in mathematics and science than did fourth- and eighth-grade students in the other participating countries. Although it is not possible to make exact comparisons between the PISA and TIMSS results[2], they suggest that U.S. performance relative to that of other countries declines after grade eight.

Recent results from the National Assessment of Educational Progress (Grigg, Donahue, & Dion, 2007) are not encouraging. Only 35% of twelfth-grade students in 2005 had reading

[2] The countries participating in PISA and TIMSS differ, as do the target student populations. Moreover, the PISA tests emphasize problem-solving / reasoning skills more than the TIMSS tests do.

skills at or above a level deemed to be proficient according to NAEP's policy definition. In mathematics, moreover, only 23% of the students performed at or above the proficient level. The average reading skills of twelfth-grade students in 2005 showed no significant change from the previous assessment in 2002, and declined in comparison to 1992.

The scores of most students on the ACT college-entrance test are below what is needed for them to be adequately prepared for college. Only 20% of ACT-tested high school graduates in 2002 were prepared to take standard first-year courses in college English, mathematics, social studies, and science, as indicated by their ACT scores; among ACT-tested high school graduates in 2007, this percentage was 23% (ACT, 2007e). Academically underprepared students not only have to spend more time and money taking remedial courses in college; they also earn lower grades and have lower retention rates (Noble & Radunzel, 2007).

These percentages pertain to ACT-tested students, but the situation would likely be worse if extended to all high school graduates in the U.S. Although not all students need to go to college directly after high school, the reading and mathematics skills they need to hold high-paying jobs are similar to those that they need to succeed in college (ACT, 2006). Moreover, about 55 percent of job openings between 2004 and 2014 will require some postsecondary education or training (U.S. Department of Labor, Bureau of Labor Statistics, 2006).

A concomitant disturbing phenomenon is that different demographic groups have markedly different levels of literacy and numeracy skills. In the PISA 2003 survey, for example, the average mathematics, reading, and science scores of U.S. Hispanic fifteen-year-olds were .5 to .6 standard deviations lower than those of the total group of U.S. fifteen-year-olds. The average scores of U.S. black fifteen-year-olds were .7 to .8 standard deviations lower than those of the total group (Lemke et al., 2004). Among 2007 high school graduates who took the ACT, only 10% of Hispanic students and 3% of black students were prepared to take standard first-year courses in college, as compared to 23% of all tested students (ACT, 2007d).

Furthermore, the demographic composition of U.S. students is expected to shift. For example, the percentage of public high school graduates who are Hispanic is projected to increase from about 12% in 2002 to about 25% in 2022 (Western Interstate Commission for Higher Education, 2008). The corresponding percentage for blacks is projected to remain stable at about 13%. The percentage for white non-Hispanics is projected to decline from 69% in 2002 to 52% in 2022.

Evidence like that described here has prompted several organizations and commissions to warn that the academic skills of America's young adults are insufficient to maintain its economic strength and technological leadership (National Commission on Excellence in Education, 1983; Carnegie Forum on Education and the Economy, 1986; Council on Competitiveness, 2007; Committee on Prospering in the Global Economy of the 21st Century, 2007; National Center on Education and the Economy, 2007). Other organizations and individuals have investigated the relationship between countries' educational development and economic prosperity (Ramirez, Luo, Schofer, & Meyer, 2006; The World Bank, 2006; Tienken, 2008).

Enhancing Students' Educational Achievement Through Course Work and Grades

Many different recommendations have been made for improving students' academic skills. They include restructuring public education bureaucracies, charter schools, voucher systems, accountability systems for schools, improving teacher training, raising requirements for licensure, accountability systems for teachers, higher salaries for teachers, special incentives for excellent teaching, improving curricula, new methods of instruction, greater involvement of parents, special programs and services for at-risk students, raising requirements for graduation, stricter policies on student behavior, and raising public awareness about the importance of education, among others. Acrimonious debate accompanies many of these suggestions, particularly those that involve punitive sanctions, spending large sums of money, or shifts in political power.

One seemingly obvious, and less controversial, way to improve high school students' academic achievement is to encourage them to take more challenging courses: Taking more challenging courses exposes students to more of the content that they need to master before college or work. There would seem to be room for improvement in high school students' course work. As of 2006, only 26 states required high school students to take any mathematics courses at all in order to graduate (ACT, 2007c). Of these 26 states, 12 required Algebra II, and only four states required any mathematics beyond Algebra II. ACT research has shown that Algebra II has a substantial impact on student readiness for college (ACT, 2004). In science, while 30 of the 50 states required at least one course for graduation, only 17 explicitly required Biology, one explicitly required Chemistry, and two explicitly required Physics. The Education Commission of the States (2006) summarized the alignment of states' high school graduation requirements with any statewide requirements for unconditional admission to college.

Merely enrolling in and sitting through more challenging courses might not improve students' educational achievement, however. Another important consideration is the grades they earn: The extra effort required to earn higher grades should result in more learning. Among high school graduates whose high school GPA was below 2.0, for example, the average ACT Composite score was 16.0 (on a scale from 1-36). Among students whose high school GPA was above 3.5, the average ACT Composite score was 24.2, a 1.7 standard deviation difference (ACT, 2005).

Of course, course work, course grades, and test scores are all affected by prior educational achievement. Students who begin high school with stronger academic skills will likely take more challenging courses, earn higher grades, and have higher test scores later. To estimate the extent to which taking more challenging courses and earning higher grades is likely to benefit students' test scores, one must, at a minimum, control for prior educational achievement.

Two recent reports from the U.S. Department of Education illustrate this point. The percentage of high school students who completed a standard academic curriculum increased from 31% to 51% between 1990 and 2005, and overall GPA increased from 2.68 in 1990 to 2.98 in 2005 (Shettle et al., 2007). Nevertheless, the average NAEP Reading score of twelfth-grade students declined from 292 to 286 (on a 500-point scale) between 1992 and 2005 (Grigg, Donahue, & Dion, 2007). Combining these two results apparently contradicts the expectation that taking more courses and earning higher grades will improve academic achievement (Landsberg, 2007, February 23). To make this inference, however, one would need to control for prior educational achievement, grade inflation, and other variables.

A more subtle consideration is that the benefit of taking more challenging courses and earning higher grades might not be the same for all students, but instead might depend on their prior achievement. For example, although all students might benefit to some degree from taking an additional mathematics course or from earning higher grades in their mathematics courses, students with high prior mathematics achievement might benefit *more* from doing so than students with low prior mathematics achievement. We have estimated these interaction effects in this study.

Some courses are taught at a more advanced level or at an accelerated pace. Examples include courses designed by the Advanced Placement Program (College Board, 2006) or by the

IB Diploma Program (International Baccalaureate Organization, 2007). Alternatively, high schools often offer accelerated or honors courses designed by their own faculty. The enriched content of such courses could make them more beneficial than standard courses.

Typically, but not always, advanced/honors courses are taken by students with above-average academic achievement. Carbannaro (2005) modeled students' mathematics test scores in tenth grade from their background characteristics, their eighth-grade test scores in different subject areas, curriculum track (e.g., advanced/honors vs. general academic), and effort (defined by behavioral and psychosocial characteristics). As one would expect, Carbannaro found that advanced/honors-track students had significantly higher eighth-grade test scores and expended greater effort than did other students. He also found that students' tenth-grade scores were strongly related to their eighth-grade scores and to effort, but only modestly related to curriculum track. Carbannaro did not include in his model interactions of prior test scores by track, but he did find that interactions of effort by track were not important predictors of tenth-grade scores.

We also need to consider that the benefit of taking more challenging courses and earning higher grades could vary by high school. Taking an additional course at a school in a safe neighborhood with excellent facilities, a rigorous curriculum, and well-prepared teachers might be more beneficial than taking a similar course in another school without these advantages. Raudenbush (2004) distinguished between two kinds of school-level characteristics: context (over which schools have no control) and practice (over which they do have control).

Other Variables Related to Achievement

Grading standards differ among subject areas, teachers, and schools. Ideally, an analysis should take into account all these sources of variation. The data in this study permitted studying variation by subject area and school, but not by teacher.

We can think of high school course grades as composite indicators of cognitive achievement and of behavioral and psychosocial characteristics, such as self-discipline, effort, attendance, conformity, and motivation (Stiggins, Frisbie, & Griswold (1989); Carbannaro (2005), Duckworth & Seligman (2006), Noble, Roberts, & Sawyer (2006))[3]. Although ACT scores are direct indicators of academic achievement, their underlying constructs (academic achievement in different subject areas) might also be influenced, either directly or indirectly, by psychosocial characteristics and behavior. For this study, we did not have data on psychosocial and behavioral variables, but they are promising characteristics to research.

Another issue is the extent to which students' background characteristics (e.g., family income, race/ethnicity, gender) are related to their test scores, given their course work and course grades. Besides defining politically important subgroups of the population, background characteristics are related to behavioral and psychosocial characteristics that drive educational achievement. For example, high-income families might be able to provide their children with more educationally beneficial activities than low-income families; or, they might have different expectations for males' and females' achievement in mathematics and science, which could affect their motivation and effort. Because we did not have data on psychosocial characteristics and behavior in this study, we are restricted to using background characteristics as proxies.

Research on the Benefit of Course Work and Grades

This report summarizes exploratory research on the benefit of additional high school course work and higher grades for increasing students' academic achievement in grades eleven/twelve, as measured by their scores on the ACT college entrance test (ACT, 2007a). The research investigated the benefit of enhanced preparation activities, given students' background

---

[3] ACT staff members are also currently investigating the relationships among psychosocial characteristics and behavior, family involvement, and academic success in grades six and higher.

characteristics, prior academic achievement, and high school attended. This report provides information on the following questions:

- How important is academic achievement in grade eight for predicting academic achievement in grades eleven/twelve?
- How important are course work and grades in high school, given eighth-grade achievement, for predicting achievement in grades eleven/twelve?
- How much improvement in achievement could we expect from students' taking additional rigorous courses and earning higher grades, given their other characteristics?
- Does prior achievement in grade eight affect the benefit of subsequent additional course work and higher grades?
- Does high school attended affect the benefit of additional course work and higher grades?

The research is exploratory, in the sense that it involves identifying (and confirming through replication in a second data set) important predictors of academic achievement, rather than testing formal theoretical models that explain how the variables affect each other. The results could, however, inform the development of formal models.

Table 1 on the following two pages summarizes fourteen studies published since 1990 on the relationship of high school students' test scores with their course work and course grades, given other relevant variables. Among the previous studies, only eight controlled for prior achievement test scores[4], and only one of them (Burkam & Lee, 2003) controlled for both prior

---

[4] The study by Chaney, Burgdorf, and Atash (1997) used Grade 9 GPA as a measure of prior achievement.

TABLE 1
Research on Predicting Educational Achievement from High School Course Work (CW) and Other Variables

Study	Educational achievement test score outcome variable(s)	Breadth of data	Predictor variable(s): Standard CW	Prior achievement test scores	Adv./Hon. CW	Subj.-area grades	Background variables	HS characteristics	Hierarchical modeling
Previous research									
Schiel, Pommerich, & Noble (1996)	ACT Mathematics, ACT Science	National	X	X			X		(HS dummy vars.)
Chaney, Burgdorf, & Atash (1997)	NAEP Mathematics, NAEP Science	National	X		X		X	X	X
Madigan (1997)	NELS:88 Science	National	X	X			X	X	
Meyer (1999)	High School and Beyond (Mathematics)	National	X	X			X		
Girotto & Peterson (1999)	ITED Reading ITED Mathematics ITED Science	Local	X	X		(overall GPA)	X		
Burkam & Lee (2003)	NELS:88 Mathematics	National	X	X		X	X		
Burkam (2003)	NELS:88 Reading	National	X		X				
Perkins et al. (2004)	NAEP Mathematics, NAEP Science	National	X		X	X	X	X	
Carbannaro (2005)	NELS:88 Mathematics	National	(track only)	X	X		X		
Ma & McIntyre (2005)	Can. Ach. Test (Mathematics)	Regional	X			X	X	X	X

(continued on next page)

TABLE 1 (continued)
Research on Predicting Educational Achievement from High School Course Work (CW) and Other Variables

Study	Educational achievement test score outcome variable(s)	Breadth of data	Predictor variable(s)						Hierarchical modeling
			Standard CW	Prior achievement test scores	Adv./Hon. CW	Subj.-area grades	Background variables	HS characteristics	
Previous research (continued)									
Leow, Marcus, Zanutto, & Boruch (2004)	TIMSS Mathematics, TIMSS Science	National	X				X		
Shettle et al. (2007)	NAEP Mathematics NAEP Science	National	X		X	X	X		
Noble & Schnelker (2007)	ACT Mathematics ACT Science	National	X	X			X	X	X
Bozick & Ingels (2007)	ELS: 2002 Mathematics	National	X	X	X		X	X	
Current study	ACT English ACT Mathematics ACT Reading ACT Science	National	X	X	X	X	X		X

achievement test scores and subject-area grades.[5] None of the fourteen previous studies controlled for all the variables listed in Table 1. In all of the studies, there was a positive relationship between educational achievement and course work.

The studies differ in how they measure course work. One of the studies (Carbannaro, 2005) analyzed course track (general academic/honors/vocational), rather than course work in particular subject areas. In contrast, Noble and Schnelker (2007) studied particular courses (e.g., calculus), and Bozick and Ingels (2007) studied particular course sequences (e.g., geometry – algebra II – precalculus vs. precalculus paired with another course). The mathematics and science course work variables in the current study are based on the highest-level course taken (see Table A-1 in the appendix). The course work definitions used in this study, while not as detailed as those used by Noble and Schnelker and by Bozick and Ingels, permit us to investigate at a general level the extent to which taking higher-level courses increases educational achievement, while keeping the complexity of the models at a manageable level.

Curricula, instructor effectiveness, and grading standards can vary among high schools. Therefore, relationships between achievement and course work, and between achievement and course grades, also vary among high schools. Some of the studies summarized in Table 1 addressed this issue by including high school characteristics as predictors in their models. Doing this essentially had the effect of adjusting the intercept terms according to high school characteristics. Three of the studies summarized in Table 1 also used hierarchical modeling, a more powerful approach in which intercepts and slopes can all vary among high schools, both randomly and according to high school characteristics.

[5] The study by Girotto and Peterson (1999) controlled for overall GPA, rather than for course grades.

Many states and localities have value-added accountability systems (Center for a Greater Philadelphia, 2007), in which students' test scores are interpreted with respect to their scores on an earlier test, rather than with respect to an absolute standard or a distribution of scores in a norm group. Value-added accountability systems require collecting data on students' prior educational achievement and the schools they attend. The organizations administering the systems likely also collect data on students' background characteristics, course work, and grades. Therefore, these systems are a potential source of data that could be used to investigate the benefits of increased course work and higher grades.

Data

Our research is based on data from students who took the EXPLORE, PLAN, and ACT tests. Each of these tests measures students' achievement in written English[6], mathematics, reading, and science. Students typically take EXPLORE in grade eight, PLAN in grade ten, and the ACT in grades eleven/twelve. The contents of all three tests are aligned to measure knowledge and skills that are typically taught in the targeted grades, and that are related to the knowledge and skills that students eventually will need to succeed in college (ACT, 2007b; ACT, 2008; ACT, 2007a).

Source Files and Analysis Files

To construct predictive models, we used data on students who took all three tests (EXPLORE, PLAN, and the ACT) and who graduated from high school in 2005 or 2006. The source data set for 2005 contains records for 112,734 students at 4,515 high schools. The source data set for 2006 contains records for 132,441 students at 4,992 high schools. We constructed models from the 2005 data, and replicated them using the 2006 data.

[6] The English tests in EXPLORE, PLAN, and the ACT consist of multiple-choice items measuring students' knowledge of the conventions of standard written English. The ACT also includes an optional direct-writing (essay) component.

For various reasons, we estimated models from a subset of each year's source data set:

- All records in the original source files had valid ACT scores, and about 99% had valid EXPLORE scores, PLAN scores, and race/ethnicity code. (There were no records with missing gender code.) We selected for analysis records with valid EXPLORE and PLAN scores and race/ethnicity code.
- Some students take the ACT as early as age 12 for admission to gifted-and-talented programs; others wait until they enroll in college. Because we were interested in making inferences about high school students who test in grades eleven or twelve, we selected for analysis the records of only these students.
- Finally, we eliminated records with high school codes that we could not locate in a file maintained by Market Data Retrieval, Inc. For the most part, these codes either corresponded to small special-purpose schools or schools in foreign countries, or else were not correctly entered by students.

In the resulting analysis file for the 2005 graduates, there were records for 98,812 students at 4,191 high schools. In the analysis file for the 2006 graduates, there were records for 117,280 students at 4,638 high schools. About 60% of the excluded records were associated with students who did not take the ACT in grades eleven or twelve.

Retested Students

Many students who take the ACT do so more than once. A typical retesting scenario occurs when a student takes the ACT in the spring of the eleventh grade and retests in the fall of the twelfth grade. In general, students retest because they want to increase their scores. As a result of more time for learning and greater familiarity with the test, scores on a second testing average about 0.7 - 0.8 score points higher than scores on the initial testing (Lanier, 1994;

Andrews & Ziomek, 1998). Whenever possible, we used students' *last* ACT record so that the results of the analyses would reflect the most current information about their course work, grades, and academic achievement.

A complicating factor is that students who retest are not required to update their course work and course grade information.[7] To ensure that the course work, course grade, and test score information for a given student were contemporaneous, we used the first record of multiple-tested students whenever the amount of course work and course grade information on the last testing did not exceed the amount reported on the first testing. In the final data sets, about 43% of the records were from single-tested students, 32% were the first records of multiple-tested students, and 25% were the last records of multiple-tested students.

Analysis Variables

The outcome variables in the study were students' ACT English, Mathematics, Reading, and Science scores. ACT scores measure students' cognitive achievement in the respective subject areas. ACT scores predict academic success in college, as measured by grades in first-year courses, by first-year and longer-term GPA, and by retention (ACT, 2007a). Future research at ACT will consider how the predictor variables in this study, as well as ACT scores, relate to success in college.

