Available online at www.sciencedirect.com
ScienceDirect
Contemporary Educational Psychology 32 (2007) 599–629
Contemporary Educational Psychology
www.elsevier.com/locate/cedpsych

The relative importance of predictors of math and science achievement: An opportunity–propensity analysis

James P. Byrnes [a,*], David C. Miller [b]

[a] Temple University, USA
[b] American Institutes of Research, USA

Available online 2 November 2006

Abstract

In the present study, the authors propose a new framework that integrates literature on achievement, supports the testing of novel hypotheses, and stresses the importance of examining a large number of factors in the same study. This framework assumes that high achievement is a function of three categories of factors: (a) opportunity factors (e.g., coursework), (b) propensity factors (e.g., prerequisite skills, motivation), and (c) distal factors (e.g., SES). A secondary analysis of the National Longitudinal Educational Study (NELS:88) using hierarchical regression and structural equation modeling revealed that 58–81% of the variance in achievement was explained by family variables and specific opportunity and propensity factors. The findings are discussed in terms of their potential implications for intervention efforts.

Keywords: Achievement; Educational opportunity; Propensity; Ethnic and racial differences

1. Introduction

As any comprehensive handbook of educational research illustrates (e.g., Alexander & Winne, 2006), the field of educational psychology is subdivided into distinct research areas such as motivation, instruction, reading achievement, math achievement, and so on. Scholars who specialize in one of these areas tend not to specialize in others. In addition,

* Corresponding author.
E-mail address: jpbyrnes@temple.edu (J.P. Byrnes).

0361-476X/$ - see front matter
doi:10.1016/j.cedpsych.2006.09.002

researchers within each of these areas often focus on specific components of some predictor of achievement (e.g., motivational goals) to the exclusion of other components of that same predictor (e.g., motivational attributions), and also rarely include constructs from other research areas in their studies (e.g., domain-specific skills and aptitudes).

Because the problem of student achievement is so complex, it makes sense that various subgroups of researchers would try to make this problem initially more tractable by examining individual or small sets of factors in their studies of achievement. Indeed, much has been learned about these aspects of achievement in the process. However, the continued tendency to focus on a limited number of predictors within each study of achievement has led to two related problems. One is that scientists and policy makers do not have a sense of how all of the various pieces of the achievement puzzle fit together. A second problem is that the relative importance of various predictors is still largely unknown because researchers have not typically included adequate controls in their studies. For example, when the present authors reviewed the literature to discover all longitudinal studies of science achievement that had been conducted in the past 15 years, they located 12 studies; of these, researchers examined the role of eight or fewer predictors ($M = 5$ predictors per study). It is nearly always the case that the correlation between a predictor and an achievement outcome is reduced when the effects of other potent predictors are statistically controlled. Without such controls, there is no way to tell the difference between important, authentic predictors and relatively minor or even spurious predictors. In our view, discovering and utilizing a comprehensive set of authentic predictors is the first step in building an accurate, multivariate theoretical model of achievement. Creating such a model, in turn, is the first step in developing more effective forms of intervention. Interventions can only make a difference if they target the most important sources of individual and group differences in achievement.

The primary goal of the present study was to promote the process of integrative theory-building in the area of academic achievement through a combination of top-down (i.e., logical) and bottom-up (i.e., empirical) methods. In doing so, we follow the lead of researchers from a variety of other disciplines (Kuhn, 1992). The pervasiveness of the logical and empirical approaches to scientific progress was first recognized in the philosophy of science literature when Reichenbach (1938) created his classic distinction between "the context of justification" (i.e., advancing science through logical argumentation) and "the context of discovery" (i.e., advancing science through data collection). In the area of linguistics, for example, considerable progress was made when linguists argued for the (logical) distinction between semantic, syntactic, and pragmatic aspects of language. Understanding word meaning (semantics) is not the same as knowing that third person singular requires an "-s" on the end of regular verbs (syntax) or knowing how to get one's way by being polite or complimentary (pragmatics). Once this distinction was agreed upon, language development researchers created theoretical models of these aspects and conducted studies to test these models and determine when children seemed to acquire expertise in semantics, syntax, and pragmatics (Hoff, 2001). In essence, top-down aspects of theory building are more meta-theoretical than theoretical. Meta-theoretical claims and assumptions function in ways similar to epistemological stances such as constructivism or empiricism by highlighting the kinds of phenomena that should be investigated (e.g., children's invention of concepts).

In the present case, the top-down (meta-theoretical) aspect of our approach involved a logical analysis of the categories of possible predictors and their interconnections. Consistent with other recent proposals (e.g., Byrnes, 2003; Corno et al., 2002), we began with the assumption that high achievement in a particular domain, such as math or science, is more likely to occur when two necessary conditions have been met: (a) children are given multiple opportunities to acquire knowledge and skills in that domain (the opportunity condition) and (b) children are able and willing to take advantage of these opportunities (the propensity condition). We define opportunities to learn as *culturally defined contexts in which an individual is presented with content to learn* (e.g., by a teacher or parent, an author, a narrator of an educational TV program, etc.) or *given opportunities to practice skills*. Thus, opportunities can occur both within school (e.g., in a classroom context) and outside of school (e.g., in the home context). Three questions one could ask about such contexts are (1) has the content required on achievement tests been presented in these contexts? (2) has this content been presented accurately? and (3) has the content been presented effectively? Thus, any variables related to exposure (e.g., coursework, a teacher's emphasis, homework, amount of repetition, etc.) or teaching quality (e.g., use of proven techniques, communication skills, classroom management, equitable treatment of students) would fall into the domain of an opportunity factor (Opdenakker, Van Damme, De Fraine, Van Landeghem, & Onghena, 2002; Pressley, Wharton-McDonald, & Rafael, 2002; Tate, 1995).

Propensity factors, in contrast, are *any factors that relate to the ability or willingness to learn content once it has been exposed or presented in particular contexts*. Thus, cognitive factors such as intelligence, aptitude, cognitive level, and pre-existing skills would qualify, as would motivational factors such as interest, self-efficacy, values, and competence perceptions (Byrnes, 2003; Corno et al., 2002; Eccles, Wigfield, & Schiefele, 1998; Jones & Byrnes, 2006; Sternberg, Grigorenko, & Bundy, 2001). Self-regulation is a hybrid of cognitive (e.g., beliefs) and motivational (e.g., efficacy) orientations (Pintrich, 2000), so it would also qualify as a propensity factor. We further assume that when the opportunity and propensity conditions are fulfilled in an individual (i.e., they have been exposed to content in an effective manner and were willing and able to take advantage of this learning opportunity), higher achievement will follow directly. As a result, opportunity factors and propensity factors are considered to be *proximal causes* of achievement.

In essence, then, this framework is useful for pointing out the kinds of factors that *might* matter, in advance of data collection. To move beyond the level of meta-theory, however, it is important to collect data to see which of the proposed factors actually *do* predict achievement when other potent factors have been controlled. This framework is also useful for summarizing or organizing the extant empirical literature. Whereas many of the variables examined to date can be categorized either as aspects of learning contexts (opportunity factors) or as student characteristics related to their willingness or ability to take advantage of learning opportunities (propensity factors), we argue that some factors identified in the literature are more properly viewed as *factors that enable or explain the emergence of opportunities to learn or propensities*. To identify the latter factors, one asks questions such as, "How can we explain the fact that a particular child regularly finds him- or herself in contexts in which a competent teacher presents the content needed for achievement tests in an effective manner?" and "How can we explain the fact that a particular child entered these learning contexts with the skills and motivation needed to benefit from these opportunities?" We argue that factors such as family socio-economic status, parental educational expectations for their children, children's own educational expectations, and prior educational experiences provide answers to such questions (Eccles et al., 1998; Oakes, 1985; Roscigno, 2000). Because these factors operate earlier in time and explain the emergence of opportunities and propensities, we call these factors *distal* factors.

The chronological and logical arrangement of distal, opportunity, and propensity factors is depicted in Fig. 1. The testable parts of the model are the links (arrows) between particular factors. As the figure shows, we grouped factors that might be hypothesized to be predictive of achievement into distal, opportunity, and propensity categories using boxes with dashed lines to indicate that these categories are meta-theoretical claims. We assume that individual factors operate independently but conjointly to determine achievement outcomes. Use of solid line boxes might imply that we expect the factors to be different manifestations of the same latent factor (which we do not).