We used students' four EXPLORE scores as measures of prior achievement in all the models. The contents of the EXPLORE tests align with those of the corresponding ACT tests, but at a lower level of educational development.

Students' average ACT scores, as well as the relationship of ACT scores with course work, course grades, and other predictor variables, differ among high schools. We used students' self-

[7] Effective in fall 2007, students are required to provide course work and course grade information when registering for the ACT through the Internet.

reported high school code to identify the high school in which they enrolled. We excluded from the analyses records without a valid high school code.

Students provide data on their background characteristics, high school course work, and high school grades when they register for the ACT. Students also provide data on their background characteristics and high school course work when they take PLAN, and data on their background characteristics when they take EXPLORE. We used data elements from EXPLORE and PLAN to supplement missing data on the ACT. For example, if a student did not provide racial/ethnic information when taking the ACT, we used racial/ethnic information from EXPLORE or PLAN.

The variables we used to measure standard high school course work were: highest-level English course taken, highest-level mathematics course taken, number of social studies courses taken, highest-level science course taken, and a foreign language dummy variable (1=took any foreign language course). We did not use non-core curriculum courses, such as business mathematics and speech, in constructing the course work variables.

When students register for the ACT, they also indicate whether they have taken Advanced Placement, accelerated, or honors courses in English, mathematics, science, social studies, or foreign languages. We used their responses to create Adv./Hon. dummy variables corresponding to these subject areas. Note that each Adv./Hon. dummy variable indicates only whether a student took any Adv./Hon. course in a general subject area (e.g., mathematics); it does not identify a particular Adv./Hon. course (e.g., Trigonometry) that a student might have taken.

The course grade variables are averages of the grades that students reported for the courses they took in the subject areas of English, mathematics, social studies, and science. The range of each of these variables is 0.0 – 4.0.

We used the following background variables as predictors: gender, race/ethnicity, parents' educational level, family income, and English as the primary language spoken at home. Deferring analyses for particular racial/ethnic groups to a later time, we considered race/ethnicity as a dummy variable, with value 1 corresponding to the two high-scoring racial/ethnic groups (Caucasian-American/White; Asian American, Pacific Islander) and value 0 corresponding to all other groups (African-American/Black (non-Hispanic); American Indian, Alaska Native; Mexican-American/Chicano; Puerto-Rican, Cuban, Other Hispanic Origin; Other; and Multiracial). Two variables related to parents' education: the number of parents who graduated from high school and the number of parents who attended college.

Another class of variables relates to the context of students' ACT testing. Students can take the ACT as many times as they wish, and at any age and educational level. Students' choices on these variables depend on their educational achievement, psychosocial characteristics (e.g., motivation, goals, and self-discipline), interactions with others (e.g., counselors), and other behavior (e.g., test preparation).

Students who choose to take the ACT at a younger age tend to score higher than students who choose to take the ACT at an older age. On the other hand, among students of a given age, those who test in grade 12 tend to score higher than those who test in grade 11. Age in this study retains the fractional part of years (e.g., 16.37).

The relationship between students' scores and their retesting status is more complex. Retested students who update their course work and course grade information tend to score higher on retesting than on the initial testing. In contrast, retested students who do not update their course work and course grade information tend to score lower on retesting than on the initial testing. Of course, merely providing course work and course grade data does not cause students to possess more academic skills: Providing these data could be driven by noncognitive characteristics, such as compliance with rules and conscientiousness, which also help students acquire academic skills.[8] In our models, we summarized retesting status with two dummy variables: The first dummy variable is equal to 1 for retested students who did not update their course work and course grade information. The second is equal to 1 for retested students who did update their course work and course grade information. For students who did not retest, both dummy variables are equal to 0.

Another special situation affected students' ACT scores. In Colorado and Illinois, all students took the ACT during the eleventh grade during the time period associated with our data; in all other states, students self-selected to take the ACT. We therefore constructed two dummy variables corresponding to Colorado and Illinois; the dummy variables were defined by the state in which a student's high school was located.

Table A-1 in the appendix defines the outcome and predictor variables, and shows the steps taken to edit the 2006 data.

- The column under the heading "Source data set" pertains to the original file of all students who took EXPLORE, PLAN, and the ACT (N=132,441).
- The columns under the heading "Imputation data set" refer to the subset of the data with non-missing EXPLORE scores, PLAN scores, and the race/ethnicity variable, and with educational level at time of testing equal to 11 or 12 (N=123,748). We imputed any missing values of the other variables in the "Imputation data set," creating five versions of it (see *Missing Data* below).

[8] Another possibility is that some students choose to update their course work and course grade information only if they surpass some threshold in taking additional rigorous course work after the initial testing.

- The column under the heading "Analysis data set" (N=117,280) refers to the further subset of records with valid high school codes that could be found in the MDR file; it was from this file that we estimated the predictive models. There are five versions of the "Analysis data set," one for each imputation.

We followed a similar procedure for the 2005 data (not summarized in Table A-1).

Missing Data

Some student records had missing values on one or more predictor variables in the models. To simplify the analyses and to increase the number of records at small high schools, we imputed missing values using SAS PROC MI (SAS Institute, 2007). The MI (multiple imputation) procedure replaces missing values of variables with estimates based on non-missing data. The procedure retains the variance of each variable by replacing missing values with plausible values that reflect the uncertainty in their estimation. We created five imputed versions of the 2005 data set and five imputed versions of the 2006 data set.

Table A-1 in the appendix shows the percentage of cases, by predictor variable, with missing data in the 2006 data set. The largest percentages of missing data were associated with the background variables "Number of parents who attended college" and "Family income" (24% and 21%, respectively). Approximately 6% of cases had missing data on the standard course work variables (except for "Studied any foreign language" which had 22% missing data). The Adv./Hon. course work variables had about 15% missing data, and the subject-area grade averages had 15%-19% missing data.

Table A-1 also shows, for each imputed predictor variable, the difference between the mean value among cases with non-missing data and the mean value among cases with imputed data. Table A-1 also shows standardized differences for each predictor variable. The standardized difference was obtained by dividing the aforementioned difference by the standard deviation based on all cases (non-missing data and imputed data). The standardized differences were, for the most part, small: about 0.1 or less for the background, standard course work, and Adv./Hon. course work variables, and about 0.2 for the subject-area grade averages. The exception was the variable "Age at time of ACT testing," for which the standardized difference was 0.35; less than 1% of cases had missing data on this variable, however.

Method

Model Construction

Using data for the high school graduates of 2005, we developed predictive models of students' ACT scores, given their background characteristics, EXPLORE scores, ACT testing characteristics, standard course work in high school, Adv./Hon. course work, and subject-area grade averages. In follow-up research currently underway, models will be estimated that also consider PLAN scores, as well as course work taken before PLAN and after PLAN.

Because we anticipated that the predictive relationships might differ among high schools, we developed hierarchical linear models, in which regression weights relating predictor variables to outcome variables can vary among high schools. In addition to providing estimates of the variability of regression weights across high schools, hierarchical models lead to more accurate inferences about the statistical significance of the weights at typical high schools. We used the HLM6 software (Raudenbush, Bryk, Cheong, & Congdon, 2004) to estimate the hierarchical models.

We considered all four EXPLORE scores, all the background variables, and all the ACT testing characteristics for inclusion in the models. To reduce the potential complexity of the models, however, we considered course work and course grade average variables according to their alignment with the dependent variables. Thus, for example, we did not consider the potential contribution of mathematics course work for predicting ACT Reading score.[9]

Selection of predictor variables in main-effects models. Using data from the first imputation data set from 2005, we first estimated random-intercept main-effects models, in which all regression coefficients, except for the intercept (constant term), were constrained to be the same for each high school. Our goal was to construct parsimonious models in which all regression coefficients were highly statistically significant ($p<.001$ for student-level variables; $p<.01$ for school-level variables). Because there are many subsets of predictor variables, it was not feasible to examine all of them. We therefore used the following strategy for student-level variables in building the random-intercept main-effects models:

- We started with a random-intercept model that included all the predictors shown in Table A-1.
- Predictors whose estimated regression coefficients were statistically significant at $p<.001$ were retained.
- Predictors whose estimated regression coefficients were not statistically significant at $p<.01$ were dropped from further consideration.
- Predictors whose estimated regression coefficients were statistically significant at $p<.01$, but not at $p<.001$, were tentatively retained for consideration in subsequent cycles. If they did not reach $p<.001$ in subsequent cycles, they were dropped.

Cycles were repeated until all predictor variables either were statistically significant at $p<.001$ or were dropped. For school level variables, we used a similar scheme, with significance level

[9] Some reading and writing experts (e.g., Barton, Heidema, and Jordan (2002)) advocate incorporating reading-to-learn and writing-to-learn strategies in mathematics and science courses. Because our data do not identify these kinds of courses, we are not able to estimate their effects.

cutoffs $p<.01$ for outright retention, $p<.05$ for tentative retention, and $p>.05$ for outright rejection.

We next considered adding a random effect for each predictor variable in the main-effects model. (The random effect for a predictor variable indicates the variability of its regression weight across high schools.) We used a scheme similar to that described in the preceding paragraph, with significance level cutoffs $p<.01$ for outright retention, $p<.05$ for tentative retention, and $p>.05$ for outright rejection. While this strategy might not have detected the maximal subset of predictor variables whose student-level effects are statistically significant at $p<.001$ and whose school-level effects are statistically significant at $p<.01$, we believe it is a reasonable approximation.

Interaction effects. We next constructed interaction models, which include interactions among the predictor variables in the main-effects models. The interactions we considered were:

- Age at time of testing *x* educational level at time of testing
- Age at time of testing *x* retesting dummy variables
- Educational level at time of testing *x* retesting dummy variables
- EXPLORE score *x* standard course work
- EXPLORE score *x* Adv./Hon. course work
- EXPLORE score *x* course grade averages
- Standard course work *x* course grade averages
- Adv./Hon. course work *x* course grade averages

In considering the benefit of standard course work, for example, the main-effects models assume that the average increase in ACT Mathematics score, given additional mathematics course work, is the same no matter what EXPLORE Mathematics scores students might have. Estimating the

EXPLORE Mathematics score by mathematics course work interaction allows us to determine whether additional mathematics course work is more beneficial for students with higher EXPLORE Mathematics scores than for students with lower EXPLORE Mathematics scores.

In deciding which of these potential interaction terms to include, we followed a strategy similar to that for the main-effects models.

Other interaction relationships are possible. Leow, Marcus, Zanutto, and Boruch (2004), in a study of mathematics and science achievement in high school, found evidence of interactions between advanced course work and various demographic, psychosocial, and behavioral variables. In a predictive model for the Dallas, Texas, value-added accountability system, Webster and Mendro (1997) included interaction terms among background variables. To reduce the complexity of the models, we did not include interactions like these.

Centering. Analysts often recommend centering predictor variables (i.e., subtracting from each variable either its group (high school) mean or the overall grand mean). Benefits of centering include more precise computation and easier interpretation of the intercept (i.e., the constant term in the model). In group-mean centering, each predictor variable is interpreted as a deviation from the group (high school) mean. In grand-mean centering, a predictor variable is interpreted as a deviation from a single, fixed constant (the grand mean), and therefore has the same meaning across high schools.[10] In this study, we used grand-mean centering, so that a predictor variable would have the same meaning across high schools.

Level-2 predictors. Hierarchical linear models can describe the variation of regression coefficients across schools with purely random effects and with fixed effects (i.e., school-level characteristics as predictors of the coefficients). We considered potential random effects for all

[10] In principle, grand-mean centered random-slope models and uncentered random-slope models yield the same estimated slope fixed effects. Group-mean centered random-slope models, however, yield estimated slope fixed effects different from those obtained from grand-mean centered or uncentered random-slope models.

coefficients in the models. We did not attempt, however, to investigate school-level characteristics, with the following two exceptions, both of which pertain to the intercept term:

- *School means.* The slope estimates in a grand-mean centered random-intercept, fixed-slope model are a blend of within-group person effects and between-group effects (Raudenbush & Bryk, 2002). When group means are included as Level-2 predictors of the intercept, however, the slope estimates are unbiased estimators of within-group person effects. Because we are interested in estimating within-group person effects, we included the group means of the predictor variables as potential Level-2 predictors of the intercept.
- *Statewide testing.* All eleventh-grade students in public high schools in Colorado and Illinois took the ACT. Consequently, the tested populations in these two states differ from the self-selected populations tested in other states. We therefore included two state dummy variables corresponding to Colorado and Illinois as Level-2 predictors of the intercept.

We have deferred more extensive Level-2 modeling to the future.

Analysis of residuals. Associated with each student is an observed ACT score and a predicted ACT score (based on the hierarchical model). The difference between the observed and predicted ACT score is the Level-1 residual. Using a 1% sample of the data, we standardized the Level-1 residuals to have mean 0 and variance 1. We then compared the ordered standardized residuals to the corresponding quantiles of the normal distribution. We also created scatterplots of the estimated Level-1 residuals against the most important student-level predictor variables.

Some of the regression coefficients in the hierarchical models are the same for all high schools, but others vary among high schools. For the regression coefficients that vary among high schools, we studied the estimated Level-2 residuals (differences between the coefficients for a typical high school and the corresponding coefficients estimated for individual schools). To investigate the plausibility of the model assumption that the Level-2 residuals have a normal distribution, we first standardized the estimated Level-2 residuals to have mean 0 and variance 1, then compared the ordered standardized residuals to the corresponding quantiles of the standard normal distribution. To examine the plausibility of the multivariate normality assumption for the joint distribution of the random effects, we plotted the Mahalanobis distance against the corresponding quantile of the chi-square distribution (Raudenbush & Bryk, 2002). We also created scatterplots of the estimated Level-2 residuals against the school-level means used as predictors of the intercept.

Stability Over Time

An important question about any statistical relationship is whether it holds up over time. Because the final models involved extensive comparisons among many alternative potential models, there was an inevitable capitalization on chance: Although we set the thresholds conservatively, relationships that appear to be statistically significant at a particular threshold in the 2005 data might not be in the future. Moreover, irrespective of model fitting artifacts, relationships might themselves change over time. We therefore re-estimated all the models using data from the 2006 graduates. We did not repeat the entire model construction procedure, but instead simply re-estimated the fixed-effects coefficients and Level-2 variances using each of the five imputations of the 2006 data.

Aggregate Benefit of Additional Course Work and Higher Grades

One might also ask "What would be the overall benefit, aggregated over students at all schools, of their taking additional course work and earning higher grades in the subject areas identified by the predictive models?" One way to answer this question is to compare the percentage of all students whose ACT scores would meet the ACT College Readiness Benchmarks under various scenarios. The College Readiness Benchmarks are the minimum ACT scores at which students are likely to succeed in beginning college-level courses at typical postsecondary institutions. The Benchmark scores are: English-18, Mathematics-22, Reading-21, and Science-24 (Allen & Sconing, 2005).

From the interaction models estimated from the 2006 data, we calculated the percentage of students who would meet the various College Readiness Benchmarks under the following scenarios:

- With current EXPLORE scores, course work, and grades
- Increasing EXPLORE scores by two points in each subject area
- Taking one additional standard college-preparatory course of each type present in the model
- Taking Adv./Hon. courses in each relevant subject area
- Increasing the grade average in each relevant subject area by one letter grade

These forms of enhanced preparation, although feasible for some students, probably exceed what others are able or willing to do. They represent maximum credible levels of enhanced preparation. We refer to these scenarios as "High effort."

In setting up the scenarios, we capped the value of each predictor variable at its maximum (see Table A-1 in the appendix): For example, a student who already had a 4.0 grade average could not increase her or his grade average beyond 4.0. We then calculated for each student a predicted ACT score under each scenario, using the relevant hierarchical regression model. To each predicted ACT score we added a random error term representing the residual variation of actual ACT scores around the predicted ACT scores; the resulting quantity was a simulated ACT score. We then calculated the percentage of the simulated ACT scores that met or exceeded the relevant College Readiness Benchmark.

We also calculated the percentage of students who would meet the College Readiness Benchmarks under alternative, somewhat less ambitious, assumptions about enhanced preparation:

- Meeting the EXPLORE Benchmark scores in each subject area. The EXPLORE Benchmark scores are the EXPLORE scores that students need to have a good chance of later meeting the ACT Benchmark scores. The EXPLORE Benchmark scores are: English-13, Mathematics-17, Reading-15, and Science-20.
- Taking the minimum recommended standard college-preparatory courses in the subject areas relevant to the model[11]
- Earning a B or higher grade average in each subject area relevant to the model

We refer to these scenarios as "Moderate effort."

Note that the increase in the percentage of students meeting a Benchmark score depends not only on the regression coefficients associated with a particular type of enhanced preparation, but also on the distribution of the predictor variables. For example, students whose values on the predictor variables cause their predicted ACT scores to fall just below the Benchmark could be

[11] The minimum recommended college-preparatory courses in mathematics are Algebra I, Geometry, and Algebra II. The minimum recommended college-preparatory courses in science are Biology and Chemistry. (Course work in English and course work in social studies were not statistically significant predictors in the models for ACT English and ACT Reading.)

pushed over the Benchmark with a particular type of enhanced preparation, even if the regression coefficient associated with that type of preparation is small. Conversely, students whose predicted ACT scores are already above the Benchmark cannot contribute toward increasing the percentage. Moreover, students whose predicted ACT scores are well below the Benchmark are unlikely to contribute toward increasing the percentage, even if the regression coefficients associated with a particular type of preparation are large.

An alternative approach to simulation for estimating the percentage of students who meet an ACT Benchmark is to estimate the percentage directly, using a nonlinear model. For example, one could estimate the probability of meeting an ACT Benchmark, given various characteristics, using a logistic regression model. In our experience, however, hierarchical nonlinear models are difficult to fit, especially with large numbers of predictor variables and with random slopes. For this reason, we chose to estimate the percentage of students who meet an ACT Benchmark using simulation based on linear models.