As noted above, the foregoing analysis of possible categories of predictors and their interrelations is largely a top-down enterprise (though many of the entries in these categories such as prior knowledge and motivation are derived from empirical research). The bottom-up or empirical aspect of theory-building comes into play when multivariate studies are conducted to determine (a) the relative importance of various factors in the three categories and (b) the total amount of variance explained by a set of predictors. Those fac-

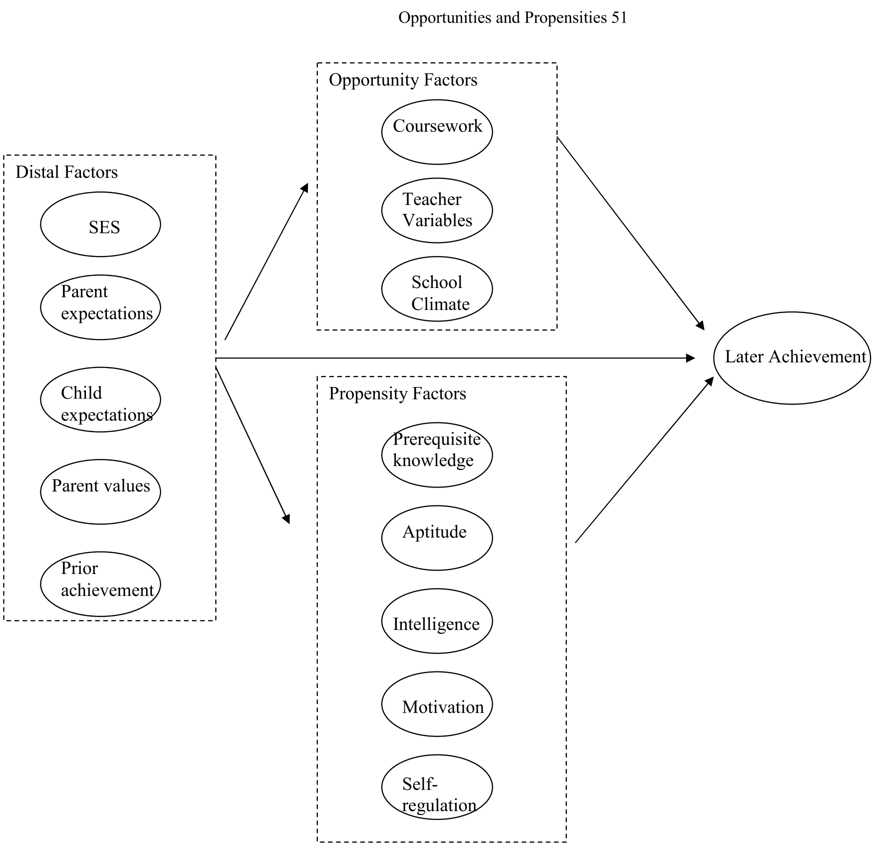

Fig. 1. An opportunity–propensity model of achievement.

tors that remain significant predictors after controls are retained in the model. Moreover, if the total amount of explained variance is less than 100%, follow-up exploratory studies are conducted to identify additional factors that may have been overlooked in the literature. The key to explaining substantial amounts of variance in achievement is to develop detailed and accurate definitions of what it means to have opportunities to learn and what it means to be prone to take advantage of these opportunities. In addition, it is important to use precise, reliable, and valid measures of the variables in question in order to maximize the explained variance and avoid wrongly assuming that a poorly measured factor does not matter.

As noted above, we define learning opportunities as culturally defined contexts occurring at particular times and locations (e.g., 4th-grade math class, soccer practice after school, etc.) when children are presented with content to learn and may be given opportunities to practice skills. Although an opportunity to learn is often equated with being enrolled in a course or school, not all opportunities are considered equal (Oakes, 1985; Tate, 1995). Clearly, children placed in the classrooms of highly skilled teachers are likely to progress faster in their development than children placed in classrooms of marginal teachers (Taylor, Pressley, & Pearson, 2000). Therefore, studies examining the role of opportunities should ideally consider the extent to which teachers are skilled classroom managers, use a balance of effective approaches, and so on. Because of possible teacher biases, however, it does not follow that children placed in the same classroom have been given equal opportunities (Leder, 1992; Tate, 1995). To insure that the role of opportunity is not underestimated, it is preferable to supplement indices such as coursework with more in-depth forms of measurement to enhance the amount of explained variance accounted for by opportunities to learn.

Even when opportunities have been carefully defined and measured, however, the opportunity–propensity (O–P) model assumes that indices of opportunities will only account for a portion of the variance in achievement because children still have to have the propensity to take advantage of any opportunities that may arise. In our view, there are three primary aspects of this propensity: (a) factors related to the *ability* to take advantage of learning opportunities (e.g., pre-existing skills, domain-specific aptitudes, and intelligence), (b) factors related to the *willingness* to take advantage of learning opportunities (i.e., motivational aspects such as interest, expectations, and self-efficacy), and (c) factors related to the construct of *self-regulation* (e.g., being organized, allocating study time effectively, and utilizing effective learning strategies). In other words, knowledge growth is expected to be maximized when students enter learning contexts with adequate levels of prerequisite knowledge, when they have high levels of domain-specific aptitude for learning the material, and when they are capable of processing various kinds of material quickly and strategically (Corno et al., 2002; Jones & Byrnes, 2006; Miller & Byrnes, 2001; Pintrich, 2000). Furthermore, knowledge growth will be maximized when students are interested in the material, when they consider the material to be relevant to their educational expectations, and when they feel that they are capable of mastering the material (Eccles et al., 1998).

With respect to distal factors that could be used to provide answers to the question, "Why were high achievers presented with more or better learning opportunities than low achievers?", prior research (Eccles et al., 1998; Oakes, 1985; Roscigno, 2000) suggests that opportunities to learn are a function of variables such as (a) parent education and income (i.e., well-educated, affluent parents are more equipped to recognize and provide significant educational resources for their children and to place them in high-quality educational environments than less educated, less affluent parents), (b) parent expectations (i.e., parents who have high educational and career goals for their children will likely choose high quality educational environments for them and provide substantial support and encouragement to their children for getting good grades), (c) child expectations (i.e., children are likely to seek out opportunities to learn if they have high educational expectations themselves), and (d) prior achievement (i.e., children who attain high levels of achievement in earlier grades either select more advanced learning opportunities or are placed in more advanced tracks by their teachers).

However, these same distal factors would also be expected to account for higher propensity levels in high achievers as well. For example, children who have the pre-requisite knowledge they need to benefit from instruction as 9th- or 10th-graders are likely to have parents who held high educational expectations for them when they were in earlier grades (e.g., 8th-grade). These same children would be likely to have high expectations for themselves and come from higher SES homes. In sum, then, the distal factors lead to the emergence of opportunities and propensities; opportunities and propensities in turn lead to higher levels of achievement (see Fig. 1).

In addition, we hypothesize that achievement differences among demographic subgroups of the population (e.g., males versus females; white students versus minority students) largely reflect the fact that gender and race/ethnicity are confounded with distal, opportunity, and propensity factors. As such, it is likely that gender and race/ethnicity will not explain a substantial amount of variance in achievement outcomes once distal, opportunity, and propensity factors have been controlled.

Ideally, the optimal design for examining the merits of all of the aforementioned hypotheses should have the following features: First, the O–P account necessitates the use of a comprehensive longitudinal design in which researchers assess factors in the distal group at a particular point in time (e.g., early adolescence) and then assess opportunities, propensities, and achievement levels at later points in time (e.g., middle and late adolescence). With such a design, researchers can determine whether the distal factors predict the proximal factors, and whether the proximal factors, in turn, predict the developmental outcomes. Many of the distal factors also have enduring effects (e.g., family income), so it is important to test this possibility (as indicated by the arrow from the leftmost distal box to the developmental outcome box in Fig. 1). The second feature of the optimal design is that the factors listed in Fig. 1 should be measured in the same study to see if they continue to explain variance after controls. Moreover, the indices of these factors should measure all aspects of a construct and be precise and valid. Third, the measures should be given to a large, representative sample of adolescents to ensure that the findings generalize to the larger population of interest.

Such an ideal design would be rather difficult to achieve given the typical resources of a team of researchers. Fortunately, the federal government has conducted several large-scale longitudinal studies that have generated databases that contain indices of many of the factors identified in Fig. 1. One such study, the National Education Longitudinal Study of 1988 (NELS:88), followed a nationally representative sample of over 12,000 adolescents for 12 years. Children in the study were first surveyed in 1988 when the respondents were in the 8th-grade and most were 14 years old. They were then resurveyed in four follow-ups that were conducted in 1990, 1992, 1994, and 2000. At the earlier waves of testing, respondents answered questions pertaining to various issues such as their performance in school and their expectations. Parents completed questionnaires that tapped into other factors such as their occupation and their educational expectations for their child. Hence, the NELS dataset provides information on many of the distal factors identified in the O–P model (the leftmost box in Fig. 1). As for indices of proximal factors (the column of middle boxes in Fig. 1), NELS includes information on opportunities and propensities pertinent to academic skills (e.g., coursework and academic self-concept). Finally, the dataset contains information on important developmental outcomes such as students' achievement in four academic subject areas. In the present study, we used the model to predict both 10th-grade and 12th-grade achievement in math and science given our interest in determining whether gender and race/ethnicity remain significant predictors of achievement after controlling for distal, opportunity, and propensity factors. Hence, the NELS dataset largely fulfills the requirements of the optimal design for examining the conjoint effects of distal, opportunity, and propensity factors, though not all of the factors that we wanted to examine were assessed in NELS and some were assessed with more precision than others (a point we return to later). Nevertheless, the dataset allows one to consider the utility of examining the individual and collective roles of a wide range of distal, opportunity, and propensity factors. Should imprecision of measurement be a problem for any of these factors, the result will be a reduction in the amount of explained variance accounted for by these factors. At worst, a relatively weak predictor will be judged (wrongly) to have no effect while stronger, significant predictors will be said to have a somewhat weaker effect than they actually have. Regardless, the results of our analyses should provide extremely useful information to researchers interested in conducting follow-up, prospective studies that are designed to continue the theory-building efforts initiated here. Given limited resources, researchers may opt to limit their focus in follow-up studies to just those predictors that proved to be significant or strong predictors in our study; others may strive to enhance measurement precision.

In addition to issues of imprecision, though, a possible other shortcoming of NELS:88 is that the data examined in the present study are approximately 15 years old. The age of the data would be an issue if there were reason to think that any of the important variables in the four categories (i.e., distal, opportunity, propensity, outcome) or the interrelations among these variables have undergone significant change since 1992. We see no inherent reason to suspect that parent expectations, family SES, grade point averages, coursework, motivation, and knowledge in math or science are fundamentally different in their basic nature. In addition, there is very good empirical evidence that little has changed in the content covered or teaching styles of teachers in areas such as math and science (see Jacobs et al., 2006). Moreover, long-term trends in NAEP scores show that absolute levels of student skills and abilities have changed only modestly and researchers measure motivation in 2006 in much the same way they did in the early 90s.

We selected variables from the NELS:88 database for consideration if (a) they met the definitional criteria (see above) for the categories of distal, opportunity, and propensity factors and (b) other researchers found that these variables were correlates of achievement in prior studies. Our goal was to account for as much variance as possible using these prior studies and the O–P framework as a guide. We expected that whereas some of these variables would account for a significant amount of variance even when many other factors were statistically controlled (e.g., prior knowledge), other variables would not account for substantial variance after controls (e.g., gender). In this way, we hoped to reveal the set of predictors that seem to be more important and, therefore, the central elements of a multifactorial model of achievement.

2. Method

2.1. Data source and participants

Data are from the NELS:88. Sponsored by the National Center for Education Statistics (NCES), NELS:88 provides longitudinal data about the critical transitions experienced by the 8th-grade class of 1988 as they developed, attended school, embarked on careers, and formed families. The survey was initiated with a nationally representative sample of students attending traditional public and private schools in the United States who did not have handicapping conditions that prevented them from completing the survey. The first wave of data collection for NELS began in the Spring of 1988. Almost 25,000 students from about 1000 schools (800 public and 200 private) were selected using a two-stage, stratified probability design, with schools as the first-stage unit and students within schools as the second-stage unit (Curtin, Ingels, Wu, & Heuer, 2002). The first follow-up was conducted in 1990, when most respondents were in 10th-grade, and the second follow-up was conducted in 1992, when most respondents were in their senior year of high school. We attempted to predict 10th-grade and 12th-grade science and math achievement using variables that reflected factors operative prior to that point (e.g., parental expectations in 8th-grade and coursework taken in 9th-grade).

As mentioned above, NELS employs a complex sampling design, with respondents nested within schools and schools nested within metropolitan statistical areas. In addition, there was deliberate over-sampling of certain subgroups of the population (i.e., specific racial/ethnic minority groups) and types of schools (i.e., private schools) to increase the precision of estimates for these low-frequency segments. To account for this unequal probability of selection and reduce the level of bias in estimates, sampling weights were used. The sampling weights also adjust for participant non-response (see Ingels et al., 1994, for further details about the study design and sample). All estimates in this report were produced using the panel weight for 8th-grade members of the NELS sample who also participated in the first- and second follow-ups. We also narrowed the sample to those individuals still in school at the first follow-up, when most were in 10th-grade. The unweighted sample size was 15,855 cases, representing approximately 2.8 million members of the 8th-grade class of 1988 who were in school 2 years later and participated in the 1990 and 1992 follow ups.

Also as a result of the complex sampling design of NELS, it is important to make certain adjustments in statistical estimates when conducting analyses. Notably, the characteristics of students within each school, including their achievement scores, share similarities because of their similar backgrounds and experiences. This enhanced homogeneity creates standard errors that are thought to be biased and smaller than they would be if unstratified random sampling were utilized (Johnson, 1989; Linn & Dunbar, 1992). Combined with the large sample size of NELS, the problem of slightly biased standard errors increases the chances of detecting significant predictors that may have little practical value. To reduce the level of bias, it is recommended that researchers utilize alternative algorithms for generating standard errors such as Jackknifing or Taylor Series approximations. The regression analyses carried out in this report used the Taylor Series procedure to calculate standard errors, and were done using the AM statistical software package (Cohen et al., 2003), which can be downloaded for no charge (http://am.air.org). In addition, we assessed the fit of several structural equation models (SEMs) using version 4.1 of the

Mplus commercial software that also makes adjustments for complex nested designs (Muthén & Muthén, 2006).

In the present study, respondents were about evenly divided by gender. As for race/ethnicity, 72% were White, 13% were Black, 10% were Hispanic, 4% were Asian/Pacific Islander, and 1% were Native American.

2.2. Analytical approach

In the present study, we attempted to predict 10th-grade and 12th-grade math and science achievement scores using hierarchical regressions and SEM. For consistency and in order to demonstrate how specific indices of distal and proximal factors are more or less predictive of various outcomes as postulated by the O–P model, we entered the same or a very similar set of predictors into each hierarchical regression. In the next section, we describe the measures that we used as indices of distal factors, opportunity factors, and proximal factors. The order of description that we use moves from right to left (back in time, starting with the developmental outcome—achievement) in Fig. 1. The reader may also wish to consult Tables 2–5 to have a clear idea of the predictors in the specific categories of distal, opportunity, and propensity.

2.3. Developmental outcomes

2.3.1. Tenth-grade and twelfth-grade math achievement

Math achievement scores were derived from standardized assessments. NCES commissioned a technical review panel comprised of subject matter experts, university faculty, and experts in test construction from the Educational Testing Service (ETS). As detailed in several technical reports (e.g., Owings, 1995), this panel was asked to create items that would (a) reflect the course content normally covered in typical American high schools, (b) range from relatively easy to relatively hard, and (c) result in a battery that had excellent psychometric properties (e.g., high reliability, high discriminant validity, little evidence of racial/ethnic or gender bias). In addition, given the longitudinal design of NELS, the assessment was constructed to have some overlap in the items between successive waves of data collection (some old, some new) to allow for assessment of growth in skills. The content coverage for the math items included arithmetic, algebra, geometry, and data analysis/probability. These content areas were then crossed with three types of items: those that required factual knowledge or procedural skills, those that required understanding of mathematical concepts or principles, and those that required problem solving. The expert panel reached consensus as to the task requirements of each item, but, to our knowledge, they did not conduct a formal assessment of item categorizations (e.g., computing kappas). After field-testing a pool of 82 potential items, students were eventually given an assessment in which they were asked to complete 40 questions in 30 min. The score assigned to students was constructed through an application of Item Response Theory (IRT) methods to the number of items a student correctly answered.

The technical manuals also document the excellent psychometric properties of the instrument. For example, the reliabilities for the math assessments were .89, .93, and .94 for base year, first follow-up, and second follow-up, respectively (Owings, 1995). Evidence for discriminant and convergent validity derive from the fact that correlations between adjacent administrations within the same content areas (e.g., 8th- and 10th-grade math) tend to be higher than those found between different content areas within the same administration (e.g., 8th-grade math and 8th-grade science). Also, the gain scores were higher for students taking courses at the time of the assessment as opposed to several years before, and for students who earned better grades in these courses (as determined from transcripts). With respect to content validity, teams of curriculum experts from school systems, the ETS, and colleges of education identified content that should be reflected on the test because this content is normally taught in high school programs. Finally, Differential Item Functioning analyses showed that there was very little evidence of gender or racial/ethnic bias in content.

At the 10th-grade, the overall mean was 43.22 (range = 16.37–72.76) and the means for males and females were 43.50 and 42.94, respectively ($d = .04$); Analyzed by racial/ethnic group, the means were: 49.05 (Asian), 45.35 (White), 37.28 (Hispanic), 34.80 (Black), and 34.19 (Native American). The largest difference between Asian and Native American students yields an effect size of $d = 1.08$. At the 12th-grade, the overall mean was 48.20 (range = 16.77–78.10) and the means for males and females were 48.84 and 47.54, respectively ($d = .09$). For the five racial/ethnic groups, the means were 53.56 (Asian), 50.36 (White), 42.10 (Hispanic), Native American (39.41) and Black (39.16). The largest difference between Asian and Black students yielded an effect size of $d = 1.01$.

2.3.2. Tenth-grade and twelveth-grade science achievement

The process for constructing the science assessment was very similar to that used for the math assessment. It, too, had excellent psychometric properties (e.g., reliabilities of.73,.81, and.82, respectively, for the base year, first follow-up, and second follow-up, respectively). There were 25 items in the final battery that were designed to assess concepts and reasoning in earth science, chemistry, scientific methods, life science, and physical science. Students were given 20 min to complete the science assessment and were assigned an IRT score. The overall means at the 10th- and 12th-grades were 21.49 and 23.38, respectively (ranges = 10.05–34.68 and 10.03–35.96, respectively). The gender gap at the 10th-grade yielded an effect size of $d = .27$ (M's of 22.30 for males and 20.66 for females); at the 12th-grade, the effect size was slightly larger at $d = .30$ (M's of 24.31 and 22.43). The largest effect size between racial/ethnic groups at the 10th-grade was $d = .92$ for the Asian versus Black contrast (M's= 22.84 and 17.34); the largest effect size between racial/ethnic groups at the 12th-grade was $d = .94$, again for the Asian versus Black contrast (M's = 24.54 and 18.74).

2.4. Proximal predictors

To predict math and science achievement scores at the 10th- and 12th-grades, we considered the role of several kinds of opportunity factors and propensity factors. In what follows, we describe the factors in each of these two categories in turn.

2.4.1. Opportunities

To provide an index of opportunities to learn math and science skills, we identified items from the 10th-grade questionnaire pertaining to (a) coursework in mathematics and science, (b) students' perceptions of the issues emphasized by their teachers in their math and science courses, and (c) students' perceptions of the responsiveness of teachers. There were several items corresponding to coursework in general math, pre-algebra, geometry, algebra I, algebra II, trigonometry, pre-calculus, calculus, and business math. However, only a small percentage of students took courses in trigonometry and calculus, and preliminary analyses showed that only three courses explained unique variance in math achievement: *general math*, *geometry*, and *algebra II*. These findings are to be expected given the content of items on the achievement test (see above). To maximize sample size and limit the analyses to the unique predictors, coursework was indexed using just the three items relating to general math, geometry, and algebra II. The wording of the question was, "From the beginning of 9th-grade to the end of this school year, how much coursework will you have taken in each of the following subjects?" For each course, the choices were "none," "1/2 year," "1 year," "$1\frac{1}{2}$ years," and "2 years." Student responses to these items were verified using an analysis of transcripts (Owings, 1995).