Underrepresented Minority Groups

We were interested in whether the relationships observed among the total group of students also pertained to underrepresented minority students (i.e., all students other than Caucasian-American/White or Asian American, Pacific Islander). We therefore re-estimated the main-effects models and the interaction models for this subset of the data. The minority-group models were based on data from 25,173 students at 2,723 high schools in the 2006 data set.

Limitations of the Data and Method

Representativeness of data. These data are not representative of all high school students in the U.S., especially because they mostly pertain to students who were considering attending college immediately after high school.[12] In general, the college-bound population is more academically able than the non-college-bound population. For example, in the two states where all eleventh-grade students in public schools take the ACT, the average ACT Composite score of students who plan to go to college is 1.2 standard deviations higher than the average Composite score of students who do not plan to go to college. The U.S. college-bound population is significant in its own right, however, because it includes about 67% of all high school graduates (U. S. Department of Education, National Center for Education Statistics, 2006)

The data in this study, from students who took the ACT, are also not completely representative of all college-bound students in the U.S. ACT-tested students are more concentrated in the central regions of the U.S., and less so on the east and west coasts. There is less variation in ACT scores by region, however, than by type of community, family income, and ethnicity. The data sets for this study encompass variation with respect to all three of these variables.

One can speculate about how our conclusions about the effectiveness of enhanced preparation might differ from conclusions about non-college-bound students, if data from this latter group could somehow be obtained. In all the simulations of enhanced preparation, the greatest possibility for increasing the percentage of students who meet the ACT Benchmarks occurs among students whose predicted ACT scores are just below the Benchmarks. Because most students in our data set already met the ACT English Benchmark, simulations done on non-college-bound students would likely have resulted in greater estimated increases in the percentage meeting the ACT English Benchmark than those reported here. On the other hand, most students did not meet the ACT Benchmarks in Mathematics, Reading, and Science.

---

[12] Exceptions are two states, Illinois and Colorado, in which all eleventh-grade students in public schools took the ACT in 2006.

Therefore, simulations done on non-college-bound students would likely have resulted in smaller estimated increases in the percentage meeting the ACT Benchmarks in these subject areas.

Reporting error in the predictors. The predictor variables in this study (background / demographic characteristics, course work, and course grades) were self-reported by students when they took EXPLORE, PLAN, and the ACT. Although self-reported information is not as accurate as information gathered from high school transcripts, we have found that students who take the ACT report their background characteristics, course work, and course grades with a reasonably high degree of accuracy (Laing, Sawyer, & Noble, 1988; Sawyer, Laing, & Houston, 1989). In the study by Sawyer et al., 71% of student-reported course grades were identical to those obtained from student transcripts, 97% were within one grade, and student-reported grades averaged 0.23 grade units higher (on a 0-4 scale) than transcript grades. Small systematic over-reporting of course work and grades is unlikely to affect the weights in the prediction models. Random misreporting of course work and grades, however, would make the weights smaller than they actually are.[13] We do not know the magnitude of random misreporting, and are therefore unable to incorporate it formally into our models.

Relatively large percentages of students reported taking Adv./Hon. course work: About 42% of students reported taking Adv./Hon. course work in English, and 38% reported taking Adv./Hon. course work in mathematics. These percentages suggest that students might have overstated their taking Adv./Hon. courses. If true, the most likely effect would be that the benefit of taking Adv./Hon. course work is underestimated. It is also possible that some students underreported taking Adv./Hon. course work. Underreporting could result from the fact that on

[13] Course grades are also affected by other sources of variation (e.g., student-teacher interactions, teachers' grading standards, and within-teacher measurement error).

the ACT registration form, the items related to Adv./Hon. course work are physically separated from the items related to standard course work.

In the future, we hope to analyze data obtained from official transcripts; models estimated from transcript data would not be subject to bias due to reporting error. Alternatively, we could compare the transcript data to the self-reported data to estimate the magnitudes of random reporting errors. We could then adjust the results using the estimated magnitudes of random reporting errors.

Potential extrapolation. The credibility of models with many predictor variables depends on, among other things, the extent to which data exist for different combinations of the predictor variable values. For example, course work and EXPLORE scores are predictors of ACT scores. Course work, however, is related to EXPLORE scores: Students with low EXPLORE scores are less likely to take upper-level courses than are students with high EXPLORE scores. In a hypothetical situation where no students with low EXPLORE scores took upper-level courses, inferences about the benefit of upper-level course work for students with low EXPLORE scores would be based purely on extrapolation.

One example we have investigated suggests that pure extrapolation does not occur, because of the large sample size. Although EXPLORE Mathematics score and the mathematics course work variable in this study are moderately correlated ($r = .41$), some students took upper-level mathematics courses, regardless of their EXPLORE scores. For example, among the 2,870 students with EXPLORE Mathematics scores of 10 (the 12th percentile), 188 eventually took Trigonometry or Calculus. Moreover, 372 of them took an Adv./Hon. course in mathematics.

Unobserved variables. In a linear model based on a randomized experiment (in which individuals are randomly assigned to treatment groups), the estimates of treatment main effects are unbiased, even if the model omits important predictor variables. The analyses in this study are based on observational data, rather than on data from a randomized experiment. Indeed, it is difficult to conceive of how students could be randomly assigned to enroll at particular schools, to take particular courses, and to earn particular grades.

An important consideration in interpreting analyses of observational data is that unobserved variables (i.e., variables not included in the models) could change the relationships between the outcome variables (ACT scores) and the treatment effects (e.g., course work and grades) if they could be included in the models. Our analyses statistically controlled for prior achievement (as measured by EXPLORE scores), high school attended, and a variety of background variables, all of which are related to ACT scores. Other variables could, however, change the relationships between ACT scores and course work and grades if they could be included in the models.

One way to assess the potential biasing effects of an unobserved variable u is through a sensitivity analysis, such as that proposed by Marcus (1997). Suppose T denotes a treatment, with T=1 corresponding to the treatment group, and T=0 corresponding to the control group. Marcus showed that the expected bias in the treatment effect resulting from not including covariate u in the model is:

$$\begin{aligned} Bias &= \gamma * \{ E[u|T=1] - E[u|T=0] \} \\ &= \gamma * \textit{Difference in } E[u|T], \end{aligned}$$

where γ is the regression coefficient for u. If the bias due to an unobserved variable u behaves like the bias due to the observed covariates $x_1,..,x_K$ in the model, then it should not exceed

$$max\{ b_i * \textit{Assumed difference in } E[x_i|T] \}, (i=1,...,K).$$

where *Assumed difference in* $E[x_i|T]$ is an assumed extreme difference (not the actual observed difference) between the treatment and non-treatment groups in the expected value of x_i. If *max*{ b_i * *Assumed difference in* $E[x_i|T]$ } is less than the regression coefficient for T, then Marcus would consider it implausible that the effect of the unobserved variable u dominates the treatment effect.

The most extreme assumption about *Difference in* $E[u|T]$ is that it is equal to the range of u (i.e., we assume that all the treatment group members have the maximum value of the unobserved covariate and that all the control group members have the minimum value). Less extreme assumptions are that *Difference in* $E[u|T]$ equals the midrange of u or the interquartile range of u. Leow, Marcus, Zanutto, and Boruch (2004) used the midrange assumption in a sensitivity analysis of a model of the effect of advanced course work on mathematics and science achievement. The data set they analyzed included a variety of background variables, but did not include a covariate measuring prior achievement. Because taking advanced course work is influenced by prior achievement, inferences from these data could be biased. Leow et al. showed that the biasing effect associated with the midrange of any observed background variable was smaller than the estimated treatment effect, and therefore concluded that prior achievement did not explain away the estimated treatment effect.

We did a sensitivity analysis of the main-effects models in this study, in a manner similar to that proposed by Leow et al. The various "treatment groups" were defined according to the scenarios in the first set of simulation studies:

- Increasing their EXPLORE scores in each subject area by two points
- Taking an additional standard college-preparatory course in each subject area relevant to the model
- Taking an Adv./Hon. course in each subject area relevant to the model

- Increasing grade average by one letter grade in each subject area relevant to the model

We used the background variables and the testing context variables as the observed covariates. For the dichotomous covariates, we assumed that the treatment and control groups would differ by 0.50 in the proportion of students belonging to one category or the other of the dichotomy. For the interval-scale covariates "Age at time of ACT testing" and "Family income," we took *Assumed difference in* $E[x_i|T]$ to be the interquartile range (the difference between the 75th and 25th percentiles).

Indirect effects. The models in this study summarize only the direct effects of various variables in predicting ACT scores. Some of the predictor variables could also have indirect effects. For example, EXPLORE scores are likely related to subsequent course work, which are, in turn, related to ACT scores. Therefore, increasing EXPLORE scores will likely increase ACT scores more than is predicted by the models in this study, because of their indirect effects through course work. In future research, we intend to estimate structural models that describe both the direct and indirect effects of the predictor variables.

Results

We used data from the 2005 graduating class to select variables for predicting the ACT English, Mathematics, Reading, and Science scores. We then re-estimated these models using data from the 2006 graduating class. Tables A-2 through A-5 in the appendix summarize the resulting coefficients, as estimated from the five imputation data sets for the 2006 data. The regression coefficients in Tables A-2 through A-5 are averages of the regression coefficients associated with the five imputation data sets for 2006.

In each table, there are two models, a main-effects model and an interaction model. For both models, there are estimated "fixed effects," which correspond to the values of the regression coefficients at typical high schools. Associated with each estimated fixed effect is an estimated standard error (SE), which measures the precision of the estimated fixed effect. The HLM software calculates each SE from the standard errors estimated from each imputation data set and from the variation of the estimated parameters across all five imputation data sets. From the estimated fixed effects and their SEs, the HLM software calculates significance levels, based on the usual *t* statistic.

Each table also contains estimated standard deviations of "random effects," which indicate the variability of some of the regression coefficients across high schools. In every model, there is a random effect associated with the intercept term. Each model also has random effects associated with some of the predictor variables; the particular predictor variables with random effects vary across models.

The variables listed in the left-most column of each table include the following:

- all potential main effects considered for the outcome variable, regardless of their statistical significance
- statistically significant ($p<.001$) interaction variables
- statistically significant ($p<.01$) high school characteristics used as predictors of the intercept.

The random effects shown in each table are those for which the estimated standard deviations are statistically significant ($p<.01$).

The symbol *NNN* in the tables indicates that an estimated coefficient was not statistically significant at the prescribed level in both the 2005 data and the 2006 data. The symbol *XXX* indicates that an estimated coefficient was statistically significant in the 2005 data, but not in the 2006 data. Among the four interaction models, the latter result occurred for only five of the 127 fixed effects and for four of the 19 random effects. This suggests that both the fixed effects and the random effects in the models were stable over time, but that the fixed effects were more stable than the random effects.

In predicting ACT English score and ACT Reading score, neither English standard course work nor social studies standard course work was statistically significant ($p<.001$), either in the 2005 data or in the 2006. Foreign language course work was statistically significant ($p<.001$) in predicting ACT English score, but not ACT Reading score. In contrast, both mathematics standard course work and science standard course work were statistically significant ($p<.001$) in predicting both ACT Mathematics score and ACT Science score.

Adv./Hon. course work in English and Adv./Hon. course work in social studies were statistically significant ($p<.001$) in predicting both ACT English score and ACT Reading score. Adv./Hon. course work in mathematics and Adv./Hon. course work in science were statistically significant ($p<.001$) in predicting both ACT Mathematics and ACT Science score.

Grade averages in English and social studies were statistically significant ($p<.001$) in predicting both the ACT English and ACT Reading scores. Grade averages in mathematics and science were statistically significant ($p<.001$) in predicting both the ACT Mathematics and ACT Science scores.

There were four variables related to the context of students' testing: age at time of testing, educational level at time of testing, a dummy variable identifying students who retested and updated their course work and grade information, and a dummy variable identifying students who retested but did not update their course work and grade information. All four testing context variables were statistically significant ($p<.001$) in all models except for ACT Mathematics.[14] The estimated coefficients indicate that:

- Students who test at an older age tend to score lower than students who test at a younger age (by 0.4 to 0.8 score points per year, depending on subject area).
- Students of a given age who test in grade twelve tend to score higher on the ACT English, Reading, and Science tests than students of the same age who test in grade eleven (by 0.4 to 0.9 score points).
- Students who retest and update their course work and course grade information tend to score higher than similar students who test only once (by 0.2 to 1.0 score points).
- Students who retest, but who do not update their course work and course grade information, tend to score lower than similar students who test only once (by 0.4 to 0.9 score units).

With regard to the last two results, the decision to retest is likely driven by initial test score, by a variety of psychosocial characteristics (such as goals, motivation, and academic discipline), by advice from others (counselors, friends, and parents), and by subsequent learning-related behavior (such as additional course work or review). The higher average scores of students who retest and update their course work and course grade information, and the lower average scores of students who retest but do not update their information, are also likely driven, both directly and indirectly, by these characteristics.

Relative Importance of Predictor Variables

The predictor variables in a regression model typically have different scales and are measured in different contexts. One way to assess the relative importance of the predictor

[14] ACT Mathematics score was not related to educational level, although it was related to the other three testing context variables.

variables is to compare their fixed effects after transforming all variables to have variance 1. The resulting standardized coefficient for each particular predictor variable (often called a "beta weight") indicates the expected change in the outcome variable, expressed in standard deviation units, given a one standard deviation increase in the particular predictor while holding the other predictors constant.[15]

Table A-6 in the appendix shows the magnitudes (absolute values) of the beta weights for the main-effects models. The sum of the magnitudes of the beta weights for a particular class of variables (e.g., grade averages) can be used as an indicator of the relative importance of the class. Table A-6 also shows the sums of the beta weight magnitudes for the different classes of predictor variables.

For all four ACT scores, prior educational achievement (as measured by EXPLORE scores) was much more important than any other class of predictor variables, including standard course work. Prior achievement was more strongly related to ACT English and Reading scores (beta weight sum = 0.74 and 0.72, respectively) than to ACT Mathematics and Science scores (beta weight sum = 0.54 and 0.64, respectively). Given that EXPLORE scores likely also affect ACT scores indirectly through course work and grades, the total effects of EXPLORE scores are larger than the direct effects reported here.

As one might expect, eighth-grade academic achievement tends to predict best eleventh/twelfth-grade academic achievement in the same subject area. For example, EXPLORE English scores (which measure writing skills in a multiple-choice format) more strongly predict ACT

---

[15] A limitation of beta weights in a data set such as this one in which the predictors are correlated, is that a change in one predictor (e.g., EXPLORE score) will result in a change in the predictors correlated with it (e.g., course work). A structural model, in which both direct and indirect effects were estimated, would overcome this limitation. See also Azen and Budescu (2003) for alternative measures of importance based on R^2. For simplicity, we have used beta weights in this analysis.

English scores than they predict ACT Mathematics, Reading, or Science scores. Nevertheless, EXPLORE scores in all four subject areas predict ACT scores in all four subject areas. In particular, EXPLORE English and Reading scores are weakly predictive of ACT Mathematics and Science scores, even given EXPLORE Mathematics and Science scores and the other predictor variables in the models.

The relative importance of the other classes of predictor variables depends on the ACT score being predicted. For predicting ACT English and Reading scores:

- ACT testing characteristics constitute the second most important class of predictor variables. These variables measure students' choices about when they take the ACT (age and grade at time of testing), whether they retest, and whether they update their course work and course grade information.
- Subject-area grade averages and Adv./Hon. course work are about equally important behind ACT testing characteristics.
- Background variables are somewhat less important than subject-area grade averages and Adv./Hon. course work.
- Standard course work is the weakest class of predictor variables. Indeed, English and social studies course work are not even in the models for ACT English and ACT Reading scores[16], and foreign language course work has only a weak relationship with ACT English score.

[16] High schools typically require their students to take English courses and certain social studies courses. For this reason, English and social studies course work are strongly related to educational level, which is in the models for both ACT English and ACT Reading.

For predicting ACT Mathematics score:

- Background characteristics and subject-area grade averages (in mathematics and science) are about equally important, behind EXPLORE scores.
- ACT testing characteristics, standard course work, and Adv./Hon. course work are about equally important, behind background characteristics and subject-area grade averages.

For predicting ACT Science score:

- Background variables and ACT testing characteristics are about equally important, behind EXPLORE scores.
- Standard course work and subject-area grade averages (in mathematics and science) are about equally important, behind background variables and ACT testing characteristics.
- Adv./Hon. course work is somewhat less important than standard course work and subject-area grade averages.

It is interesting to compare our results in predicting ACT Mathematics score to those obtained by Burkam and Lee (2003) in predicting NELS:88 mathematics scores. Using different data and modeling techniques, Burkam and Lee obtained beta weights of 0.54, 0.32, and 0.11, respectively, for prior achievement, mathematics course work, and mathematics grades. Our results for prior achievement (0.54), course work (0.28=0.16+0.12), and course grades (0.16) are similar to Burkam and Lee's results.

Carbannaro (2005) predicted mathematics achievement in grade ten from achievement in various subject areas in grade eight, from honors track course work, and from other variables. He reported standardized coefficient sums of 0.80 for grade eight achievement and 0.11 for honors track course work. This ordering is consistent with our result of 0.54 and 0.12 for EXPLORE scores and Adv./Hon. course work, respectively.

The results suggest that taking standard and Adv./Hon. mathematics and science courses improves ACT Mathematics and Science scores, but that taking English, social studies, and foreign language courses is of little or no benefit in improving ACT English and Reading scores. Earning higher grades in standard courses and taking Adv./Hon. courses do provide modest benefit. Given the strong relationship between EXPLORE scores and ACT English and Reading scores, however, major improvements in reading and writing need to occur *before* grade eight.

Interaction Models

Tables A-2 through A-5 in the appendix also show the interaction models for predicting ACT scores. The interaction models permit us to determine whether the benefit of enhanced preparation (e.g., additional course work) is roughly the same for all students, or whether it depends on other characteristics (e.g., students' EXPLORE scores).