For science courses, it was again the case that only a small percentage of students took certain courses and only three courses were predictive. To maximize sample size and limit the analyses to just the unique predictors, science coursework was indexed using three items relating to *general science*, *biology*, and *chemistry*. The wording of these items was identical to that described above for math courses.

In preliminary analyses, we observed that there were prototypical patterns of course-taking that few students deviated from, and that taking more semesters than typical probably indicated that less skilled students were retaking courses. As a result, achievement was not monotonically related to number of semesters. To deal with this fact, we used dummy codes (0, 1) to reflect four categories of course-taking: no courses, one semester, one full year (the prototype), and more than one year. For example, the code for one full year was "0 1 0." The reference category was no courses (0 0 0) and the dummy coded variables listed in Tables 2–5 reflect the remaining three categories.

In addition to considering students' enrollment in math and science courses, we also considered the approach and emphases of their instruction as indicators of being provided authentic opportunities to learn. The focal item regarding math emphasis was worded as follows: "In your most recent or current mathematics class, how much emphasis does/did your teacher place on each of the following objectives?" The objectives were (a) "increasing your interest in math," (b) "learning and memorizing facts, rules, and steps," (c) "preparing you for further study in math," (d) "thinking about what a problem means and ways it might be solved," and (e) "showing you the importance of mathematics in daily life." The options were "none," "minor emphasis," "moderate emphasis," and "major emphasis." There was an identical item for science courses. Given contemporary descriptions of effective teaching that emphasize a balance of meaning-based and practice-based approaches (e.g., Taylor et al., 2000), students were assigned one point for each time they indicated a major emphasis on objectives (b), (c), and (d), and either a minor or moderate emphasis on objectives (a) and (e). Therefore, the total score for *student perceptions of math emphasis* and *student perceptions of science emphasis* could range from 0 to 5. Higher scores indicate that students perceived their teachers as emphasizing regular practicing of skills, problem-solving, and conceptual understanding. The means were 2.69 and 2.13 for math and science, respectively. Teachers who fell at the extreme ends of the scales either promoted rote-learned skills at the expense of concepts and problem-solving, or emphasized concepts, problem-solving, and interest at the expense of requiring adequate practice and repetition.

The *student perceptions of teacher responsiveness* aspect of opportunities was derived from student responses to the following item: "How much do you agree with the following statements about your current school and teachers?" The statements included "students

Table 1
Selected intercorrelations among predictors

Within distal	(1)	(2)	(3)	(4)
(1) SES	—	.42	.38	.29
(2) Parent expectations		—	.50	.38
(3) Student expectations			—	.43
(4) Middle school GPA				—

Within opportunity (math)	(1)	(2)	(3)	(4)	(5)
(1) General math (1 year)	—	−.26	−.15	−.11	−.08
(2) Geometry (1 year)		—	.25	.18	.13
(3) Algebra II (1 year)			—	.11	.10
(4) Math emphasis				—	24
(5) Teacher responsiveness					—

Within opportunity (science)	(1)	(2)	(3)	(4)	(5)
(1) General science (1 year)	—	−.09	−.11	−.05	−.08
(2) Biology (1 year)		—	.14	.08	.07
(3) Chemistry (1 year)			—	.15	.08
(4) Science emphasis				—	20
(5) Teacher responsiveness					—

Within propensity	(1)	(2)	(3)	(4)	(5)	(6)
(1) Math achv. prior to 9th-grade	—	.72	.38	.39	.23	.35
(2) Science achv. prior to 9th-grade		—	.25	.35	.20	.30
(3) Math GPA (9th, 10th-grades)			—	.46	.24	.22
(4) Science GPA (9th, 10th-grades)				—	.28	.30
(5) Efficacy for graduating					—	.23
(6) Plans to take the SAT						—

Distal with opportunity (math)	(5)	(6)	(7)	(8)	(9)
(1) SES	−.19	.31	.18	.11	.06
(2) Parent expectations	−.21	.34	.22	.15	.14
(3) Student expectations	−.21	.33	.21	.14	.15
(4) Middle school GPA	−.28	.45	.32	.20	.17
(5) General math (1 year)					
(6) Geometry (1 year)					
(7) Algebra II (1 year)					
(8) Math emphasis					
(9) Teacher responsiveness					

Distal with opportunity (science)	(5)	(6)	(7)	(8)	(9)
(1) SES	−.13	.18	.18	.07	.06
(2) Parent expectations	−.16	.18	.18	.10	.14
(3) Student expectations	−.15	.18	.20	.11	.15
(4) Middle school GPA	−.20	.27	.26	.12	.17
(5) General science (1 year)					
(6) Biology (1 year)					
(7) Chemistry (1 year)					
(8) Science emphasis					
(9) Teacher responsiveness					

Table 1 (*continued*)

Distal with propensity	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) SES	.43	.38	.16	.21	.19	.30	.09
(2) Parent expectations	.40	.34	.20	.25	.19	.35	.14
(3) Student expectations	.38	.34	.21	.30	.23	.37	.16
(4) Middle school GPA	.53	.44	.42	.51	.27	.37	.28
(5) Math achv. prior to 9th-grade							
(6) Science achv. prior to 9th-grade							
(7) Math GPA (9th, 10th-grades)							
(8) Science GPA (9th, 10th-grades)							
(9) Efficacy for graduating							
(10) Plans to take the SAT							
(11) Math self-concept							

Opportunity with propensity (math)	(6)	(7)	(8)	(9)	(10)
(1) General math (1 year)	−.31	−.12	−.13	−.18	−.10
(2) Geometry (1 year)	.50	.25	.22	.31	.20
(3) Algebra II (1 year)	.38	.21	.13	.22	.20
(4) Math emphasis	.20	.19	.13	.16	.20
(5) Teacher responsiveness	.12	.21	.16	.17	.21
(6) Math achv. prior to 9th-grade					
(7) Math GPA (9th, 10th-grades)					
(8) Efficacy for graduating					
(9) Plans to take the SAT					
(10) Math self-concept					

Opportunity with propensity (science)	(6)	(7)	(8)	(9)
(1) General science (1 year)	−.18	−.14	−.11	−.14
(2) Biology (1 year)	.20	.15	.17	.19
(3) Chemistry (1 year)	.23	.17	.09	.18
(4) Science emphasis	.12	.13	.09	.14
(5) Teacher responsiveness	.10	.21	.16	.17
(6) Science achv. prior to 9th-grade				
(7) Science GPA (9th, 10th-grades)				
(8) Efficacy for graduating				
(9) Plans to take the SAT				

Note: Because of the large sample size, all correlations greater than $r = \pm.05$ are statistically significant.

get along well with teachers," "discipline is fair," "the teaching is good," "teachers are interested in students," "when I work hard on my schoolwork, my teachers praise my effort," "in class, I often feel 'put down' by my teachers," and "most of my teachers really listen to what I have to say." Response options ranged from "strongly agree" to "strongly disagree." The resultant index was a factor score derived from a principal components analysis of responses to these statements ($\alpha = .78$). There was a one-factor solution that explained 44% of the variance. Higher derived factor scores correspond to more perceived teacher responsiveness. Negatively worded items were reverse scored. The mean was 9.94 (range 1.65–15.43).

2.4.2. Propensities

In addition to examining the predictive role of opportunities pertaining to math and science achievement, we also considered the role of the propensity to take advantage of

Table 2
Final model for predicting 10th-grade math achievement

Predictor	Coefficient	*t*-value	Correlations	
			Zero	Partial
Distal factors				
Socio-economic status, 8th-grade	.458	3.386**	.427	.048
Parent expectations for child in 8th-grade	.284	2.492*	.406	.035
Student expectations in 8th-grade	.303	2.687**	.401	.035
Middle school GPA	.928	5.787***	.554	.082
Opportunity factors				
General math coursework (1/2 year)	−2.292	−2.775**	−.107	−.057
General math coursework (1 year)	−2.402	−9.479***	−.338	−.124
General math coursework (1.5 to 2 years)	−2.354	−6.894***	−.188	−.101
Geometry coursework (1/2 year)	1.214	3.497***	.008	.040
Geometry coursework (1 year)	2.246	9.432***	.525	.147
Geometry coursework (1.5–2 years)	3.308	3.427**	.017	.052
Algebra II coursework (1/2 year)	1.776	4.693***	.107	.059
Algebra II coursework (1 year)	1.130	4.966***	.376	.072
Algebra II coursework (1.5–2 years)	−.049	−.040 (n.s.)	.008	−.001
Student percept. of math emphasis	.260	3.893***	.217	.053
Student percept. of teacher responsiveness	.044	.856 (n.s.)	.153	.014
Propensity factors				
Math achievement before start of 9th-grade	.773	76.270***	.874	.726
Math GPA in 9th- and 10th-grades	.573	4.257***	.419	.062
Efficacy for graduating high school	1.015	4.254***	.264	.057
Plans to take the SAT	.617	3.145**	.375	.043
Math self-concept	.113	5.233***	.371	.068
Demographic factors				
Gender (female)	−.674	−4.086***	−.051	−.055
Race/ethnicity dummy1 (Black)	−2.205	−6.712***	−.207	−.104
Race/ethnicity dummy2 (Hispanic)	−1.076	−3.813***	−.134	−.048
Race/ethnicity dummy3 (Asian)	.038	.097 (n.s.)	.083	.001
Race/ethnicity dummy4 (Native American)	−1.700	−3.159**	−.052	−.026

Note: Variable labels are explained in the text: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; $N = 10{,}934$; "zero" means zero-order (uncorrected) correlations; partial correlations are correlations between the predictor and outcome with all other variables statistically controlled; the partial correlations for variables with significant *t*-values are significant as well.

opportunities. In the case of math achievement, we examined the NELS database for any indicators of the self-regulatory, cognitive, and motivational aspects of the propensity to take advantage of opportunities to learn math. The NELS database did not have specific indices of self-regulated learning. As for indices of cognitive propensity, we hypothesized that students would demonstrate faster rates of knowledge growth in their high school courses if they entered them with higher levels of prerequisite knowledge, higher levels of aptitude, and higher levels of general intelligence. The NELS database did not have distinct indices of each of these cognitive aspects, but it did have (a) *math achievement before start of 9th-grade*-students' performance on the standardized math assessment given at the end of 8th-grade (completed a few months before they entered 9th-grade; it indexes the math skills they brought to courses in 9th- and 10th-grade), and (b) *math GPA in 9th- and 10th-grades*-students' self-reported grade point average (GPA) for math courses in the 9th-grade until

Table 3
Final model for predicting 12th-grade math achievement

Predictor	Coefficient	*t*-value	Correlations	
			Zero	Partial
Distal factors				
Socio-economic status, 8th-grade	.790	5.039***	.423	.074
Parent expectations for child in 8th-grade	.478	3.628***	.412	.052
Student expectations in 8th-grade	.409	2.798**	.408	.042
Middle school GPA	1.215	6.488***	.559	.095
Opportunity factors				
General math coursework (1/2 year)	−2.826	−3.408**	−.116	−.060
General math coursework (1 year)	−3.338	−10.572***	−.351	−.151
General math coursework (1.5–2 years)	−3.477	−8.837***	−.193	−.130
Geometry coursework (1/2 year)	1.620	3.821***	.007	.045
Geometry coursework (1 year)	2.660	10.462***	.530	.154
Geometry coursework (1.5–2 years)	2.237	1.899 (n.s.)	−.002	.035
Algebra II coursework (1/2 year)	1.374	2.745**	.098	.039
Algebra II coursework (1 year)	.858	3.262**	.369	.049
Algebra II coursework (1.5–2 years)	.261	.167 (n.s.)	.006	.003
Student percept. of math emphasis	.157	2.075*	.199	.028
Student percept. of teacher responsiveness	.132	2.160*	.170	.037
Propensity factors				
Math achievement before start of 9th-grade	.678	54.596***	.832	.638
Math GPA in 9th- and 10th-grades	.943	5.859***	.451	.088
Efficacy for graduating high school	.915	3.126**	.269	.044
Plans to take the SAT	1.180	5.395***	.385	.073
Math self-concept	.175	6.184***	.405	.092
Demographic factors				
Gender (female)	−1.725	−9.228***	−.079	−.124
Race/ethnicity dummy1 (Black)	−2.184	−5.125***	−.212	−.090
Race/ethnicity dummy2 (Hispanic)	−.761	−2.279*	−.120	−.030
Race/ethnicity dummy3 (Asian)	.779	1.387 (n.s.)	.080	.022
Race/ethnicity dummy4 (Native American)	−2.078	−1.937 (n.s.)	−.038	−.027

Note: Variable labels are explained in the text: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; $N = 8969$; "zero" means zero-order (uncorrected) correlations; partial correlations are correlations between the predictor and outcome with all other variables statistically controlled; the partial correlations for variables with significant *t*-values are significant as well.

the time of report in 10th-grade. The accuracy of student self-reports of grades was confirmed through analyses of transcripts (Owings, 1995). Likewise, the database included students' *science achievement before start of 9th-grade* and *science GPA in 9th- and 10th-grades*. Whereas together these cognitive indices measure the knowledge that students brought to their classes, we do not preclude the possibility that they also may tap into students' intelligence and aptitude for math and science as well (though, to a lesser degree because the achievement tests provided adequate time and were designed to reflect course content rather than require exceptional levels of skill). The means were 35.88 (8th-grade math achievement), 18.72 (8th-grade science achievement), 2.76 (math GPA), and 2.79 (science GPA).

With respect to the motivational aspect of propensity, we used a 10th-grade item that assessed a student's *efficacy for graduating high school* ("How sure are you that you will graduate from high school?") to predict both math and science achievement. The response

Table 4
Final model for predicting 10th-grade science achievement

Predictor	Coefficient	*t*-value	Correlations	
			Zero	Partial
Distal factors				
Socio-economic status, 8th-grade	.569	7.837***	.397	.095
Parent expectations for child in 8th-grade	.212	3.017**	.326	.042
Student expectations in 8th-grade	.137	1.880 (n.s.)	.325	.026
Middle school GPA	.576	6.668***	.437	.082
Opportunity factors				
General science courses (1/2 year)	−.510	−1.566 (n.s.)	−.057	−.020
General science courses (1 year)	−.701	−5.281***	−.203	−.069
General science courses (1.5–2 years)	−.403	−1.499 (n.s.)	−.081	−.019
Biology courses (1/2 year)	−.538	−2.213*	−.107	−.027
Biology courses (1 year)	.242	1.463 (n.s.)	.195	.020
Biology courses (1.5–2 years)	−.160	−.452 (n.s.)	−.020	−.007
Chemistry courses (1/2 year)	−.070	−.256 (n.s.)	.015	−.003
Chemistry courses (1 year)	.447	3.221**	.219	.041
Chemistry courses (1.5–2 years)	−.135	−.218 (n.s.)	−.008	−.002
Student percept. of science emphasis	.038	.801 (n.s.)	.109	.011
Student percept. of teacher responsiveness	.090	3.518***	.134	.047
Propensity factors				
Science achievement before start of 9th-grade	.677	51.596***	.732	.580
Science GPA in 9th- and 10th-grades	.530	9.039***	.385	.108
Efficacy for graduating high school	.426	2.814**	.218	.037
Plans to take the SAT	.490	3.763***	.305	.055
Demographic factors				
Gender (female)	−1.343	−13.212***	−.178	−.174
Race/ethnicity dummy1 (Black)	−2.105	−8.018***	−.236	−.157
Race/ethnicity dummy2 (Hispanic)	−1.084	−5.720***	−.140	−.077
Race/ethnicity dummy3 (Asian)	−.188	−.924 (n.s.)	.046	−.009
Race/ethnicity dummy4 (Native American)	−.878	−1.833 (n.s.)	−.037	−.021

Note: Variable labels are explained in the text: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; $N = 10{,}854$; "zero" means zero-order (uncorrected) correlations; partial correlations are correlations between the predictor and outcome with all other variables statistically controlled; the partial correlations for variables that have significant *t*-values are significant as well.

options for this item ranged from "very sure I won't" to "very sure I will" ($M = 2.84$ out of 3). In addition, we used a 10th-grade item that asked about students' *plans to take the SAT* as an indicator of their educational expectations. From the five response categories (distinguished mainly by when they planned on taking it), we recoded responses into two categories: "Yes, I plan to take it" (coded 1; 61%) and "No, haven't thought about it/don't plan to take it" (coded 0; 39%). Furthermore, as a specific predictor of 10th-grade and 12th-grade math achievement, we used the factor score derived from a principal components analysis of the following four *math self-concept* items taken from the well-regarded Self-Description Questionnaire of Marsh and colleagues (e.g., Marsh & O'Neill, 1984; $\alpha = .88$ in the present sample): "Mathematics is one of my best subjects," "I have always done well in mathematics," "I get good marks in mathematics," and "I do badly on tests

Table 5
Final model for predicting 12th-grade science achievement

Predictor	Coefficient	*t*-value	Correlations	
			Zero	Partial
Distal factors				
Socio-economic status, 8th-grade	.488	4.482***	.376	.076
Parent expectations for child in 8th-grade	.413	4.288***	.334	.076
Student expectations in 8th-grade	.127	1.441 (n.s.)	.332	.022
Middle school GPA	.616	5.149***	.429	.081
Opportunity factors				
General science courses (1/2 year)	−.117	−.346 (n.s.)	−.039	−.004
General science courses (1 year)	−.731	−3.670***	−.202	−.066
General science courses (1.5–2 years)	−1.146	−3.576***	−.098	−.049
Biology courses (1/2 year)	−.483	−1.485 (n.s.)	−.103	−.022
Biology courses (1 year)	.146	.621 (n.s.)	.192	.011
Biology courses (1.5–2 years)	−.352	−.808 (n.s.)	−.019	−.014
Chemistry courses (1/2 year)	−.351	−1.098 (n.s.)	.024	−.012
Chemistry courses (1 year)	.360	2.338*	.207	.031
Chemistry courses (1.5–2 years)	−.949	−.681 (n.s.)	−.022	−.012
Student percept. of science emphasis	.052	.823 (n.s.)	.100	.014
Student percept. of teacher responsiveness	.081	2.379*	.132	.038
Propensity factors				
Science achievement before start of 9th-grade	.657	37.670***	.705	.537
Science GPA in 9th- and 10th-grades	.582	7.879***	.384	.108
Efficacy for graduating high school	−.010	−.049 (n.s.)	.193	−.001
Plans to take the SAT	.546	3.619***	.296	.057
Demographic factors				
Gender (female)	−1.640	−13.544***	−.187	−.197
Race/ethnicity dummy1 (Black)	−2.657	−12.376***	−.257	−.182
Race/ethnicity dummy2 (Hispanic)	−1.181	−5.728***	−.137	−.079
Race/ethnicity dummy3 (Asian)	.055	.227 (n.s.)	.043	.003
Race/ethnicity dummy4 (Native American)	−1.010	−1.274 (n.s.)	−.026	−.021

Note: Variable labels are explained in the text: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; $N = 8940$; "zero" means zero-order (uncorrected) correlations; partial correlations are correlations between the predictor and outcome with all other variables statistically controlled; the partial correlations for variables with significant *t*-values are significant as well.

in mathematics." For these four items, the response options were "false," "mostly false," "more false than true," "more true than false," "mostly true," and "true." A single factor emerged that explained 75% of the variance in the self-concept items. Higher scores mean more positive self-concepts for math. Negatively worded items were reverse scored. There was not a comparable self-concept scale for science in the NELS database. The mean for math self-concept was 8.90 (range −1.70 to 15.50).