Although the particular interaction terms differ by the ACT score they predict, their coefficients all indicate that students with high EXPLORE scores benefit more from standard or Adv./Hon. course work than do students with low EXPLORE scores. For example, the prediction model for ACT Reading (see Table A-4) includes the following interaction terms related to EXPLORE Reading score:

- EXPLORE Reading score *x* Adv./Hon. social studies course work
- EXPLORE Reading score *x* Social studies grade average

Table 2 shows the expected increase in ACT Reading score from taking an Adv./Hon. social studies course or from increasing social studies grade average. Note that the expected increase resulting from either form of enhanced preparation depends on EXPLORE Reading score.

TABLE 2
Expected Increase in ACT Reading Score from Enhanced Preparation, Given EXPLORE Reading Score

	Enhanced preparation	
EXPLORE Reading Score	Take Adv./Hon. social studies course	Raise social studies grade average by one letter grade
25	0.89	0.96
20	0.70	0.71
15	0.51	0.45
10	0.33	0.19

The models also include interaction terms for course work by subject area grade average. They indicate that students who earn high grades in particular subject areas benefit more from taking courses in those areas than do students with low grade averages. For example, the model for ACT Reading includes the interaction term ADV_ENG *x* ENG_AV. The coefficient for this term (0.280377) suggests that the average benefit associated with taking an Adv./Hon. course in English is approximately 0.28 higher for students whose English grade average is a 4.0, than for students whose English grade average is 3.0.

The models also include various interactions among the variables related to testing context. The interactions vary with the ACT score outcome variable. For example, the increase in predicted ACT English score of students who retested and updated their course work and grade information was larger for students who tested last in grade twelve than for students who tested last in grade eleven (positive coefficient for the interaction of EDLVL by RETEST2 in Table A-2). The increase in predicted ACT Mathematics score of students who retested and updated their course work and grade information was slightly smaller for older students than for younger students (negative coefficient for the interaction of AGE by RETEST2 in Table A-3).

Random Effects

All of the models have random effects associated with the intercept term (see the bottom of Tables A-2 through A-5). These random effects reflect variation in high schools' average achievement for reasons other than differences in students' background characteristics, prior achievement, course work, course grades, testing context, or differences in the school-level predictor variables.

For example, the standard deviation of the intercept term in the interaction model for predicting ACT Mathematics score is approximately 0.8, and the fixed effect for the intercept term is approximately 20.2 (see Table A-3). The fixed effect of 20.2 can be interpreted as the average ACT Mathematics score at a typical high school, and the standard deviation of 0.8 reflects the variation across high schools in the average ACT Mathematics scores adjusted for differences in the predictor variables. Given the assumption of the hierarchical linear model that random effects are approximately normally distributed, we would expect about 16% of the schools to have adjusted average ACT Mathematics scores less than 20.2 - 0.8 = 19.4, and about 16% to have adjusted average ACT Mathematics scores greater than 20.2 + 0.8 = 21.0.

All models also have random effects associated with one or more of the predictor variable slopes. The random effects indicate that the benefit of taking additional standard courses, taking Adv./Hon. course work, or earning higher grades varies among high schools. In some models, there are also random effects associated with background characteristics and testing context. The models for predicting ACT English and ACT Mathematics scores have more random slopes than do the models for predicting ACT Reading and ACT Science scores.

The standard deviations of the slopes are sometimes large relative to the corresponding average slopes. In the main-effects model for predicting ACT English score, for example, the average slope for ADV_FL is approximately 0.22, and the standard deviation of the ADV_FL slope across high schools is approximately 0.43 (see Table A-2). This result suggests that at some high schools, taking Adv./Hon. foreign language courses does not increase ACT English scores. (Indeed, within-school analyses revealed that at a few schools, taking Adv./Hon. foreign language courses was associated with *lower* ACT English scores.) While this is a disturbing result for the schools affected, it does not prove that their curricula, teachers, or administrators are ineffective; the negative relationship could be due to characteristics beyond the control of the schools.

Although the variation in the intercept and slope coefficients across schools is likely related in some way to the schools' effectiveness, we do not have a detailed explanation of this result. A detailed explanation would require identifying school-level characteristics that predict the intercept and the slopes associated with course work and grades. Although investigating school-level characteristics (other than means of Level-1 variables) is beyond the scope of this preliminary study, we hope to do so in the future. Examples of studies that have done this are: Raudenbush and Bryk (1986), Lee and Bryk (1989), Caldus and Bankston (1997), Kreft and de Leeuw (1998), Ma and McIntyre (2005), and Noble and Schnelker (2007).

Model Diagnostics

One of the assumptions of the hierarchical linear model is that the residuals for students within schools have a normal distribution. To study the plausibility of this assumption, we plotted the quantiles of the standardized Level-1 residuals against the corresponding quantiles of a normal distribution. The plots showed that the quantiles of the standardized Level-1 residuals and the normal quantiles corresponded closely for nearly all students. For only a tiny percentage of cases was there a moderately large difference between the two quantiles (0.5 or greater): ACT English (<0.1%), ACT Mathematics (<0.2%), ACT Reading (<0.2%), and ACT Science (<0.3%).

Scatter plots of the Level-1 residuals against predicted ACT scores and against key predictor variables revealed no apparent relationships. This result suggests no obvious inadequacy in a linear model.

Another assumption of the hierarchical linear model is that the random effects associated with high schools have a multivariate normal distribution. To study the plausibility of this assumption, we first plotted the quantiles of the standardized Level-2 residuals from the estimated models against the corresponding quantiles of the standard normal distribution. We examined quantiles of the Level-2 residuals for the random intercept and random slopes in each model.

The plots for the intercept showed that the quantiles of the standardized residuals corresponded closely to the quantiles of the standard normal distribution. For only a very small percentage of outlier high schools was there a moderately large difference (0.5 or greater) between the two quantiles: ACT English (1%), ACT Mathematics (<0.5%), ACT Reading (<0.5%), and ACT Science (<0.5%).

For most high schools, the quantiles of the standardized residuals associated with the random slopes also corresponded closely to the quantiles of the standard normal distribution. There were, however, more schools with a moderately large difference (0.5 or greater) in the quantiles of the random slope residuals: ACT English (7%), ACT Mathematics (4%), ACT Reading (2%), and ACT Science (5%).

To examine the plausibility of the multivariate normality assumption for the joint distribution of the random effects, we plotted the Mahalanobis distance against the corresponding quantiles of the chi-square distribution (Raudenbush & Bryk, 2002). The plots revealed that virtually all the high schools fell on the diagonal, but that a handful of outlier schools had very large discrepancies (e.g., differences greater than two standard deviations of the chi-square distribution).

The estimates of fixed effects in the hierarchical model are unbiased, even if the multivariate normality assumption is incorrect (Raudenbush & Bryk, 2002). The validity of hypothesis tests about the fixed effects could be affected, however, by the failure of this assumption. To diagnose this potential problem, the HLM software calculates robust estimated standard errors that are consistent even if the model assumptions are incorrect. The standard errors reported in the Appendix are robust standard errors. The median difference between model-based standard errors and the robust standard errors (as a proportion of the robust standard errors) was 0.04, 0.07, 0.05, and 0.04 for the ACT English, Mathematics, Reading, and Science score models, respectively. This result suggests that the partial failure of the multivariate normality assumption due to outlier high schools had a minor effect on inferences about the fixed effects.

Model Fit

The variation of observed ACT scores about the regression surfaces defined by the fixed and random effects is an indicator of model fit. The residual variances for the interaction models were 9.15, 5.70, 13.27, and 7.77 for ACT English, Mathematics, Reading, and Science, respectively.

Another index of model fit is Level-1 R^2:

$$R^2 = 1 - \frac{\text{Res.var.(Model)}}{\text{Res.var.(Random ANOVA)}}.$$

R^2 is the decrease in residual variance relative to a random-effects ANOVA model (in which only the school mean is included in the model). It can be interpreted as the proportion of variance in the outcome variable explained at Level 1. The Level-1 R^2 statistics for the interaction models are: .67 (ACT English), .70 (ACT Mathematics), .55 (ACT Reading), and .54 (ACT Science). Therefore, as judged by the R^2 statistic, the models for ACT English and ACT Mathematics are better fitting than the models for ACT Reading and ACT Science. Some of the differences in R^2 values are likely due to differences in the reliabilities of the four ACT scores: English (.91); Mathematics (.91); Reading (.85); and Science (.80).

Underrepresented Minority Groups

Fixed effects. We estimated both main-effects and interaction models for the subset of underrepresented minority students. In the main-effects models, the sums of the absolute values of the beta weights for all classes of predictor variables (except for background characteristics[17]) corresponded closely to those calculated for the total group. For example, the beta-weight sums associated with course work and course grades for predicting ACT Mathematics score were 0.16

[17] The beta weight sums associated with background characteristics were lower for the minority students because the ethnicity variable was constant in the minority group subset.

and 0.17, respectively, for the underrepresented minority group, as compared to 0.16 and 0.16, respectively, for the total group. Therefore, the relative importance of the predictor variable classes for minority students was the same as for the total group.

Nearly all of the coefficients for the interactions of EXPLORE scores with course work and course grade variables that we estimated in the total group models were also statistically significant in the minority group models. Therefore, prior achievement had similar effects on the benefit of enhanced preparation of underrepresented minority students as for the total group of students.

Random effects. We were able to detect fewer random effects in the minority group models than in the total group models. The minority group interaction models for the ACT English, Mathematics, Reading, and Science scores had 3, 6, 2, and 1 random effects, respectively, as compared to 7, 7, 4, and 4 random effects, respectively, for the total group interaction models. (In all the models, one of the random effects was associated with the intercept.) One likely reason for this result is that the minority-group models were based on data from only 2,723 high schools, as compared to 4,638 high schools for the total-group models.

Simulation Study Results

The College Readiness Benchmarks are the minimum ACT scores at which students are likely to succeed in beginning college-level courses at typical postsecondary institutions. The Benchmark scores are: English-18, Mathematics-22, Reading-21, and Science-24. Table 3 and Figure 1 on pages 50-51 show the percentage of students who would meet the ACT College Readiness Benchmark scores under various scenarios[18]:

- Current reported course work and grades

[18] The percentages of students represented in the analysis file for this study who meet the ACT College Readiness Benchmarks differ from the percentages of all ACT-tested students who meet the Benchmarks, because the two groups of students are different.

- Higher EXPLORE scores
- Additional standard course work
- Adv./Hon. course work (if not currently taken)
- Higher grades

The particular standard course work, Adv./Hon. course work, and subject-area grade averages in Table 3 and Figure 1 are those associated with the predictor variables in the interaction model for each ACT subject area.

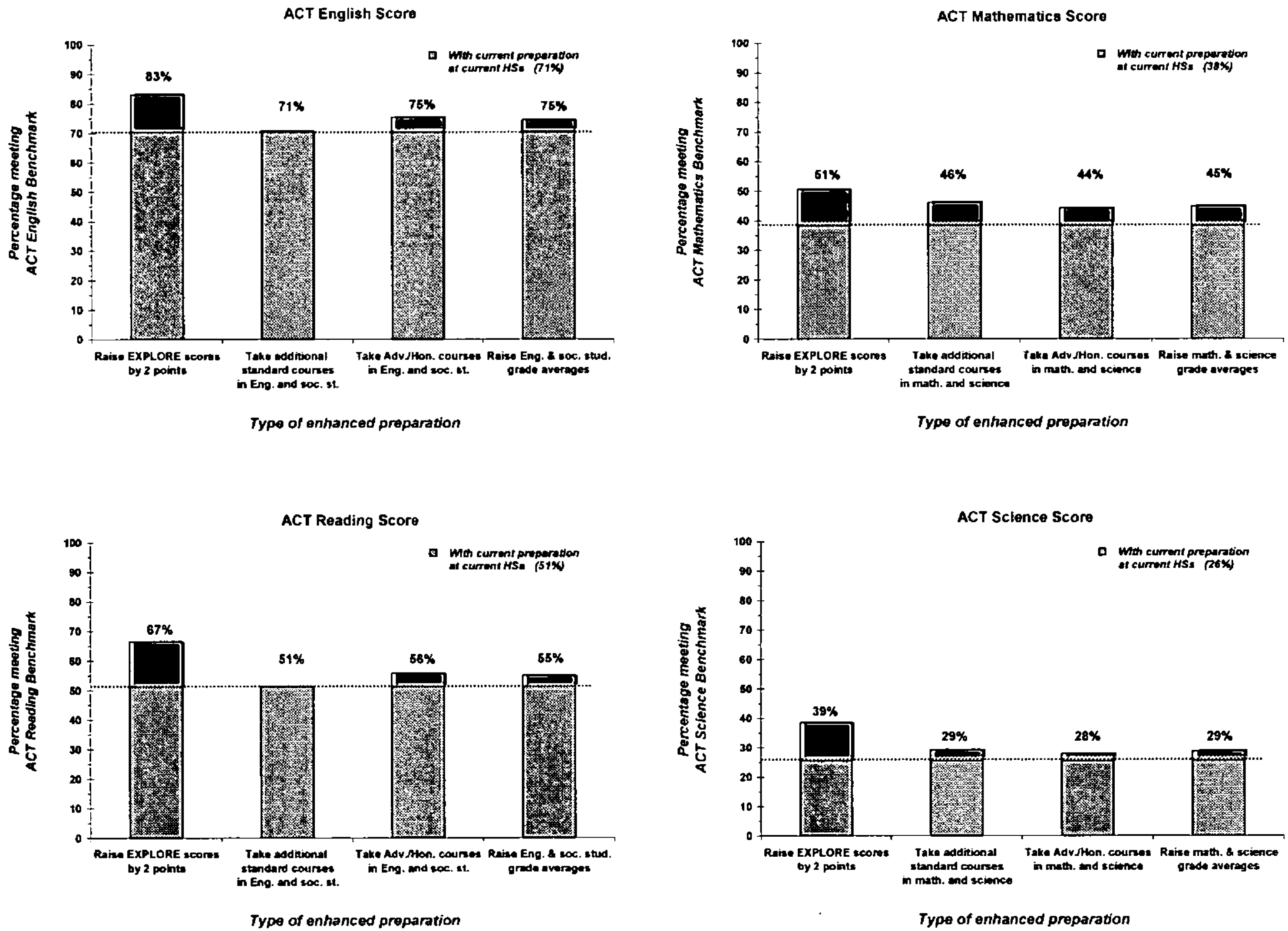
Table 3 shows results for two sets of scenarios of enhanced preparation, defined by the effort required on the part of students. In the first set of scenarios (labeled "High Effort"), all students increase each of their EXPLORE scores by two points, or they take an additional standard course in each relevant subject area, or they increase their grade averages in the relevant subject areas by one letter grade. The results for the High Effort scenarios are also shown graphically in Figure 1.

The second set of scenarios in Table 3 is less ambitious. In them, all students only need to meet the EXPLORE Benchmarks in each subject area, or meet the minimum recommended college-preparatory course work in each relevant subject area, or maintain a B or higher grade average in each relevant subject area.

TABLE 3
Percentage of Students Meeting the ACT College-Readiness Benchmarks at Current High Schools, by ACT Subject Area, Current Preparation, and Type of Enhanced Preparation

ACT subject area	Current preparation	Enhanced preparation effort	Enhanced preparation activity: Higher EXPLORE scores	Additional standard course work	Adv./Hon. course work	Higher subject-area grade averages
English	71	High	Increase EXPLORE scores by 2 points: 83	Not applicable	Take Adv./Hon. courses in English and social studies 75	Increase grade averages in English and social studies by 1 letter: 75
		Moderate	Meet EXPLORE Benchmarks: 83	Not applicable	Not applicable	Maintain B or higher grade average in English and social studies 72
Mathematics	38	High	Increase EXPLORE scores by 2 points: 51	Take one additional course in mathematics and natural science 46	Take Adv./Hon. courses in mathematics and natural science 44	Increase grade averages in mathematics and natural science by 1 letter: 45
		Moderate	Meet EXPLORE Benchmarks: 47	Take mathematics and natural science courses through Alg. II and Chemistry 39	Not applicable	Maintain B or higher grade average in mathematics and science 40
Reading	51	High	Increase EXPLORE scores by 2 points: 67	Not applicable	Take Adv./Hon. courses in English and social studies 56	Increase grade averages in English and social studies by 1 letter: 55
		Moderate	Meet EXPLORE Benchmarks: 63	Not applicable	Not applicable	Maintain B or higher grade average in English and social studies 52
Science	26	High	Increase EXPLORE scores by 2 points: 39	Take one additional course in mathematics and natural science 29	Take Adv./Hon. courses in mathematics and science 28	Increase grade averages in mathematics and natural science by 1 letter: 29
		Moderate	Meet EXPLORE Benchmarks: 33	Take mathematics and natural science courses through Alg. II and Chemistry 26	Not applicable	Maintain B or higher grade average in mathematics and science 26

FIGURE 1. Percentage of students meeting the ACT College-Readiness Benchmarks at current high schools, by ACT subject area, current preparation, and type of high-effort enhanced preparation



Among the forms of enhanced preparation, increasing EXPLORE scores would result in the greatest increase in the percentage of students meeting the ACT College Readiness Benchmarks. Increasing EXPLORE scores by two points yields increases of 12, 13, 16, and 13 percentage points in meeting the English, mathematics, reading, and science Benchmarks, respectively. If all students achieved the less ambitious goal of meeting the EXPLORE Benchmarks, the percentage increases would be 12, 9, 12, and 7. Note that increasing EXPLORE scores would increase the percentage of students who would meet the ACT English Benchmark, even though 71% of students already met the Benchmark with current preparation.