2.5. Distal predictors

Four indices of *distal factors* were judged to have the potential to be predictive of opportunities, propensities, and math and science achievement because, as noted earlier, they could plausibly explain why some students had been presented with more opportunities or entered high school with higher levels of propensity. The first was a composite variable of *socioeconomic status* (SES) that is available in the NELS database. This variable is constructed from the following information obtained in the base-year parent questionnaire: highest level of education for mother and father, occupation of mother and father, and family income. If parent-reported information was not available in the base year of the survey, then student-reported data from the base year or first or second follow up was used (substituting the number of selected items in the household for parental income in instances of student-reported data). Using this information, a single SES score was derived for each participant ($M = 0$; range from −2.93 to 2.75).

The second distal factor was an item taken from the parent questionnaire that asked about *parent expectations for the child in 8th-grade*, and the third distal factor was an item taken from the 8th-grade student questionnaire that reflected the *student's expectations in 8th grade*. Each item, respectively, was worded as follows: "How far in school do you expect your 8th-grader to go?" and "How far in school do you think you will get?" From the response options, we created the categories of high school or less (12% and 10%, respectively), vocational, technical or business school/some college (27% and 22%, respectively), finish college (40% and 45%, respectively), and graduate school (21% and 23%, respectively). The fourth distal factor was the student's *middle school GPA*, that is, an average of the student's grades in English, math, science, and social studies from the 6th-grade until the time of report in 8th-grade. This was a composite variable available in the NELS database. The mean GPA was 2.94, with a range from .5 to 4.0. Again, accuracy of the student self-reports was verified through transcript analysis (Owings, 1995). Whereas the 10th-grade expectations and GPA items described earlier may explain a child's degree of willingness and ability to benefit from learning opportunities (and therefore are indices of propensity), the 8th-grade expectations and middle-school GPA items described in this section are conceptualized as distal factors because they can explain why a student had specific learning opportunities after 8th-grade (e.g., taking particular courses in 9th- and 10th-grade) and took advantage of these opportunities.

2.6. Demographic factors

In addition to the collection of proximal and distal factors, we also considered the extent to which a student's *gender* and *race/ethnicity* predicted math achievement after the proximal and distal factors had been entered in a hierarchical regression.

Table 1 shows intercorrelations among various predictors and Tables 2–5 show the complete list of predictors used for each of the four outcomes.

3. Results

The results are organized as follows. In the first section, the same hierarchical regression approach was used to predict math and science achievement scores at the 10th- and 12th-grades. The ordering of variables was based on the chronological and logical arrangement of factors in the O–P framework. In the first step of each regression, the distal factors were entered because these include student and family background factors that were assessed earlier in time. As such, they are thought to enable the emergence of opportunities and propensities. In the second step, indices of opportunities were entered because students cannot take advantage of opportunities unless these opportunities present themselves first. With opportunities controlled, we then entered propensities to see the extent to which factors such as cognitive abilities and motivation predict who took advantage of opportunities. In the final step, gender and race/ethnicity were entered to determine the extent to which these factors are confounded with distal factors, opportunities, and propensities (and would, therefore, predict little meaningful variance in the final step).

In these hierarchical regressions, we did not examine all possible orders of entry due to the chronology of indices. For example, it would not make sense to predict factors assessed in the 8th-rade using factors assessed in the 10th- or 12th-grade. In addition, columns for zero-order correlations and partial correlations were included in tables for regressions. The zero order correlations are correlations between the predictor and outcome without other factors controlled. Partial correlations are correlations between a predictor and the outcome with all other factors controlled. Inspection of the two columns demonstrates the importance of controlling for other factors (see Tables 2–5).

In the second section, we describe the results of structural equation modeling (SEM) that was employed to extend the results of the hierarchical regressions by testing the significance of all direct and indirect paths in the complex models simultaneously. With regression, in contrast, one must examine only portions of such models in several successive analyses (e.g., direct first, then indirect second). The simultaneous approach is preferred because path coefficients are adjusted for all other factors, not just some of these factors. As such, the simultaneous approach is less likely to retain a spurious or relatively weak predictor than the multi-step approaches. In addition, we wanted to provide a test of the claim that distal factors predict opportunity and propensity factors which, in turn, predict achievement.

3.1. Hierarchical regression analyses

3.1.1. Tenth-grade math achievement

In the first step of the hierarchical regression for math achievement, the distal factors assessed in 8th-grade were entered as a block and found to account for 41.7% of the variance in 10th-grade achievement scores. When indices of opportunities to learn math (e.g., courses) were entered on the second step as a block, the amount of explained variance increased by 11.2% to a total of 52.9%. When indices of the propensity to learn math were entered on the third step, the amount of explained variance increased by 27.5–80.4%. On the final step, gender and race/ethnicity were found to add less than 1% of the variance after all of the other factors had been entered (total = 80.6%). The final model with numerous significant predictors is shown in Table 2.

As can been seen in Table 2, all four of the distal factors were significant predictors of 10th-grade math achievement. That is, students tended to attain higher levels of achievement in math when (a) they came from higher SES homes, (b) their parents had higher educational expectations for them, (c) they had higher educational expectations themselves, and (d) they had higher GPAs in middle school.

Controlling for these distal factors, the model suggests that students attained high levels of math achievement when (a) they avoided taking any general math courses, (b) they took at least one semester of geometry, (c) they took one or two semesters of algebra II, and (c) they had math teachers who emphasized regular practicing of skills, problem-solving, and conceptual understanding (see opportunity factors in Table 2).

Moreover, as indicated by the difference in additional explained variance for opportunity and propensity factors reported above (i.e., 11.2% vs. 27.5%), students did particularly well if they brought a high level of skill to their learning opportunities (i.e., indexed by math achievement scores before the start of 9th-grade), performed well in their math courses in the 9th- and 10th-grades (indexed by their math GPA), felt efficacious about graduating high school, planned to take the SAT, and held positive math self-concepts. Thus, opportunities to learn are clearly necessary, but students have to be able and willing to benefit from these classes to attain higher levels of achievement.

As for gender and race/ethnicity, the model indicates that, even after controls, males tended to perform better on the math assessment than females, and White students tended to perform better than Black, Hispanic, and Native American students. White students and Asian students did not differ. (Note again, though, that gender and race/ethnicity explained less than 1% of the variance after all other variables were controlled).

Collectively, these findings demonstrate the value of taking an integrative and multifactorial approach. The O–P framework combined with prior research helped us to identify the categories of predictors (distal, opportunity, propensity) and members of these categories that should be selected and evaluated. The findings also demonstrate the relative strength of the predictors as well (indicated by the *t*-values and correlations shown in Table 2).

The value of controlling for other factors also becomes apparent when the order of entry is modified somewhat. For example, when indices of opportunities were entered in the first step (without controlling for distal factors), the amount of explained variance was 44.1% rather than the 11.2% reported above when entering distal factors first. Similarly, when propensities were entered without controlling for distal or opportunity factors, the amount of explained variance was 77.9% rather than the 27.5% reported above. When gender and race/ethnicity were entered without controlling for distal, opportunity, and propensity factors, the amount of explained variance was 9.3% rather than less than 1%. Thus, in every case, controls lead to substantial reductions in explained variance.

3.1.2. Twelfth-grade math achievement

In the first step of the hierarchical regression for 12th-grade math achievement, distal factors were found to account for 43.0% of the variance (slightly more than the 41.7% found for 10th-grade math). In the second step, indices of opportunities added another 11.2% of the variance (the same as found for 10th-grade math). On the third step, indices of the propensity to take advantage of opportunities added an additional 21.9% of the variance (somewhat less than the 27.5% found for 10th-grade math). On the final step, gender and race/ethnicity again added less than one percent of the variance (total = 76.6%). Thus, there was a high degree of stability in the prediction of math achievement when assessed another two years later. As can be seen in comparing Table 3 to Table 2, the predictors were highly consistent at both the 10th- and 12th-grade. An additional predictor that emerged at grade 12 was students' perceptions of teacher responsiveness. That is, students who as 10th-graders perceived their teachers as more responsive tended to have higher achievement scores as 12th-graders.

To again show the importance of controlling for other factors, we varied the order of entry of factors for 12th-grade math. When indices of opportunities were entered in the first step (without controlling for distal factors), the amount of explained variance was 45.2% rather than the 11.2% reported above when entering distal factors first. Similarly, when propensities were entered without controlling for distal or opportunity factors, the amount of explained variance was 72.4% rather than the 21.9% reported above. When gender and race/ethnicity were entered without controlling for distal, opportunity, and propensity factors, the amount of explained variance was 9.3% rather than less than 1%.

3.1.3. Tenth-grade science achievement

On the first step of the hierarchical regression for 10th-grade science achievement, the distal factors were found to explain 29.8% of the variance (considerably less than the 41.7% found for 10th-grade math, but still substantial). On the second step, indices of opportunities added an additional 1.3% of the variance (much less than the 11.2% found for 10th-grade math). On the third step, indices of propensities added another 27.6% of the variance (essentially the same as the 27.5% added for 10th-grade math). On the final step, gender and race/ethnicity added an additional 2.1% of the variance (compared to .2% found for 10th-grade math, but still fairly modest). The full model explained a total of 60.8% of the variance.