Taking additional standard courses would result in only a modest increase in the percentage of students meeting the ACT Benchmarks. Standard course work in English and social studies is not even in the models for predicting ACT English and Reading scores. Taking an additional mathematics and science course would increase the ACT Mathematics Benchmark rate by 8 percentage points. Taking an additional mathematics and science course would increase the ACT Science Benchmark rate by only 3 percentage points. Thus, taking additional course work does not necessarily lead to sufficient college readiness for many students. This overall result is consistent with that reported in *Rigor at Risk: Reaffirming Quality in the High-School Curriculum* (ACT, 2007c).

Taking Adv./Hon. courses would also result in only a modest increase in the percentage of students meeting the ACT Benchmarks. The increases are 4, 6, 5, and 2 percentage points, respectively, for the four subject areas.

We obtained a similar result for higher subject-area grade averages. Increasing relevant subject-area grade averages by one letter grade would increase the ACT Benchmark rates by 4, 7, 4, and 3 percentage points, respectively.

The results for the moderate-effort enhanced preparation scenarios follow the same pattern. As one would expect, of course, the estimated percentages of students meeting the ACT Benchmarks are lower than the percentages for the corresponding high-effort enhancements.

The general result is that taking additional standard course work, taking Adv./Hon. course work, or increasing subject-area grade averages only modestly increases the percentage of students who would meet the ACT Benchmarks. A likely reason for this result is that enhanced preparation can raise the Benchmark rates only by increasing the ACT scores of students whose predicted ACT scores are just below the Benchmarks. Increasing the ACT scores of students whose predicted ACT scores already meet the Benchmarks, or whose predicted ACT scores are much lower than the Benchmarks, will not affect the percentage meeting the Benchmarks. Moreover, the interaction terms in the models imply that the principal benefits of enhanced preparation accrue to students who already have high ACT scores.

Although there are only modest increases in the ACT Benchmark rates associated with taking additional standard course work, taking Adv./Hon. course work, or increasing subject-area grade averages *individually,* doing all of these enhancements *together* would add up to a larger increase. For example, all three enhancements done together would increase the proportion of students meeting the ACT Mathematics Benchmark from 38% to approximately 49% (not shown in Table 3). One would question, however, the feasibility of students' accomplishing all three enhancements (particularly if they start with below-average prior achievement).

Enhanced preparation assuming improved high schools. As was noted previously, students at some high schools have higher average ACT scores than at others, for reasons other than the students' background characteristics, prior achievement, testing context, standard course work, advanced course work, or subject-area grade averages. Furthermore, standard course work, advanced course work, and subject-area grade averages are more strongly related to ACT scores at some high schools than at others. The estimated random effects for the intercepts and slopes in the hierarchical models indicate these among-school differences.

We did simulations similar to those summarized in Table 3, but with the additional assumption that a below-average intercept or slope for any school could be brought up to that of a typical school. Note that some of the differences among high schools might be due to characteristics that they can control (e.g., curriculum and instruction), but some of the differences might be caused by factors over which they have no control (e.g., neighborhood SES). Therefore, the second set of simulations provides *upper bounds* for the ACT Benchmark rates that could result from schools' improving the characteristics they can control.

The results of the second set of simulations, shown in Table 4, are typically 1 or 2 percentage points higher than the corresponding results in Table 3. This pattern suggests that improving underperforming high schools would not significantly increase the proportion of students who are ready to take first-year college courses. One possible reason for this result is that the ACT scores of students attending underperforming high schools are not close enough to the Benchmarks for potential improvements in the high schools to push the students' scores over the Benchmarks. Another reason, stated previously, is that the interaction terms in the models imply that the benefits of enhanced preparation accrue mostly to students who already meet the ACT Benchmarks.

TABLE 4
Percentage of Students Meeting the ACT College-Readiness Benchmarks at Improved High Schools, by ACT Subject Area, Current Preparation, and Type of Enhanced Preparation

ACT subject area	Current preparation	Enhanced preparation effort	Enhanced preparation activity: Higher EXPLORE scores	Additional standard course work	Adv./Hon. course work	Higher subject-area grade averages
English	72	High	Increase EXPLORE scores by 2 points: 84	Not applicable	Take Adv./Hon. courses in English and social studies 77	Increase grade averages in English and social studies by 1 letter: 76
		Moderate	Meet EXPLORE Benchmarks: 84	Not applicable	Not applicable	Maintain B or higher grade average in English and social studies 73
Mathematics	40	High	Increase EXPLORE scores by 2 points: 52	Take one additional course in mathematics and natural science 48	Take Adv./Hon. courses in mathematics and natural science 46	Increase grade averages in mathematics and natural science by 1 letter: 47
		Moderate	Meet EXPLORE Benchmarks: 48	Take mathematics and natural science courses through Alg. II and Chemistry 41	Not applicable	Maintain B or higher grade average in mathematics and science 42
Reading	52	High	Increase EXPLORE scores by 2 points: 68	Not applicable	Take Adv./Hon. courses in English and social studies 57	Increase grade averages in English and social studies by 1 letter: 56
		Moderate	Meet EXPLORE Benchmarks: 64	Not applicable	Not applicable	Maintain B or higher grade average in English and social studies 53
Science	26	High	Increase EXPLORE scores by 2 points: 39	Take one additional course in mathematics and natural science 30	Take Adv./Hon. courses in mathematics and science 29	Increase grade averages in mathematics and natural science by 1 letter: 29
		Moderate	Meet EXPLORE Benchmarks: 34	Take mathematics and natural science courses through Alg. II and Chemistry 27	Not applicable	Maintain B or higher grade average in mathematics and science 27

Sensitivity Analysis

Table 5 shows the potential sensitivity of the four main-effects models to unobserved covariates. The "maximum plausible bias" is the maximum, over all the observed covariates, of the product of each covariate's coefficient and its assumed extreme difference. The observed covariates included the background variables (gender, ethnicity, family income, parents' educational levels, and primary language at home) and the testing context variables (age and educational level at time of testing, retesting, and updating of course work and course grade information). The "covariate name" in Table 5 is the covariate whose product is a maximum. The "treatment effects" were calculated from the coefficients of the treatment variables.

TABLE 5
Potential Sensitivity of Main-Effects Models to Unobserved Covariates

			Treatment effects			
ACT subject area	Maximum plausible bias	Covariate name	Increase EXPLORE scores 2 pts.	Take addtl. standard courses	Take Adv./Hon. courses	Increase subj-area grd. avgs.
English	0.61	AGE	2.33	n/a	1.33	0.93
Mathematics	0.57	GENDER	1.58	0.95	1.23	1.02
Reading	0.47	RETEST1	2.37	n/a	1.14	0.91
Science	0.59	GENDER	1.72	0.57	0.70	0.67

Note: RETEST1=1 if a student took the ACT more than once, but did not update course work and course grade data.

The results indicate that the effects associated with increasing each EXPLORE score by 2 points, taking Adv./Hon. course work, and raising subject-area grade averages by one letter grade are robust to the plausible effects of unobserved covariates. For predicting ACT Mathematics scores, the effect of additional standard course work is also robust to the plausible effects of unobserved covariates. The effect of additional standard course work is not robust in the ACT Science model, however.

In our opinion, the unobserved variables that most likely could change the model are psychosocial characteristics (e.g., motivation, self-discipline) and behavioral variables other than course work (e.g., attendance, doing homework, conforming to rules). Variables such as these likely affect students' choices of course work, their grades, and their testing context variables, as well as their test scores. The effects of psychosocial characteristics and behavioral variables likely are both direct and indirect.

Implications for Preparing Students for College

We investigated activities that could improve high school students' academic readiness for postsecondary education. Our simulation study suggests that taking additional standard college-preparatory courses in high school, taking advanced/honors courses, and earning higher grades would, by themselves, only modestly increase the percentage of students who leave high school adequately prepared to take credit-bearing courses in the first year of college. Moreover, taking additional courses and earning higher grades mostly benefit students who by grade eight are already well "on-target" in preparing themselves for college. Our results also suggest that improving high schools would only modestly increase the proportion of students who are adequately prepared to take college-level courses. Among the enhancements we studied, the only one that appears likely to result in a substantial increase is to increase students' academic skills by eighth grade. These results pertain to underrepresented minority students as well as to the total group of students.

Taking rigorous courses in high school, earning higher grades in these courses, and improving high schools are all important in their own right. All of these enhancements would improve students' academic preparation by the time they graduate from high school. Given current levels of preparation in grade eight, however, these enhancements would not individually suffice to increase significantly the proportion of students who are adequately prepared to take college-level courses. Accomplishing this goal will require improving students' academic skills before eighth grade. Furthermore, if we could improve students' academic skills before eighth grade, then the other enhancements would be even more effective.

In this section, we briefly discuss implications of these results for bringing the academic skills of more students to a level where they can succeed in college-level courses. The discussion is organized by grade level and age of students: after grade eight, during elementary school, and during early childhood. Following this discussion, we comment on likely important effects of psychosocial characteristics and behavior.

Implications for Underprepared Students after Grade Eight

An alternative to taking standard college-preparatory courses in high school is to take remedial courses. Balfanz, Legters, and Jordon (2004) investigated a program of remedial instruction in reading and mathematics in grade nine in four large urban school districts. After adjusting for various covariates, they found gain score beta weights of 0.28 in reading and 0.18 in mathematics. Similarly, the on-line Best Evidence Encyclopedia (BEE) found small to moderate effects among the studies that met their evidence standards (Slavin, Lake, & Groff, 2007; Slavin, Cheung, Groff, & Lake, 2007).

To interpret the results of Balfanz et al. in the context of this study, consider the EXPLORE scores of a nationally representative sample of students (ACT, 2007b). The average and standard deviation of the EXPLORE Reading score in this sample were 13.8 and 3.7; the corresponding statistics for the EXPLORE Mathematics score were 15.1 and 4.0. Suppose we could increase the average Reading and Mathematics scores by 0.28 and 0.18 standard deviation, respectively, through massive remedial instruction of underprepared students. Accomplishing this goal would raise the average EXPLORE Reading score to 14.8, and the average EXPLORE Mathematics score to 15.8. Both of these hypothetical average scores are still below the EXPLORE Reading and Mathematics Benchmark scores of 15 and 17, respectively, that we associated with the "Moderate effort" scenario of enhanced preparation. Although remedial instruction could benefit students with academic deficiencies, most of them would still not be on target in preparing for college.

It appears that a significant proportion of current eighth-grade students have academic deficiencies, and that neither additional college-preparatory course work nor remedial instruction will prepare them for college or high-paying jobs by the time they graduate from high school. Policy makers need to devise educational and social policies that accommodate these students. For example, given the difficulty of making substantial large-scale improvements in students' achievement, it is likely that the need for remedial instruction in college will persist for some time. According to the U.S. Department of Education (Parsad & Lewis, 2004), 28% of first-year college students in fall 2000 enrolled in reading, writing, or mathematics remedial courses.

Implications for Students in Elementary School

A principal result of this research is that achievement in grade eight is very important in driving subsequent achievement. One obvious strategy for ultimately reducing the number of academically underprepared high school graduates is to increase diagnosis and remediation of academic weaknesses well before grade eight. Another strategy is to implement curricula and instructional techniques that are more effective for all students.

The BEE Web site provides bibliographies and ratings of research related to elementary school mathematics (Slavin & Lake, 2007). BEE typically found modest effect sizes (approximately 0.10) for different mathematics curricula; there were stronger effect sizes (0.20 or higher), however, for several different instructional process programs. The What Works Clearinghouse (WWC) Web site (2007a) provides bibliographies of research and ratings of programs for both beginning reading and elementary school mathematics. WWC found evidence of large gains in percentile rank for one of the reading interventions and small to moderate gains in several others. WWC found evidence of small gains in mathematics interventions in elementary school.

The most significant funding stream for remedial instruction in K-12 education is Title I-A (Improving Basic Programs)[19] of the federal Education and Secondary Act of 1965 and the No Child Left Behind Act of 2001 (U.S. Department of Education, 2002). In FY 2008, it is expected to provide nearly $14 billion to assist state and local education agencies to meet the needs of low-achieving students enrolled in schools with high concentrations of poverty (U.S. Government Accountability Office, n.d.). Although education agencies and schools may use Title I funds to serve children from preschool age through high school, about three-fourths of Title I participants are in preschool through grade six. Title I funds are most commonly used to support instruction in reading and mathematics.

It has been difficult to measure the effectiveness of Title I programs because of variation in their designs, the complicating effects of other variables (including changes in the law and in

[19] Although other programs are funded under Title I (e.g., Reading First), Title I-A is the largest, and is sometimes referred to simply as "Title I."

implementation of programs), and the near ubiquity of participation by eligible schools (Raudenbush, 2002). Borman (2000) found a modest typical effect size (0.15) for reading skills instruction during the period 1963 – 1995. In a study of New York schools during the years 1993, 1997, and 2001, van der Klaauw (2005) found no benefit for reading and mathematics achievement.

Two other significant federal funding streams are Reading First (Title I-B-1) and Early Reading First (Title I-B-2). The purpose of Reading First is to ensure that children can read at grade level or above by the end of third grade; the purpose of Early Reading First is to prepare young children to enter kindergarten with the language, cognitive, and early reading skills needed to succeed in reading. Reading First is a formula grant program; Early Reading First is a competitive grant program.

There has been controversy over how Reading First was implemented by federal officials and consultants, with charges of interference in state and local decision making and of favoritism toward particular reading programs. As a result, the U.S. Congress reduced funding for the program by over 60% for the fiscal year that began in October 2007 (Manzo, 2008, Jan. 16). There was early positive anecdotal evidence on the effectiveness of Reading First, but a recently released interim study on the effectiveness of Reading First did not yield clear results (Manzo, 2008, June 4).

It appears therefore that applying certain new curricula or teaching methods in elementary school could result in small to moderate improvements in reading and mathematics achievement. In particular, applying appropriate interventions to underachieving students could improve their educational achievement by grade eight, and therefore, their eventual readiness for college and work. There is, however, only meager and inconsistent evidence on the effectiveness of the Title

I and Reading First funding streams in improving elementary school students' educational achievement.

Implications During Early Childhood

Many authorities believe that academic achievement in school is strongly influenced by cognitive skills in early childhood. For example, Duncan et al. (2007) found that pre-mathematics skills at entry to kindergarten are an important predictor of achievement in both mathematics and reading between ages seven and fourteen. (They also found that pre-reading skills predict future achievement, but less so than pre-mathematics skills.) Given our result that achievement in grade eight strongly predicts achievement in high school, the findings of Duncan et al. suggest that efforts to improve achievement in high school might need to begin in early childhood.

Many authorities also believe that efforts to improve later achievement in school must attend to young children's health needs and psychosocial development, as well as to their cognitive development (Schweinhart, Barnes, & Weikart, 1993; Reynolds et al., 2007). They also advocate involving parents closely in their children's education. Advocates of "whole-child" programs believe that the programs result not only in long-term gains in cognitive skills, but also in greater long-term success in society (e.g., higher graduation rates from high school, fewer arrests).

The WWC Web site (2007a) web site provides a bibliography of research related to many different early childhood education programs. The web site also identifies early childhood programs for which the Clearinghouse believes there is strong evidence of benefit.

The principal funding stream for early childhood education is the federal Head Start program (U.S. Department of Health and Human Services, 2007, August 15). In fiscal year

2006, Head Start served more than 900,000 children, 45% of whom were age three or younger. Its budget of more than $6.5 billion funded more than 1,600 programs with a paid staff of more than 200,000 and more than six times as many volunteers. The U.S. Department of Health and Human Services (2005) has undertaken a large-scale study of the benefits of participation in Head Start programs. Preliminary results from the first year of data collection indicated small to moderate benefits in cognitive skills, behavior, and health, as well as parents' behavior, for the children participating in Head Start programs.

The U.S. Government Accountability Office (2008) has criticized several aspects of the management of Head Start. Its report recommended that the program administrators develop a more strategic approach to assessing risks, expand efforts to collect data on and estimate improper payments, improve the accuracy of data reported by grantees, and develop clear criteria for providing assistance to high-risk grantees.

Psychosocial Characteristics and Behavior

Psychosocial characteristics such as motivation and social connectedness, and behavioral variables such as non-use of drugs, attendance, and obedience of rules, are important predictors of academic success in middle school and high school (Rumberger, 1995; Worrell & Hale, 2001; Jones & Byrnes, 2006; see also footnote 3). If educators could intervene effectively with young children whose psychosocial characteristics and behavior predict high risk of academic failure, they could be set on a course by which they could eventually benefit from a rigorous curriculum in high school. Findings from clinical psychology suggest that interventions for improving psychosocial characteristics and behavior are more effective for young children than for adolescents (Dadds & Fraser, 2003; Dunn & Mezzich, 2007).

In a preliminary report, the Collaborative for Academic, Social, and Emotional Learning (CASEL, 2007) summarized the results of a meta-analysis of research on 207 programs that promote positive youth development in school, family, or community settings. CASEL reported that among the studies evaluating academic outcomes, students who participated in a program designed to enhance social and emotional learning scored 11 percentile points higher on standardized achievement tests relative to peers who did not participate in such a program.

The WWC (2007b) rates the effectiveness of character education programs (activities and experiences organized to foster positive character development and associated core ethical values). The Clearinghouse assigns ratings on programs' effectiveness in various domains, including academic achievement (as measured by test scores and grades) and academic participation (attendance, persistence, and graduation).

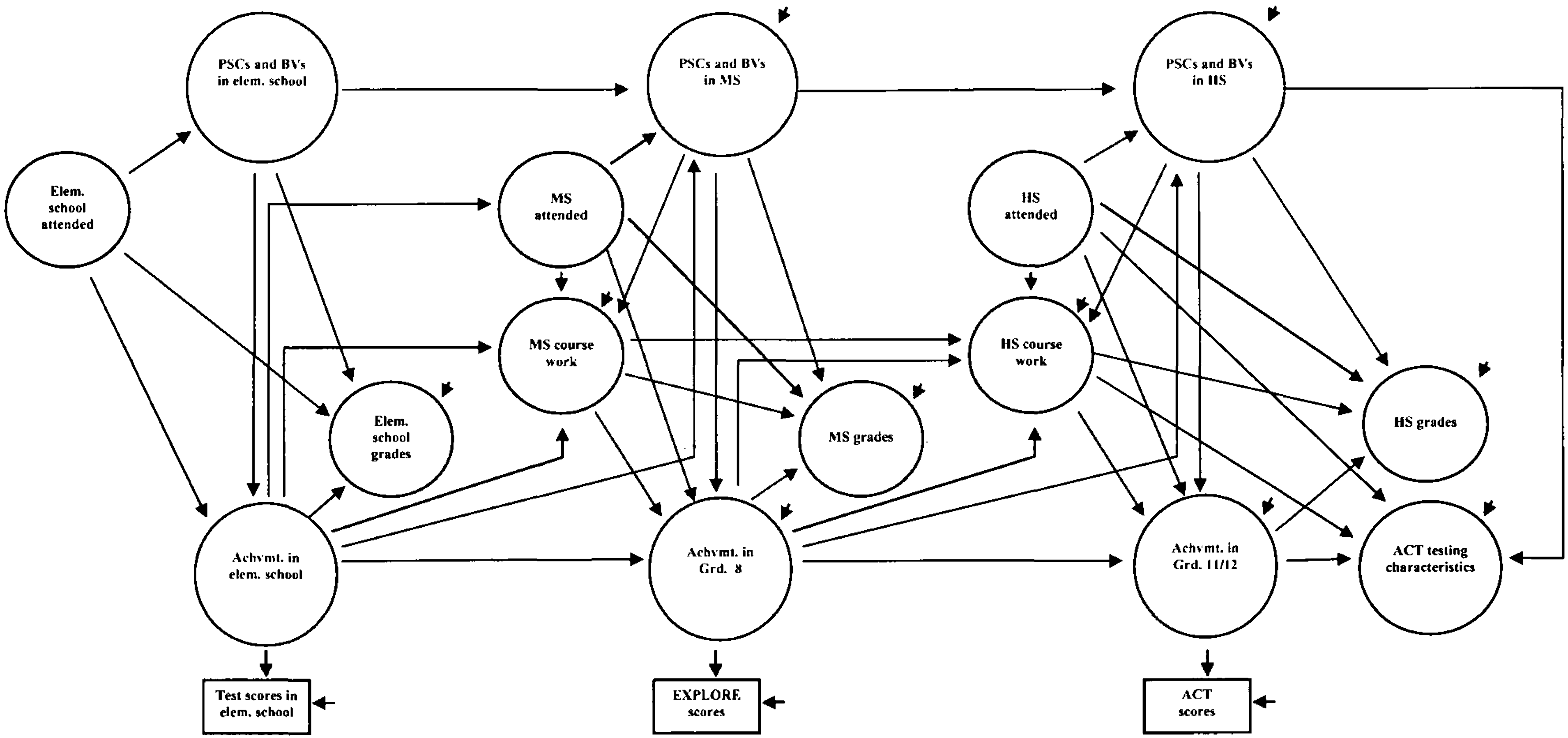
Although we did not have measures of psychosocial characteristics and behavior in this study, one of our results suggests that they could be important predictors of ACT scores: Multiple-tested students who update their course work and course grade information tend to score higher than similar students who test only once (by 0.2 to 1.0 score points). In contrast, students who retest but do not update their course work and course grade information tend to score lower than similar students who test only once (by 0.4 to 0.9 score units). Updating course work and course grade information is likely influenced by characteristics such as conformity, self-discipline, and motivation, which could also influence learning.

Potential Further Research

A principal finding of this study is that the academic skills students achieve by grade eight strongly predict the academic skills they will have when they graduate from high school. This finding suggests the need for diagnosis and intervention to improve academic skills well before grade eight. Research by others suggests that psychosocial characteristics (PSCs) and behavioral variables (BVs) are also important predictors of academic achievement. PSCs include constructs such as motivation, self-discipline, and social connectedness. BVs include characteristics such as attendance, doing homework, and conforming to rules.

Figure 2 on the following page shows relationships among some of the variables in this study (EXPLORE scores, high school attended, high school course work and grades, ACT testing characteristics, and ACT scores) and other variables that are likely important antecedents of academic achievement, but that are not represented in this study. The diagram shows a simplified structure relating PSCs, BVs, and academic achievement in elementary school, middle school, and high school. The circles in the diagram represent constructs (unobserved variables) measured by observed variables (shown as rectangles). For example, academic achievement in grade eleven/twelve is measured by ACT scores. Arrows connecting circles to rectangles or to other circles represent the effects of variables on each other. Small arrows pointing into circles represent random residual errors; small arrows pointing into rectangles represent random measurement errors.

FIGURE 2. A model for academic achievement in grades 11/12, as measured by ACT scores



There are three classes of variables in the diagram:

- *Academic achievement* in elementary school, grade eight, and grades eleven/twelve (red circles at the bottom of the diagram) is represented as a single construct, even though there are actually several distinct dimensions of academic achievement (e.g., reading skills, mathematics skills, etc.). In grades eight and eleven/twelve, academic achievement is measured by EXPLORE scores and ACT scores, respectively, as shown by the indicators in the rectangles beneath the circles.
- *PSCs and BVs* at each time point (red circles at the top of the diagram) are represented as a single construct. In reality, there are several constructs related to psychosocial characteristics and several related to behavior; we have represented them by a single circle to simplify the diagram. Furthermore, the diagram does not show the relevant observed variables measuring the constructs.
- *Other variables.* The diagram also shows the variables school attended, course work, grades, and ACT testing characteristics (black circles). Again, to simplify the presentation, the diagram does not show the observed variables measuring these constructs.

Note the following relationships among the variables:

- Academic achievement in grades eleven/twelve is directly influenced by academic achievement in grade eight, current PSCs and BVs, high school attended, and high school course work. Academic achievement in grades eleven/twelve is also indirectly influenced by variables at earlier points in time.
- Course work is directly influenced by school attended, by current PSCs and BVs, and by past course work and academic achievement. Grades are directly influenced by

school attended, current PSCs and BVs, course work, and achievement. Course work and grades are also indirectly influenced by variables at earlier points in time.

- Two of the predictor variable classes in this study (ACT testing characteristics and high school grades) do not actually drive ACT scores themselves, but instead are themselves consequences of achievement in grades eleven/twelve, high school attended, high school course work, PSCs, and BVs. Thus, ACT testing characteristics and high school grades are proxies for more fundamental variables that are not in our current prediction models.

The model in Figure 2 could be extended to include achievement at Grade 10 (as measured by PLAN scores), as well as high school course work and grades before PLAN and after PLAN. The model could also be extended to include background characteristics, such as those considered in this report.

Models such as this could suggest potential interventions that would increase students' academic achievement more effectively than the enhanced preparation activities considered in this study. For example, interventions that improve behavior in middle school could improve academic achievement in grade eight, which would, in turn, improve achievement in grades eleven/twelve. Improved behavior in middle school would also improve achievement in grades eleven/twelve by improving behavior and course work in high school. Though not shown in the diagram, improved academic achievement in grade eight would also increase the effectiveness of additional course work in high school.

Analogous interventions in elementary school could yield earlier and ultimately greater benefits by preparing students better for middle school. For example, students who have borderline literacy skills in grade four, but who have favorable psychosocial and behavior profiles, might be more likely to benefit from academic remedial interventions, and then go on to take rigorous courses in high school, than would students with less favorable psychosocial and behavior profiles. On the other hand, students with less favorable psychosocial and behavior profiles might benefit more from intervention on those characteristics. If sufficient appropriate data could be collected, we could also investigate these hypotheses through structural models.

References

ACT (2004). *Crisis at the core: Preparing all students for college and work.* Retrieved May 1, 2008 from http://www.act.org/research/policymakers/pdf/crisis_report.pdf

ACT (2005). *ACT high school profile report.* Retrieved June 14, 2007 from http://www.act.org/news/data/05/data.html

ACT (2006). *Ready for college and ready for work: Same or different?* Iowa City, Iowa: Author. Retrieved February 25, 2007 from http://www.act.org/path/policy/pdf/ReadinessBrief.pdf

ACT (2007a). *ACT technical manual.* Iowa City, Iowa: Author. Retrieved August 5, 2007 from http://www.act.org/aap/pdf/ACT_Technical_Manual.pdf

ACT (2007b). *EXPLORE technical manual.* Iowa City, Iowa: Author. Retrieved August 5, 2007 from http://www.act.org/explore/pdf/TechManual.pdf

ACT (2007c). *Rigor at risk: Reaffirming quality in the high school curriculum.* Iowa City, IA: Author. Retrieved June 1, 2007 from http://www.act.org/path/policy/pdf/rigor_report.pdf

ACT (2007d). *2007 ACT national and state scores. College readiness indicators.* Retrieved February 4, 2008 from http://www.act.org/news/data/07/benchmarks.html

ACT (2007e). *2007 ACT national profile report.* Retrieved August 21, 2007 from http://www.act.org/news/data/07/pdf/one.pdf

ACT (2008). *PLAN technical manual.* Iowa City, Iowa: Author. Retrieved March 17, 2008 from http://www.act.org/plan/pdf/PlanTechnicalManual.pdf

Allen, J. & Sconing, J. (2005). *Using ACT Assessment scores to set benchmarks for college readiness.* (ACT Research Report Series 2005-3). Iowa City, IA: ACT, Inc. Retrieved December 12, 2006, from http://www.act.org/research/reports/pdf/ACT_RR2005-3.pdf

Andrews, K. M. & Ziomek, R. L. (1998). *Score gains on retesting with the ACT Assessment.* (ACT Research Report Series 98-7). Iowa City, Iowa: ACT, Inc. Retrieved June 6, 2007 from http://www.act.org/research/reports/pdf/ACT_RR98-07.pdf.

Azen, R. & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological methods, 8*(2), 129-148.

Balfanz, R., Legters, N. & Jordon, W. (2004). Catching up: Effect of the Talent Development ninth-grade instructional interventions in reading and mathematics in high-poverty high schools. *NASSP Bulletin, 88* (December 2004), 3-30.

Barton, M. L., Heidema, C., & Jordan, D. (2002). Teaching reading in mathematics and science. *Educational Leadership*, 60 (3), 24-28.

Bernanke, B. S. (2007, May). *Embracing the challenge of free trade: Competing and prospering in a global economy.* Speech delivered at the Montana Economic Development Summit, 2007. Retrieved January 1, 2008 from http:/www.federalreserve.gov/newsevents/speech/bernanke20070501a.htm.

Borman, G. D. (2000). Title I: The evolving research base. *Journal of Education for Students Placed at Risk, 5* (1 & 2), 27-45.

Bozick, R. & Ingels, J. (2007). *Mathematics coursetaking and achievement at the end of high school: Evidence from the Education Longitudinal Study of 2002 (ELS: 2002).* U.S. Department of Education, Institute of Education Statistics, National Center for Education Statistics. Retrieved February 4, 2008 from http://nces.ed.gov/pubs2008/2008319.pdf

Burkam, D. T. (2003). *English coursetaking and the NELS:88 transcript data.* U.S. Department of Education, Institute of Education Statistics, National Center for Education Statistics. Retrieved October 10, 2006 from http://nces.ed.gov/pubs2003/200302.pdf

Burkam, D. T. & Lee, V. (2003). *Mathematics, foreign language, and science coursetaking and the NELS:88 transcript data.* U.S. Department of Education, Institute of Education Statistics, National Center for Education Statistics. Retrieved October 10, 2006 from http://nces.ed.gov/pubs2003/200301.pdf

Caldus, S. & Bankston III, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, 90, 269-277.

Carbannaro, W. (2005). Tracking, students' effort, and academic achievement. *Sociology of Education, 78* (January), 27-49.

Carnegie Forum on Education and the Economy (1986). *A nation prepared: Teachers for the 21st Century.* New York: Carnegie Corporation of New York.

Center for a Greater Philadelphia (2007). *Value-added assessment and student progress.* Retrieved February 28, 2007 from http://www.cgp.upenn.edu/ope_nation.html.

Chaney, B., Burgdorf, K., Atash, N. (1997). Influencing achievement through high school graduation requirements. *Educational Evaluation and Policy Analysis, 19*(3), 229-244.

Collaborative for Academic, Social, and Emotional Learning (2007). *The benefits of school-based social and emotional learning programs: Highlights from a forthcoming CASEL report.* Retrieved February 28, 2008 from http://www.casel.org/downloads/metaanalysissum.pdf

College Board (2006). *Advanced Placement report to the nation.* Retrieved March 1, 2007 from http://www.collegeboard.com/prod_downloads/about/news_info/ap/2006/ 2006_ap-report-nation.pdf

Committee on Prospering in the Global Economy of the 21st Century (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future.* Washington, DC: National Academies Press.

Council on Competitiveness (2007). *Competitiveness Index: Where America stands.* Washington, DC: Author. Retrieved March 18, 2008 from http://www.compete.org/images/uploads/File/PDF%20Files/Competitiveness_Index_Where_America_Stands_March_2007.pdf

Dadds, M. R. & Fraser, J. A. (2003). Prevention programs. In Essau, C. A. (Ed.), *Conduct and oppositional defiant disorders: Epidemiology, risk factors, and treatment.* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

de León, A. G. (2002). The *urban high school's challenge: Ensuring literacy for every child.* Retrieved June 27, 2008 from http://www.carnegie.org/pdf/literacy.pdf

Duckworth, A. L. & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98* (1), 198-208. Retrieved February 28, 2007 from http://www.sas.upenn.edu/~duckwort/images/GenderDifferencesFeb2006.pdf

Duncan, G., Claessan, A., Huston, A. C., Pagani, L., Engel, M., Sexton, H., Dowsett, C., Magnuson, K. Klebanov, P., Feinstein, L., Brooks-Gunn, Ja., and Duckworth, K. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446.

Dunn, M. G. & Mezzich, A. C. (2007). Development in childhood and adolescence: Implications for prevention research and practice. In Tolan, P., Szapocznik, J. & Sambrano, S. (Eds.), *Preventing youth substance abuse: Science-based programs for children and adolescents.* Washington, DC: American Psychological Association.

Education Commission of the States (2006). *Alignment between high school graduation and college admissions course requirements.* Retrieved May 5, 2008 from http://mb2.ecs.org/reports/Report.aspx?id=745

Girotto, J. R. & Peterson, P. E. (1999). Do hard courses and good grades enhance cognitive skills? In Susan E. Mayer and Paul E. Peterson (Eds.), *Earning and learning: How schools matter.* Washington, DC: Brookings Institution Press.

Grigg, W., Donahue, P., & Dion, G. (2007). *The Nation's report card: 12th-grade reading and mathematics 2005* (NCES 2007-468). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

International Baccalaureate Organization (2007). *IB Diploma Program.* Retrieved March 1, 2007 from http://www.ibo.org/diploma

Jones, K. K. & Byrnes, J. P. (2006). Characteristics of students who benefit from high-quality mathematics instruction. *Contemporary Educational Psychology, 31,* 328-343.

Kreft, Ita G. G. & de Leeuw, Jan. (1998). *Introducing multilevel modeling.* London: Sage (1998).

Laing, J. R., Sawyer, R. L., & Noble, J. P. (1988). Accuracy of self-reported activities and accomplishments of college-bound students. *Journal of College Student Development,* 29, 362-368.

Landsberg, M. (2007, February 23). Grades are rising but learning is lagging, federal reports find. *Los Angeles Times.* Retrieved March 30, 2007 from http://www.latimes.com/news/local/la-me-students23feb23,1,1407397.story

Lanier, C. W. (1994). *ACT Composite scores of retested students.* (ACT Research Report Series 94-3). Iowa City, Iowa: ACT, Inc. Retrieved June 6, 2007 from http://www.act.org/research/reports/pdf/ACT_RR94-03.pdf

Lee, V. E. & Bryk, A. S. (1989). A multilevel model of the social distribution of high-school achievement. *Sociology of Education,* 62 (3), 172-192.

Lemke, M., Sen, A., Pahlke, E., Partelow, L, Miller, D., Williams, T. Kastberg, D. & Jocelyn, L. (2004*). International outcomes of learning in mathematics literacy and problem solving: PISA 2003 results from the U.S. perspective.* (NCES 2005-003). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Leow, C., Marcus, S., Zanutto, E., & Boruch, R. (2004). Effects of advanced course-taking on math and science achievement: Addressing selection bias using propensity scores. *American Journal of Evaluation, 25* (4), 461-478.

Lochner, L. (2004). Education, work, and crime: A human capital approach. *International Economic Review, 45*(3), 811-843.

Ma, Xin & McIntyre, L. J. (2005). Exploring differential effects of mathematics courses on mathematics achievement. *Canadian Journal of Education, 28* (4), 827-852.

Madigan, T. (1997). *Science proficiency and course taking in high school: The relationship of science course-taking patterns to increases in science proficiency between 8th and 12th grades.* U.S. Department of Education, National Center for Education Statistics. Retrieved February 20, 2007 from http://nces.ed.gov/pubs97/97838.pdf

Manzo, K. K. (2008, January 16). Massive funding cuts to 'Reading First' generate worries for struggling schools. *Education Week*, 27 (19), pp. 1, 10.

Manzo, K. K. (2008, June 4). 'Reading First' research offers no definitive answers. *Education Week, 27* (39), p. 9.

Marcus, S. M. (1997). Using omitted variable bias to assess uncertainty in the estimation of an AIDS education treatment effect. *Journal of Educational and Behavioral Statistics, 22*, 193-201.

Meyer, R. H. (1999). The effects of math and math-related courses in high school. In Susan E. Mayer and Paul E. Peterson (Eds.), *Earning and learning: How schools matter.* Washington, DC: Brookings Institution Press.

National Center on Education and the Economy (2007). *Tough choices or tough times: The report of the New Commission on the Skills of the American Workforce.* San Francisco, CA: Jossey-Bass.

National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

Noble, J. & Radunzel, J. (2007). *College readiness = college success beyond the first year.* Paper presented at the Annual Forum of the Association for Institutional Research, June 2-6, Kansas City, MO.