As can be seen in Table 4, three out of the four distal factors were significant predictors of 10th-grade science in the same direction as was found for 10th-grade math and 12th-grade math. In contrast to math, science achievement was predicted by a limited number of indices of opportunities: taking 1 year of general science (negatively), taking one semester of biology (negatively), taking 1 year of chemistry, and judging teachers to be responsive. As for propensity factors, higher science achievement in 10th-grade was predicted by all of the measured factors: (a) higher science achievement before the start of 9th-grade, (b) getting better science grades in the 9th- and 10th-grades, (c) feeling efficacious about graduating high school, and (d) planning to take the SAT. This is the same pattern that was found for 10th-grade math and 12th-grade math. Finally, the findings suggest that males performed better than females, and that White students performed better than Black or Hispanic students. Asian and White students did not differ.

As was the case for math, entering blocks of factors without controls shows how the role of these factors could be overestimated. When indices of opportunities were entered in the first step (without controlling for distal factors), the amount of explained variance was 13.0% rather than the 1.3% reported above when entering distal factors first. Similarly, when propensities were entered without controlling for distal or opportunity factors, the amount of explained variance was 56.4% rather than the 27.6% reported above. When gender and race/ethnicity were entered without controlling for distal, opportunity, and propensity factors, the amount of explained variance was 12.6% rather than 2.1%.

3.1.4. Twelfth-grade science achievement

On the first step of the hierarchical regression for 12th-grade science, distal factors were found to account for 28.8% of the variance (slightly less than the 29.8% found for 10th-grade science but considerably less than the 43.0% found for 12th-grade math). On the second step, indices of opportunities added an additional .9% of the variance (slightly less than the 1.3% found for 10th-grade science but much less than the 11.2% found for 12th-grade math). On the third step, indices of propensities added an additional 24.9% of the variance (slightly less than the 27.6% found for 10th-grade science and slightly more than the 21.9% found for 12th-grade math). On the fourth step, gender and race/ethnicity added an additional 3.1% of the variance (total = 57.7%). The final model is shown in Table 5. As can be seen, most factors that were significant predictors for 10th-grade science were also significant predictors of 12th-grade science.

Finally, we once again varied the order of entry of variables to demonstrate the value of controlling other potent factors. When indices of opportunities were entered on the first step (without controlling for distal factors), the amount of explained variance was 13.1% rather than the .9% reported above when entering distal factors first. Similarly, when propensities were entered without controlling for distal or opportunity factors, the amount of explained variance was 52.3% rather than the 24.9% reported above. When gender and race/ethnicity were entered without controlling for distal, opportunity, and propensity factors, the amount of explained variance was 13.7% rather than 3.1%.

3.2. Structural equation models

The aforementioned hierarchical regressions primarily examined the right half of the O–P models for science and math achievement. That is, they largely considered the extent to which distal, opportunity, and propensity factors predicted achievement. However, we were also interested in determining the extent to which the entire model fit the data. That is, do distal factors predict opportunities and propensities, which in turn predict achievement? In each of the four models that we created to test this proposal, we included paths connecting each distal factor to each opportunity and propensity factor (the left half of models such as Fig. 1). These additional paths, then, describe the indirect effects of distal factors such as SES on achievement through opportunity factors (e.g., coursework) and propensity factors (e.g., math self-concept), but also test the claim that distal factors predict opportunities and propensities. The tests of the total model also served as replications of the regressions by including the direct effects of the distal, opportunity, and propensity factors on achievement as well (the right half of Fig. 1). The role of gender and race/ethnicity was examined by including paths for these factors as well. We assumed that each of the factors in the three categories operated independently but conjointly. We did, however, examine the opposite hypothesis by testing the fit of models that had single latent factors for the distal, opportunity, and propensity categories.

The results of each SEM analysis were highly consistent with the overall O–P framework, though individual paths were eliminated to improve the fit of some models. The results are presented in tabular form (Tables 6 and 7) rather than in figures because the large number of crisscrossing paths and coefficients made it hard to identify individual paths and see which were significant. z-scores are included in the tables rather than raw coefficients to put coefficients on the same scale so that their relative predictive power can be compared. Assessing the fit of each model presented in Tables 6 and 7 requires setting cutoff criteria that minimizes both Type I (too liberal) and Type II (too conservative) error. Recent analyses suggest that the minimum cutoff value for the CFI should be above.90 to avoid Type I error but be below .95 to avoid Type II error (Hu & Bentler, 1999). For SRMR and RMSEA, it is recommended that these indices be no higher than .08 and .06, respectively. Results showed that for the complete model for 10th-grade math achievement (Table 6, first column), the corresponding fit indices were .91 (CFI), .06 (SRMR), and .06 (RMSEA). For 12th-grade math achievement (Table 6, second column), the indices were .90 (CFI), .06 (SRMR), and .06 (RMSEA). For 10th-grade science (Table 7, first column), the indices were .91 (CFI), .05 (SRMR), and .04 (RMSEA). For 12th-grade science (Table 7, second column), the indices were .91 (CFI), .06 (SRMR), and .04 (RMSEA). Thus, the proposed models provided a good fit to the data, demonstrating that distal factors do

Table 6
z-Scores for significant path coefficients in SEMs for math achievement

Distal predicting opportunity	10th	12th
SES → general math course (1 year)	−5.837	−5.860
Parent expectations → general math course	−4.257	−3.454
Student expectations → general math course	−3.148	−3.552
Middle school GPA → general math course	−11.512	−9.693
SES → geometry course (1 year)	8.509	7.538
Parent expectations → geometry course	6.969	5.408
Student expectations → geometry course	5.917	5.588
Middle school GPA → geometry course	22.597	20.184
SES → algebra II course (1 year)	2.912	2.936
Parent expectations → algebra II course	2.731	2.629
Student expectations → algebra II course	3.198	2.794
Middle school GPA → algebra II course	19.114	17.129
Distal predicting Propensity		
SES → math achv. prior to 9th-grade	16.331	13.346
Parent expectations → math achv. prior to 9th-grade	7.453	7.151
Student expectations → math achv. prior to 9th-grade	4.856	3.887
Middle school GPA → math achv. prior to 9th-grade	29.143	26.637
SES → efficacy for graduating	4.328	2.489
Student expectations → efficacy for graduating	6.697	5.912
Middle school GPA → efficacy for graduating	12.566	9.971
SES → plans to take the SAT	8.012	5.972
Parent expectations → plans to take the SAT	7.340	5.876
Student expectations → plans to take the SAT	13.015	11.521
Middle school GPA → plans to take the SAT	13.157	12.090
Distal, opportunity, propensity predicting achievement		
SES → math achievement	6.071	7.614
Parent expectations → math achievement	n/s	3.131
Student expectations → math achievement	2.678	2.439
Middle school GPA → math achievement	6.668	6.415
General math course → math achievement	−8.633	−9.110
Geometry course → math achievement	9.504	10.578
Algebra II course → math achievement	3.999	2.795
Math emphasis → math achievement	4.447	2.177
Math achv. prior to 9th-grade → math achievement	78.405	58.067
Math GPA in 9th- and 10th-grade → math achievement	3.864	5.019
Efficacy for graduating → math achievement	4.917	3.642
Plans to take the SAT → math achievement	3.455	5.371
Math self-concept → math achievement	5.300	7.442

Note: Each z-score is derived by dividing the path coefficient for that variable by its standard error; the critical value for a two-tailed test at $p < .05$ is 1.96 for each z-score.

predict opportunity and propensity factors, which in turn predict achievement, and distal factors also have some direct effects on achievement. In addition, the coefficients listed in Tables 6 and 7 provide further insight into the relative predictive strength of individual factors because each coefficient includes adjustments for all other paths in the model. Moreover, it is notable that none of the four SEM models include gender or race/ethnicity as

Table 7
z-Scores for significant path coefficients for SEMs predicting science achievement

	10th	12th
Distal predicting opportunity		
SES → general science course (1 year)	−2.892	−2.609
Parent expectations → general science course	−2.284	n/s
Student expectations → general science course	−2.805	−4.454
Middle school GPA → general science course	−6.396	−6.204
SES → biology course (1 year)	6.168	4.823
Parent expectations → biology course	4.332	3.486
Middle school GPA → biology course	11.056	10.463
SES → chemistry course (1 year)	4.162	3.353
Student expectations → chemistry course	5.346	5.629
Middle school GPA → chemistry course	12.517	10.678
Distal predicting propensity		
SES → science achv. prior to 9th-grade	10.821	9.948
Parent expectations → science achv. prior to 9th-grade	4.523	4.405
Student expectations → science achv. prior to 9th-grade	5.168	4.853
Middle school GPA → science achv. prior to 9th-grade	17.098	15.609
SES → science GPA in 9th- and 10th-grade	3.716	3.698
Student expectations → science GPA in 9th- and 10th-grade	4.707	4.707
Middle school GPA → science GPA in 9th- and 10th-grade	34.162	31.466
SES → efficacy for graduating	3.799	3.646
Student expectations → efficacy for graduating	6.722	5.746
Middle school GPA → efficacy for graduating	12.769	10.540
SES → plans to take the SAT	8.740	6.522
Parent expectations → plans to take the SAT	7.127	5.833
Student expectations → plans to take the SAT	12.889	11.591
Middle school GPA → plans to take the SAT	12.215	11.570
Distal, opportunity, propensity predicting achievement		
SES → science achievement	13.575	9.550
Middle school GPA → science achievement	5.576	4.753
General science course → science achievement	−4.549	−3.660
Biology course → science achievement	2.781	2.390
Chemistry course → science achievement	2.945	2.352
Teacher responsiveness → science achievement	2.675	n/s
Science achv. prior to 9th-grade → science achievement	55.727	45.710
Science GPA in 9th- and 10th-grade → science achievement	8.682	7.955
Efficacy for graduating → science achievement	3.627	n/s
Plans to take the SAT → science achievement	2.554	3.032

Note: Each z-score is derived by dividing the path coefficient for that variable by its standard error; the critical value for a two-tailed test at $p < .05$ is 1.96 for each z-score.

significant predictors because analyses showed that including these factors produced substantial decreases in the CFI and substantial increases in the SRMR and RMSEA. Similarly, inclusion of single latent factors for the distal, opportunity, and propensity categories produced poor fits as well. Thus, the factors appeared to operate independently but conjointly.