Noble, J., Roberts, W., and Sawyer, R. (2006). *Student achievement, behavior, perceptions, and other factors affecting ACT Scores.* (ACT Research Report No. 2006-1). Iowa City, Iowa: ACT.

Noble, J. P. & Schnelker, D. (2007). *Using hierarchical modeling to examine course work and ACT score relationships across high schools.* (ACT Research Report Series 2007-2). Iowa City, IA: ACT, Inc. Retrieved August 6, 2007 from http://www.act.org/research/reports/pdf/ACT_RR2007-2.pdf

Organisation for Economic Co-operation and Development (2007). *PISA 2006: Science competencies for tomorrow's world.* Retrieved June 23, 2008 from http://www.oecd.org/dataoecd/15/13/39725224.pdf

Parsad, B. & Lewis, L. (2004). *Remedial education at degree-granting institutions in fall 2000* (NCES 2004-010). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Perkins, R., Kleiner, B., Roey, S., & Brown, J. (2004). *The high school transcript study: A decade of change in curricula and achievement, 1990-2000.* U.S. Department of Education, National Center for Education Statistics. Retrieved February 22, 2007 from http://nces.ed.gov/pubsearch/pubs2004/2004455.pdf

Ramirez, F. O., Luo, X., Schofer, E., & Meyer, J. W. (2006). Student achievement and national economic growth. *American Journal of Education, 113*, 1-29.

Raudenbush, S. (2002, April). *New directions in the evaluation of Title I.* Paper presented at the annual meeting of the American Educational Research Association in New Orleans.

Raudenbush, S. (2004). What are value-added models estimating? *Journal of Educational and Behavioral Statistics, 29* (1), 121-129.

Raudenbush, S. & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59* (1), 1-17.

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

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Reynolds, A. J., Temple, J. A., Ou, S. R., Robertson, D. L., Mersky, J. P., Topitzes, J. W., & Niles, M. D. (2007). Effects of a school-based, early childhood intervention on adult health and well-being: A 19-year follow-up of low-income families. *Archives of Pediatrics & Adolescent Medicine, 161* (8), 730-739.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal, 32(3)*, 583-625.

SAS Institute (2007). *PROC MI.* Retrieved February 28, 2007 from http://support.sas.com/onlinedoc/913/docMainpage.jsp

Sawyer, R. L., Laing, J. R., & Houston, W. M. (1989). Accuracy of self-reported high school courses and grades of college-bound students. *College and University, 64(3)*, 288-299.

Schiel, J., Pommerich, M., & Noble, J. P. (1996). *Factors associated with longitudinal educational achievement, as measured by PLAN and ACT Assessment Scores* (ACT Research Report Series 96-5). Iowa City, IA: ACT, Inc.

Schweinhart, L. J., Barnes, H. V., & Weikart, D. P., with Barnett, W. S. & Epstein, A. S. (1993) *Significant benefits: The High/Scope Perry Preschool Study through age 27.* Ypsilanti, Michigan: High/Scope Press.

Shettle, C., Roey, S., Mordica, J., Perkins, R., Nord, C., Teodorovic, J., Brown, J., Lyons, M., Averett, C., Kastberg, D. (2007). *The Nation's report card: America's high school graduates* (NCES 2007-467). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. Retrieved February 28, 2007 from http://nces.ed.gov/nationsreportcard/pdf/studies/2007467.pdf

Slavin, R., Cheung, A., Groff, C., and Lake, C. (2007). *Effective reading programs for middle and high school: A best-evidence synthesis.* Retrieved February 22, 2008 from http://www.bestevidence.org/_images/word_docs/EffectiveProgramsforMSandHSReading%20%2002%2017%2008%20BEE%20version.doc

Slavin, R. & Lake, C. (2007). *Effective programs in elementary school mathematics: A best-evidence synthesis.* Retrieved February 27, 2008 from http://www.bestevidence.org/_images/word_docs/Eff%20progs%20ES%20math%20Version%201.2%20for%20BEE%2002%2009%2007.doc

Slavin, R., Lake, C., & Groff, C. (2007). *Effective programs in middle and high school mathematics: A best-evidence synthesis.* Retrieved February 22, 2008 from http://www.bestevidence.org/_images/word_docs/EffProgsMSandHSMath%2007%2030%2007.doc

Steen, L. A. (1999). Numeracy: The new literacy for a data-drenched society. *Educational Leadership, 57*(2), 8-13.

Stiggins, R. J., Frisbie, D. A., and Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice, 8* (2), 5-14.

The World Bank (2006). *Where is the wealth of nations? Measuring capital for the 21st century.* Retrieved July 24, 2007 from http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/ENVIRONMENT/EXTEEI/0,,contentMDK:20744819~pagePK:210058~piPK:210062~theSitePK:408050,00.html

Tienken, C. H. (2008). Rankings of international achievement test performance and economic strength: Correlation or conjecture? *International Journal of Education Policy & Leadership, 3*(4). Retrieved April 26, 2008 from http://www.ijepl.org

TIMSS and PIRLS International Study Center (2004). *TIMSS 2003 highlights.* Retrieved July 1, 2008 from http://timss.bc.edu/timss2003i/conference_IR.html

U.S. Department of Education (2002). *No Child Left Behind: A desktop reference.* Retrieved March 14, 2007 from http://www.ed.gov/print/admins/lead/account/nclbreference/reference.pdf

U.S. Department of Education, National Center for Education Statistics (2006). *The condition of education 2006* (NCES 2006-071). Washington, D.C. U.S. Government Printing Office. Retrieved February 27, 2007 from http://nces.ed.gov/nationsreportcard/pdf/main2005/2007468_1.pdf

U.S. Department of Health and Human Services (2005). *Head Start impact study: First year findings*. Retrieved August 5, 2007 from http://www.acf.hhs.gov/programs/opre/hs/impact_study/reports/first_yr_execsum/first_yr_execsum.pdf

U.S. Department of Health and Human Services (2007, August 15). *Head Start program fact sheet.* Retrieved August 22, 2007 from http://www.acf.hhs.gov/programs/hsb/about/fy2007.html

U.S. Department of Labor, Bureau of Labor Statistics (2006). Job outlook by education, 2004-14. *Occupational Outlook Quarterly (Fall 2006).* Retrieved March 5, 2008 from http://www.bls.gov/opub/ooq/2006/fall/art02.pdf

U.S. Department of Labor, Bureau of Labor Statistics (2007). *Education pays.* Retrieved March 18, 2008 from http://www.bls.gov/emp/emptab7.htm

U.S. Government Accountability Office. (n.d.). *Budget of the United States government Department of Education.* Retrieved March 15, 2007 from http://www.gpoaccess.gov/usbudget/fy08/pdf/budget/education.pdf

U.S. Government Accountability Office (2008). *Head Start. A more comprehensive risk management strategy and data improvements could further strengthen program oversight* (GAO-08-221). Retrieved February 28, 2008 from http://www.gao.gov/cgi-bin/getrpt?GAO-08-221

van der Klaauw (2005). *Breaking the link between poverty and low student achievement: An evaluation of Title I.* Retrieved March 14, 2007 from http://www.unc.edu/~vanderkl/brlink.pdf

Webster, W. J. & Mendro, R. L. (1997). The Dallas value-added accountability system. In Jason Millman (Ed.), *Grading teachers, grading schools*. Thousand Oaks, CA: Corwin Press, Inc.

Western Interstate Commission for Higher Education (2008). *Knocking at the college door March 2008.* Boulder, CO: Author.

What Works Clearinghouse (2007a). *Welcome to WWC.* Retrieved February 27, 2008 from http://ies.ed.gov/ncee/wwc.

What Works Clearinghouse (2007b). *Character education.* Retrieved February 22, 2008 from http://ies.ed.gov/ncee/wwc/reports/character_education/topic/

Worrell, F. C. & Hale, R. L. (2001). The relationship of hope in the future and perceived school climate to school completion. *School Psychology Quarterly, 16*, 370-388.

Appendix

Statistical Tables

Table A-1
Analysis Variables

Analysis variable	Definition	Range	Source data set: Mean (SD)	Imputation data set: Mean (SD)	Imputation data set: Pct. imputed	Imputation data set: Difference: Avg. imp. val. – Avg. nonimp. val.	Imputation data set: Std. difference: Avg. imp. val. – Avg. nonimp. val.	Analysis data set: Mean (SD)
		N (students)	132,441	123,748				117,280
		N (high schools)	4,992	4,903				4,638
Outcome variables								
A_ENG_SC	ACT English score	1 - 36	20.6 (5.6)	20.6 (5.7)	0	. . .	. . .	20.7 (5.7)
A_MTH_SC	ACT Mathematics score	1 - 36	20.3 (4.8)	20.4 (4.8)	0	. . .	. . .	20.5 (4.8)
A_RDG_SC	ACT Reading score	1 - 36	21.1 (5.8)	21.1 (5.8)	0	. . .	. . .	21.2 (5.8)
A_SCI_SC	ACT Science score	1 - 36	20.8 (4.4)	20.8 (4.4)	0	. . .	. . .	20.9 (4.4)
Background variables								
GENDER	1=Male; 0=Female	0, 1	0.44 (0.50)	0.45 (0.50)	0	. . .	. . .	0.45 (0.50)
ETH	1=Caucasian-American/White or Asian American, Pacific Islander; 0=All other categories	0, 1	0.78 (0.41)	0.78 (0.41)	0	. . .	. . .	0.79 (0.41)
PRNT_HS	Number of parents with high school diploma	0 - 2	1.84 (0.46)	1.84 (0.46)	12	0.01	0.02	1.84 (0.46)
PRNT_CO	Number of parents who have attended college	0 - 2	0.86 (0.85)	0.86 (0.84)	24	0.05	0.06	0.87 (0.84)
FAMINC	Family income, in thousands of dollars	12 - 120	57.3 (33.5)	58.1 (33.1)	21	4.2	0.13	58.4 (33.1)

(continued on next page)

Table A-1
Analysis Variables

			Source data set	Imputation data set				Analysis data set
		N (students)	132,441	123,748				117,280
		N (high schools)	4,992	4,903				4,638
Analysis variable	**Definition**	**Range**	**Mean (SD)**	**Mean (SD)**	**Pct. imputed**	**Difference: Avg. imp. val. – Avg. nonimp. val.**	**Std. difference: Avg. imp. val. – Avg. nonimp. val.**	**Mean (SD)**
Background variables (continued)								
ENGLHOME	1=English primary language spoken at home; 0=No	0, 1	0.97 (0.18)	0.97 (0.18)	1	0.00	0.00	0.97 (0.18)
EXPLORE scores								
E_ENG_SC	EXPLORE English score	1 – 25	16.3 (3.9)	16.2 (3.9)	0	. . .	. . .	16.3 (3.9)
E_MTH_SC	EXPLORE Mathematics score	1 - 25	16.4 (3.4)	16.3 (3.4)	0	. . .	. . .	16.4 (3.3)
E_RDG_SC	EXPLORE Reading score	1 - 25	15.9 (3.8)	15.9 (3.8)	0	. . .	. . .	16.0 (3.8)
E_SCI_SC	EXPLORE Science score	1 - 25	17.5 (2.8)	17.5 (2.7)	0	. . .	. . .	17.5 (2.7)

(continued on next page)

Table A-1
Analysis Variables

			Source data set	Imputation data set				Analysis data set
	N (students)		132,441	123,748				117,280
	N (high schools)		4,992	4,903				4,638
Analysis variable	**Definition**	**Range**	**Mean (SD)**	**Mean (SD)**	**Pct. imputed**	**Difference: Avg. imp. val. – Avg. nonimp. val.**	**Std. difference: Avg. imp. val. – Avg. nonimp. val.**	**Mean (SD)**
ACT testing characteristics								
AGE	Age at time of ACT testing	1 - 21	17.38 (0.67)	17.45 (0.57)	0	-0.2	-0.35	17.45 (0.57)
EDLVL	Educational level at time of testing	11, 12	11.37 (0.63)	11.44 (0.50)	0	. . .	. . .	11.44 (0.50)
RETEST1	1=Took the ACT more than once; did not update course work and course grade data. 0=No	0, 1	0.34 (0.47)	0.32 (0.47)	0	. . .	. . .	0.32 (0.47)
RETEST2	1=Took the ACT more than once; updated course work and course grade data. 0=No	0, 1	0.24 (0.43)	0.25 (0.43)	0	. . .	. . .	0.26 (0.44)
COLDV *(Level-2 variable)*	1=High school located in Colorado. 0=No	0, 1	0.04 (0.20)	0.04 (0.21)	0	. . .	. . .	0.04 (0.20)
ILLDV *(Level-2 variable)*	1=High school located in Illinois. 0=No	0, 1	0.18 (0.38)	0.19 (0.39)	0	. . .	. . .	0.20 (0.40)

(continued on next page)

Table A-1
Analysis Variables

			Source data set	Imputation data set				Analysis data set
		N (students)	132,441	123,748				117,280
		N (high schools)	4,992	4,903				4,638
Analysis variable	**Definition**	**Range**	**Mean (SD)**	**Mean (SD)**	**Pct. imputed**	**Difference: Avg. imp. val. – Avg. nonimp. val.**	**Std. difference: Avg. imp. val. – Avg. nonimp. val.**	**Mean (SD)**
Standard course work								
ENG_RK	Highest-level English course taken:	0 - 4	3.27 (0.62)	3.32 (0.57)	6	-0.06	-0.11	3.32 (0.57)
	English in grade 9							
	English in grade 10							
	English in grade 11							
	English in grade 12							
MTH_RK	Highest-level mathematics course taken:	0 - 5	3.22 (0.85)	3.24 (0.85)	6	-0.11	-0.13	3.25 (0.85)
	Algebra I							
	Geometry							
	Algebra II							
	Trigonometry							
	Calculus							
SOC_SU	Number of social studies curses taken:	0 - 7	3.84 (1.22)	3.87 (1.21)	6	-0.06	-0.05	3.86 (1.21)
	U.S. history							
	World history							
	Other history							
	American government							
	Economics							
	Geography							
	Psychology							

(continued on next page)

Table A-1
Analysis Variables

Analysis variable	Definition	Range	Source data set Mean (SD)	Imputation data set Mean (SD)	Imputation data set Pct. imputed	Imputation data set Difference: Avg. imp. val. – Avg. nonimp. val.	Imputation data set Std. difference: Avg. imp. val. – Avg. nonimp. val.	Analysis data set Mean (SD)
	N (students)		132,441	123,748				117,280
	N (high schools)		4,992	4,903				4,638
Standard course work (continued)								
NSC_RK	Highest-level science course taken: Biology Chemistry Physics	0 - 3	1.97 (0.77)	2.00 (0.76)	6	-0.08	-0.11	2.01 (0.76)
FL	1=Studied any foreign language. 0=No	0, 1	0.91 (0.28)	0.91 (0.29)	22	-0.04	-0.14	0.91 (0.29)
Adv./Hon. course work								
ADV_ENG	1=Took Adv./Hon. course(s) in English. 0=No	0, 1	0.44 (0.50)	0.43 (0.50)	15	0.01	0.02	0.44 (0.50)
ADV_MTH	1=Took Adv./Hon. course(s) in mathematics. 0=No	0, 1	0.40 (0.49)	0.39 (0.49)	15	0.02	0.04	0.40 (0.49)
ADV_SOC	1=Took Adv./Hon. course(s) in social studies. 0=No	0, 1	0.34 (0.47)	0.34 (0.47)	15	0.04	0.09	0.35 (0.48)
ADV_NSC	1=Took Adv./Hon. course(s) in science. 0=No	0, 1	0.36 (0.48)	0.36 (0.48)	15	0.04	0.08	0.36 (0.48)
ADV_FL	1=Took Adv./Hon. course(s) in foreign language. 0=No	0, 1	0.18 (0.39)	0.19 (0.39)	16	0.04	0.10	0.19 (0.39)

(continued on next page)

Table A-1
Analysis Variables

Analysis variable	Definition	Range	Source data set: Mean (SD)	Imputation data set: Mean (SD)	Imputation data set: Pct. imputed	Imputation data set: Difference: Avg. imp. val. – Avg. nonimp. val.	Imputation data set: Std. difference: Avg. imp. val. – Avg. nonimp. val.	Analysis data set: Mean (SD)
		N (students)	132,441	123,748				117,280
		N (high schools)	4,992	4,903				4,638
Subject-area grade averages								
ENG_AV	Average of grades in English courses	0.0 - 4.0	3.21 (0.72)	3.17 (0.72)	15	-0.16	-0.22	3.18 (0.71)
MTH_AV	Average of grades in mathematics courses	0.0 - 4.0	3.06 (0.80)	3.01 (0.80)	16	-0.18	-0.23	3.02 (0.80)
SOC_AV	Average of grades in social studies courses	0.0 - 4.0	3.34 (0.69)	3.30 (0.69)	16	-0.17	-0.25	3.31 (0.69)
NSC_AV	Average of grades in natural science courses	0.0 - 4.0	3.11 (0.81)	3.07 (0.81)	19	-0.18	-0.22	3.08 (0.80)

Table A-2
Models for Predicting ACT English Score

Predictor variables	Main-effects model Coefficient (SE)	Interaction model Coefficient (SE)
Fixed effects		
Intercept		
(Intercept)	20.474591 (0.020059)	20.479845 (0.020134)
Mean FAMINC	0.008055 (0.001778)	0.008164 (0.001770)
Mean PRT_CO	0.404172 (0.068955)	0.400891 (0.069031)
Mean RETEST1	0.615511 (0.116512)	0.623297 (0.116579)
Mean RETEST2	0.397313 (0.113134)	0.418565 (0.113502)
Mean E_ENG	0.091163 (0.021587)	0.108997 (0.018464)
Mean E_MTH	-0.085516 (0.021554)	-0.087699 (0.019877)
Mean E_RDG	0.071881 (0.023508)	*XXX*
Mean E_SCI	-0.078268 (0.030416)	*XXX*
Mean ENG_AV	-0.568476 (0.074460)	-0.530145 (0.072466)
COLDV	0.688053 (0.107341)	0.634711 (0.108509)
LLDV	0.827590 (0.062954)	0.772186 (0.062982)
Background variables		
GENDER	-0.171447 (0.021147)	-0.167464 (0.021094)
ETH	0.443248 (0.030352)	0.457522 (0.030399)
FAMINC	0.003254 (0.000370)	0.003131 (0.000370)
PRNT_HS	0.140605 (0.023244)	0.145884 (0.022956)
PRNT_CO	0.169786 (0.014404)	0.161005 (0.014305)
ENGLHOME	*NNN*	*NNN*
EXPLORE scores		
E_ENG_SC	0.530265 (0.004301)	0.528877 (0.004281)
E_MTH_SC	0.206368 (0.004538)	0.206021 (0.004491)
E_RDG_SC	0.239363 (0.004069)	0.231792 (0.004093)
E_SCI_SC	0.188779 (0.005493)	0.185863 (0.005398)
ACT testing characteristics		
AGE	-0.792660 (0.022193)	-0.791340 (0.022247)
EDLVL	0.909590 (0.030530)	0.950505 (0.030821)
RETEST1	-0.598453 (0.026147)	-0.575265 (0.027625)
RETEST2	0.991143 (0.032289)	0.908478 (0.037831)
Standard course work		
ENG_RK	*NNN*	*NNN*
SOC_SU	*NNN*	*NNN*
FL	0.298497 (0.044006)	0.344002 (0.043569)

(continued on next page)

Table A-2
Models for Predicting ACT English Score

Predictor variables	Main-effects model Coefficient (SE)	Interaction model Coefficient (SE)
Adv./Hon. coursework		
ADV_ENG	0.687218 (0.033022)	0.661248 (0.033529)
ADV_SOC	0.429717 (0.029209)	0.402439 (0.030906)
ADV_FL	0.215847 (0.031105)	0.198814 (0.031529)
Subject area grade averages		
ENG_AV	0.644841 (0.023933)	0.714487 (0.024020)
SOC_AV	0.285904 (0.023204)	0.315542 (0.022765)
Interactions		
EDLVL *x* AGE	*n/a*	-0.237316 (0.044465)
EDLVL *x* RETEST1	*n/a*	0.190456 (0.054752)
EDLVL *x* RETEST2	*n/a*	0.253205 (0.064663)
E_ENG_SC *x* ADV_ENG	*n/a*	0.041036 (0.007991)
E_ENG_SC *x* ADV_SOC	*n/a*	0.030411 (0.007604)
E_RDG_SC *x* ENG_AV	*n/a*	0.047838 (0.004405)
ADV_ENG *x* ENG_AV	*n/a*	0.222565 (0.038161)

Random effects	Standard deviation	
Intercept	0.71373	0.71915
EDLVL	0.37506	0.45300
RETST2	0.52846	*XXX*
ADV_ENG	0.36426	0.35562
ADV_FL	0.42598	0.39925
ENG_AV	0.26569	0.26089
E_ENG_SC *x* ADV_SOC	*n/a*	*XXX*
E_RDG_SC *x* ENG_AV	*n/a*	0.05914
ADV_ENG *x* ENG_AV	*n/a*	0.35261
Residual	3.031	3.025

Notes:

1. Estimates are based on the 2006 data. Unless otherwise noted, all estimated effects are statistically significant ($p<.001$ for student-level effects; $p<.01$ for school-level effects).
2. *NNN* denotes effects that were not statistically significant at the prescribed level in either the 2005 data or the 2006 data.
3. *XXX* denotes effects that were statistically significant at the prescribed level in the 2005 data, but not in the 2006 data.
4. *n/a* = not applicable to the main-effects model.

Table A-3
Models for Predicting ACT Mathematics Score

Predictor variables	Main-effects model Coefficient (SE)	Interaction model Coefficient (SE)
Fixed effects		
Intercept		
(Intercept)	20.197673 (0.020198)	20.225312 (0.019406)
Mean ETH	*XXX*	*NNN*
Mean FAMINC	0.009733 (0.001470)	0.010415 (0.001448)
Mean PRT_CO	0.300416 (0.061156)	0.372183 (0.057354)
Mean RETEST2	-0.500335 (0.089857)	-0.401459 (0.085654)
Mean MTH_RK	-0.226172 (0.053043)	-0.192893 (0.052008)
Mean MTH_AV	-0.338305 (0.057427)	-0.256343 (0.055064)
COLDV	*XXX*	*NNN*
ILLDV	*XXX*	*NNN*
Background variables		
GENDER	1.134703 (0.017168)	1.111755 (0.016986)
ETH	0.364623 (0.024386)	0.416699 (0.023199)
FAMINC	0.002145 (0.000294)	0.002484 (0.000290)
PRNT_HS	*NNN*	*NNN*
PRNT_CO	0.084684 (0.011544)	*NNN*
ENGLHOME	-0.328394 (0.050428)	-0.266944 (0.048927)
EXPLORE scores		
E_ENG_SC	0.118343 (0.003149)	0.120887 (0.0030990)
E_MTH_SC	0.446661 (0.005124)	0.439068 (0.0047290)
E_RDG_SC	0.052557 (0.003298)	0.047239 (0.0032330)
E_SCI_SC	0.171716 (0.004786)	0.159926 (0.0047020)
ACT testing characteristics		
AGE	-0.363946 (0.017052)	-0.402342 (0.016561)
EDLVL	*NNN*	*NNN*
RETEST1	-0.412993 (0.021927)	-0.345732 (0.021158)
RETEST2	0.535424 (0.025620)	0.586726 (0.026793)
Standard course work		
MTH_RK	0.648737 (0.015701)	0.701175 (0.015819)
NSC_RK	0.305078 (0.014761)	0.357814 (0.014990)
Adv./Hon. coursework		
ADV_MTH	0.892148 (0.031736)	0.752676 (0.031519)
ADV_NSC	0.340758 (0.024424)	0.270257 (0.023060)

(continued on next page)

Table A-3
Models for Predicting ACT Mathematics Score

Predictor variables	Main-effects model Coefficient (SE)	Interaction model Coefficient (SE)
Subject area grade averages		
MTH_AV	0.830691 (0.016337)	0.968122 (0.018461)
NSC_AV	0.186945 (0.015691)	0.271612 (0.017929)
Interactions		
AGE *x* RETEST2	*n/a*	-0.312359 (0.034733)
E_MTH_SC *x* MTH_RK	*n/a*	0.065026 (0.003460)
E_MTH_SC *x* NSC_RK	*n/a*	*XXX*
E_MTH_SC *x* ADV_NSC	*n/a*	0.027443 (0.006429)
E_MTH_SC *x* MTH_AV	*n/a*	0.102898 (0.004329)
E_MTH_SC *x* NSC_AV	*n/a*	0.032558 (0.004676)
MTH_RK *x* MTH_AV	*n/a*	0.124048 (0.030170)
NSC_RK *x* NSC_AV	*n/a*	0.212612 (0.015566)
ADV_MTH *x* MTH_AV	*n/a*	0.341137 (0.026630)

Random effects	Standard deviation	
Intercept	0.84363	0.81533
RETEST2	0.39884	0.39199
MTH_RK	0.35038	0.34220
NSC_RK	0.28712	0.28232
ADV_MTH	0.65905	0.62157
MTH_AV	0.25477	0.23297
NSC_AV	0.22811	*NNN*
E_MTH_SC *x* MTH_RK	*n/a*	0.07339
Residual	2.441	2.387

Notes:

1. Estimates are based on the 2006 data. Unless otherwise noted, all estimated effects are statistically significant (p<.001 for student-level effects; p<.01 for school-level effects).
2. *NNN* denotes effects that were not statistically significant at the prescribed level in either the 2005 data or the 2006 data.
3. *XXX* denotes effects that were statistically significant at the prescribed level in the 2005 data, but not in the 2006 data.
4. *n/a* = not applicable to the main-effects model.

Table A-4
Models for Predicting ACT Reading Score

Predictor variables	Main-effects model Coefficient (SE)	Interaction model Coefficient (SE)
Fixed effects		
Intercept		
(Intercept)	21.029727 (0.020478)	21.028369 (0.020472)
Mean PRT_COL	0.661414 (0.060028)	0.677274 (0.060064)
Mean EDLVL	-0.368181 (0.105956)	-0.398717 (0.105964)
Mean ADV_E	-0.356605 (0.102913)	-0.375100 (0.102194)
COLDV	0.479199 (0.109614)	0.483317 (0.110301)
ILLDV	0.294659 (0.066606)	0.301793 (0.066824)
Background variables		
GENDER	0.070556 (0.025111)	0.066222 (0.025074)
ETH	0.458615 (0.033305)	0.473414 (0.033293)
FAMINC	*NNN*	*NNN*
PRNT_HS	*NNN*	*NNN*
PRNT_CO	0.151858 (0.017448)	0.141982 (0.017374)
ENGLHOME	*NNN*	*NNN*
EXPLORE scores		
E_ENG_SC	0.363528 (0.004765)	0.364529 (0.004757)
E_MTH_SC	0.112439 (0.005064)	0.111378 (0.005052)
E_RDG_SC	0.436102 (0.004897)	0.425792 (0.004964)
E_SCI_SC	0.271758 (0.006666)	0.271125 (0.006655)
ACT testing characteristics		
AGE	-0.598004 (0.025629)	-0.612284 (0.025513)
EDLVL	0.848080 (0.036966)	0.846695 (0.037144)
RETEST1	-0.942199 (0.032618)	-0.915270 (0.032593)
RETEST2	0.214469 (0.034360)	0.303961 (0.042431)
Standard course work		
ENG_RK	*NNN*	*NNN*
SOC_SU	*NNN*	*NNN*
FL	*NNN*	*NNN*
Adv./Hon. coursework		
ADV_ENG	0.560311 (0.033327)	0.539130 (0.033117)
ADV_SOC	0.581588 (0.037263)	0.548906 (0.037238)
ADV_FL	*XXX*	*XXX*

(continued on next page)

Table A-4
Models for Predicting ACT Reading Score

Predictor variables	Main-effects model Coefficient (SE)	Interaction model Coefficient (SE)
Subject area grade averages		
ENG_AV	0.503876 (0.024726)	0.526074 (0.025633)
SOC_AV	0.407544 (0.024452)	0.496900 (0.025568)
Interactions		
EDLVL *x* RETEST2	*n/a*	-0.315713 (0.073098)
E_RDG_SC *x* ADV_ENG	*n/a*	*XXX*
E_RDG_SC *x* ADV_SOC	*n/a*	0.037568 (0.007074)
E_RDG_SC *x* SOC_AV	*n/a*	0.051639 (0.005597)
ADV_ENG *x* ENG_AV	*n/a*	0.280377 (0.044174)
Random effects	**Standard deviation**	
Intercept	0.67822	0.67959
GENDER	0.33631	0.33171
EDLVL	0.41947	0.40750
ADV_ENG	0.44382	0.40606
Residual	3.647	3.643

Notes:

1. Estimates are based on the 2006 data. Unless otherwise noted, all estimated effects are statistically significant ($p<.001$ for student-level effects; $p<.01$ for school-level effects).
2. *NNN* denotes effects that were not statistically significant at the prescribed level in either the 2005 data or the 2006 data.
3. *XXX* denotes effects that were statistically significant at the prescribed level in the 2005 data, but not in the 2006 data.
4. *n/a* = not applicable to the main-effects model.

Table A-5
Models for Predicting ACT Science Score

Predictor variables	Main-effects model Coefficient (SE)	Interaction model Coefficient (SE)
Fixed effects		
Intercept		
(Intercept)	20.713058 (0.016380)	20.715791 (0.016227)
Mean EDLVL	-0.409360 (0.071349)	-0.456388 (0.068997)
Mean E_MTH	-0.046089 (0.016551)	0.047627 (0.016397)
Mean E_SCI	0.096914 (0.021714)	0.100901 (0.021544)
Mean ADV_M	-0.386279 (0.087738)	-0.365424 (0.087226)
Mean NSC_AV	-0.172419 (0.053137)	-0.177519 (0.052734)
COLDV	0.319942 (0.091464)	*NNN*
Background variables		
GENDER	1.174278 (0.018998)	1.159742 (0.019016)
ETH	0.490671 (0.026012)	0.516784 (0.025710)
PRNT_HS	0.113633 (0.019817)	0.122714 (0.019738)
PRNT_CO	*NNN*	*NNN*
ENGLHOME	*NNN*	*NNN*
EXPLORE scores		
E_ENG_SC	0.158196 (0.003635)	0.159601 (0.003591)
E_MTH_SC	0.259685 (0.004489)	0.254344 (0.004465)
E_RDG_SC	0.152782 (0.003763)	0.149241 (0.003768)
E_SCI_SC	0.287788 (0.005489)	0.282234 (0.005505)
ACT testing characteristics		
AGE	-0.544168 (0.020602)	-0.559479 (0.020391)
EDLVL	0.353867 (0.027636)	0.362776 (0.027586)
RETEST1	-0.504144 (0.024233)	-0.438846 (0.023988)
RETEST2	0.292207 (0.026253)	0.366078 (0.030655)
Standard course work		
MTH_RK	0.319566 (0.014470)	0.325493 (0.014281)
NSC_RK	0.252017 (0.014852)	0.267371 (0.014855)
Adv./Hon. coursework		
ADV_MTH	0.296616 (0.022896)	0.245156 (0.023007)
ADV_NSC	0.405311 (0.025650)	0.356372 (0.026480)

(continued on next page)

Table A-5
Models for Predicting ACT Science Score

Predictor variables	Main-effects model Coefficient (SE)	Interaction model Coefficient (SE)
Subject area grade averages		
MTH_AV	0.373696 (0.016148)	0.385009 (0.016652)
NSC_AV	0.295045 (0.016776)	0.368531 (0.017843)
Interactions		
EDLVL *x* RETEST2	*n/a*	-0.204920 (0.053634)
AGE *x* RETEST1	*n/a*	0.213013 (0.035831)
E_MTH_SC *x* MTH_RK	*n/a*	0.026168 (0.003944)
E_MTH_SC *x* ADV_NSC	*n/a*	0.047218 (0.007650)
E_MTH_SC *x* NSC_AV	*n/a*	0.038197 (0.004797)
E_SCI_SC *x* NSC_AV	*n/a*	0.023633 (0.005624)
ADV_MTH *x* MTH_AV	*n/a*	0.093263 (0.026339)
Random effects	**Standard deviation**	
Intercept	0.56795	0.57185
RETEST2	0.31938	*XXX*
NSC_RK	0.22329	0.23216
NSC_AVG	0.20298	*XXX*
E_MTH_SC *x* MTH_RK	*n/a*	0.05495
E_MTH_SC *x* ADV_NSC	*n/a*	0.10605
Residual	2.799	2.787

Notes:

1. Estimates are based on the 2006 data. Unless otherwise noted, all estimated effects are statistically significant (p<.001 for student-level effects; p<.01 for school-level effects).
2. *NNN* denotes effects that were not statistically significant at the prescribed level in either the 2005 data or the 2006 data.
3. *XXX* denotes effects that were statistically significant at the prescribed level in the 2005 data, but not in the 2006 data.
4. *n/a* = not applicable to the main-effects model.

Table A-6
Magnitudes of Standardized Regression Weights in Main-Effects Models

	ACT score			
Predictor variables	**English**	**Mathematics**	**Reading**	**Science**
Background variables				
GENDER	0.02	0.12	0.01	0.13
ETH	0.03	0.03	0.03	0.04
FAMINC	0.02	0.01	0.00	0.00
PRNT_HS	0.01	0.00	0.00	0.01
PRNT_CO	0.03	0.01	0.02	0.00
ENGLHOME	0.00	0.01	0.00	0.00
Sum	**0.10**	**0.19**	**0.06**	**0.19**
EXPLORE scores				
E_ENG_SC	0.37	0.09	0.24	0.14
E_MTH_SC	0.12	0.30	0.06	0.20
E_RDG_SC	0.16	0.04	0.28	0.13
E_SCI_SC	0.09	0.10	0.13	0.18
Sum	**0.74**	**0.54**	**0.72**	**0.64**
ACT testing characteristics				
AGE	0.08	0.04	0.06	0.07
EDLVL	0.08	0.00	0.07	0.04
RETEST1	0.05	0.04	0.08	0.05
RETEST2	0.08	0.05	0.02	0.03
Sum	**0.29**	**0.13**	**0.22**	**0.19**
Standard course work				
ENG_RK	0.00	*n/a*	0.00	*n/a*
SOC_SU	0.00	*n/a*	0.00	*n/a*
FL	0.02	*n/a*	0.00	*n/a*
MTH_RK	*n/a*	0.11	*n/a*	0.06
NSC_RK	*n/a*	0.05	*n/a*	0.04
Sum	**0.02**	**0.16**	**0.00**	**0.10**
Adv./Hon. coursework				
ADV_ENG	0.06	*n/a*	0.05	*n/a*
ADV_SOC	0.04	*n/a*	0.05	*n/a*
ADV_FL	0.01	*n/a*	0.00	*n/a*
ADV_MTH	*n/a*	0.09	*n/a*	0.03
ADV_NSC	*n/a*	0.03	*n/a*	0.04
Sum	**0.11**	**0.12**	**0.09**	**0.08**

(continued on next page)

Table A-6
Magnitudes of Standardized Regression Weights in Main-Effects Models

	ACT score			
Predictor variables	**English**	**Mathematics**	**Reading**	**Science**
Subject area grade averages				
ENG_AV	0.08	*n/a*	0.06	*n/a*
SOC_AV	0.03	*n/a*	0.06	*n/a*
MTH_AV	*n/a*	0.13	*n/a*	0.07
NSC_AV	*n/a*	0.03	*n/a*	0.05
Sum	**0.12**	**0.16**	**0.11**	**0.12**
Sum, all predictor variables	**1.37**	**1.29**	**1.20**	**1.31**