4. Discussion

The primary goal of the present study was to promote the process of theory construction in the area of academic achievement by identifying the factors that individually and collectively account for large amounts of variance in math and science achievement during high school. To identify these factors, we used a combination of top-down and bottom-up approaches. The top-down (meta-theoretical) approach consisted of using an O–P framework to identify the categories or kinds of factors that might contribute to growth in math and science skills. The bottom-up (empirical) approach consisted of conducting analyses to see which of the potential candidates actually did account for variance in science and math achievement. The basic premise of the O–P framework is that high achievement is more likely if (a) children are given genuine opportunities to learn and (b) children have the propensity to take advantage of these learning opportunities. Other explanatory factors (e.g., SES and parent expectations for children) were argued to be operative as distal variables that not only predict the emergence of opportunities and propensities, but also continue to exert some influence on these outcomes as well. Finally, we assumed that gender and race/ethnicity would not explain much variance in achievement once factors from the distal, opportunity, and propensity clusters had been controlled. Our secondary analyses of the NELS database generally found considerable support for these assumptions and also revealed the relative importance of various predictors. In what follows, we consider the implications of these results by providing answers to four questions: (a) Which factors were most important? (b) Why did the results differ somewhat for math and science achievement? (c) What are the next steps in theory building? and (d) What implications do the results have for intervention efforts?

4.1. Which factors were most important?

From the standpoint of theory building, the results of the present study are very informative. In particular, the distal factors were found to have both direct effects on achievement and indirect effects through the opportunity and propensity factors. Of these, SES and middle school GPA were generally the strongest and most consistent predictors (see Tables 6 and 7). For math, more advantaged students who earned better grades in middle school were more likely to take one full year of geometry and one full year of algebra II by the end of 10th-grade (opportunity factors) and also brought a higher level of mathematical proficiency (propensity) to these courses. More proficient students who took these courses, in turn, demonstrated higher achievement in both the 10th- and 12th-grades. Even with coursework and mathematical proficiency controlled, however, SES and middle school GPA still yielded additional direct effects. They generated z-scores for paths predicting math achievement that, relative to the other significant variables, were moderately large (range from 6.071 to 7.614) (see bottom section of Table 6). Similar results occurred for science achievement. More advantaged students who earned better grades in middle school were more likely to take biology and chemistry, and also more likely to bring a higher level of proficiency in science to these courses. Students who had the propensity to take advantage of their learning opportunities in these courses subsequently showed the highest level of science achievement.

These findings for the structural equation models were corroborated by the hierarchical regressions as well. In the case of distal factors, which accounted for 28.8–43.0% of the variance, children who (a) came from higher SES homes, (b) had parents who held higher educational expectations for them, and (c) earned better grades in middle school, gained more math and science knowledge during high school than students with the opposite profile of these distal factors. In addition, students who had higher educational expectations themselves also performed better on the 10th- and 12th-grade math assessments.

The findings from the regression analyses also clarify the role of opportunities and propensities in achievement. With respect to opportunities, the results showed that students experienced more knowledge growth when they took a full year of geometry, algebra II, and chemistry than when they took general math and general science courses. But the overall findings from the regression analyses suggest that enrollment in particular courses is best viewed as a necessary but not sufficient condition for knowledge growth in math and science. As noted above, several distal variables were significant predictors of children's achievement even after controlling for the other variables in the model. Furthermore, the findings show that indicators of children's propensities to take advantage of learning opportunities in high school also accounted for considerable amounts of variance, even after controls (i.e., 21.9–27.6% of the variance added to that contributed by distal and opportunity factors). Children who entered 9th-grade with higher math and science achievement scores showed more knowledge growth in these respective areas by 10th- and 12th-grade than those with lower scores. This finding is clearly consistent with the recent results of Jones and Byrnes (2006) and with the conclusions of Corno et al.'s (2002) review in which they argue that cognitive propensities are nearly always the strongest predictors of achievement. Likewise, the present results show that students who earned better grades in their 9th- and 10th-grade math and science courses also showed more knowledge growth in these respective areas by the end of 10th-grade and 12th-grade. Given that achievement test scores and grades in courses reflect some combination of knowledge and aptitude for math and science, these findings support the idea that growth will not occur simply because children are enrolled in courses. Rather, children have to be able to benefit from their opportunities to learn. Moreover, the findings suggest that, in addition to being able to benefit from opportunities, children also need to be willing to take advantage of opportunities. Lending support to this assertion is the finding that even after controls for distal factors, opportunity factors, and cognitive propensities, motivational propensities such as math self-concept, efficacy for graduating high school, and plans to take the SAT also contributed additional, unique variance. These findings are consistent with those of Byrnes (2003) and Jones and Byrnes (2006).

As predicted, gender and race/ethnicity accounted for only a small amount of the variance in science and math achievement after distal, opportunity, and propensity factors had been controlled. Whereas gender and race/ethnicity accounted for 9.3–13.7% of the variance without controls, they accounted for only .2–3.1% of the variance with controls. This substantial reduction after controls is quite similar to that reported by Byrnes (2003) in a secondary analysis of 12th-grade math NAEP scores, and suggests that gender and race/ethnicity are largely confounded with distal, opportunity, and propensity factors. Note, for example, that the percentages of students in each racial/ethnic group who took a full year of geometry by the 10th-grade assessment were 56.6% (Asian), 48.4% (White), 34.4% (Black), 32.2% (Hispanic), and 24.2% (Native American). Also, the percentages falling into the top two SES quintiles were 53.2% (Asian), 46.3% (White), 22.0% (Native American), 20.9% (Black), and 17.9% (Hispanic). It is also notable that inclusion of gender and race/ethnicity in the structural equation models produced a substantial reduction in fit indices, suggesting again that other factors such as SES and the skills students brought to high school mattered more.

4.2. Why did the results differ somewhat for math and science achievement?

The total amounts of explained variance were 80.6% for 10th-grade math and 76.6% for 12th-grade math. For 10th- and 12th-grade science, the corresponding figures were about 20% points lower (i.e., 60.8% and 57.7%, respectively). This difference in total explained variance seems to be largely due to more explained variance for math than for science with respect to distal factors (on average, 42% versus 29%) and opportunity factors (11% versus 1%). The z-scores in the structural equation models for math and science show a similar differential pattern whereby some distal and opportunity variables were either significant predictors of achievement in math but not science (e.g., student expectations) or stronger predictors for math than science (e.g., middle school GPA, coursework) (see Tables 6 and 7). In addition, as the correlations in Tables 2–5 suggest, the math and science proficiency that students brought to high school (indexed by achievement scores before the start of 9th-grade) was a stronger predictor for 10th- and 12th-grade math than it was for 10th- and 12th-grade science. Whereas the partial correlations between early achievement and later achievement (with all other variable controlled) were $r\text{s} = .726$ and .638 for 10th-grade and 12th-grade math, respectively, the corresponding partials for 10th-and 12th-grade science were $r\text{s} = .580$ and .537, respectively.

The differential findings for opportunities and cognitive propensities for science and math may reflect the fact that topics covered in later courses in science (e.g., chemistry) may not build on topics covered in prior courses (e.g., biology) in the same way that topics covered in later courses in math (e.g., algebra II) build on topics covered in prior courses (e.g., algebra I). But other possibilities need to be explored in follow-up studies.

The demographics were the only set of factors where there was less explained variance for math than for science (less than 1% versus 2–3%). This finding seems to partly be the result of the demographics, particularly gender, correlating a little stronger with science than with math (though still modest, even for science). However, this finding may also reflect the fact that the models for math achievement contained a predictor that was not included for science: self-concept. There were significant gender and racial/ethnic differences in math self-concept, so controlling for this variable would tend to reduce the unique variance in math achievement that was explained by gender and race/ethnicity. As previously noted, a comparable science self-concept factor was not controlled in the case of science achievement because the NELS database did not contain such a factor. Another issue to consider is that the other factors common to both models were generally stronger predictors for math than for science, so this left less variance to explain for math upon entering the demographics. Understanding these differences could yield important clarifications and refinements of an analysis of achievement.

4.3. What are the next steps in theory building?

The analyses reported in this paper (and other preliminary analyses not reported here) were useful in identifying a comprehensive set of authentic (i.e., non-spurious) predictors of math and science achievement. As noted above, predictors in the categories of distal, opportunity, and propensity factors collectively explained between 55% and 80% of the variance in achievement. This high amount of explained variance, which is substantially more than that typically found in longitudinal studies of achievement, suggests that we have identified many of the factors that seem to be responsible for different patterns of knowledge growth during high school. However, the fact that we have not accounted for 100% of the variance implies that there is more work to be done. The primary ways to increase explained variance are to (a) identify and include other potent factors that we did not examine in this study and (b) increase the precision of measurement for all of the factors that we identified, if possible.

Regarding the first of these possibilities, it would be useful to conduct studies that examine the role of all of the significant factors that we identified, plus three others that have been identified in previous studies of achievement but do not have adequate indices in the NELS database: self-regulation (Miller & Byrnes, 2001; Pintrich, 2000), domain-specific interest (Alexander, Sperl, Buehl, Fives, & Chiu, 2004; Byrnes, 2003; Eccles et al., 1998), and self-efficacy (Berry & West, 1993; Zimmerman & Kitsantas, 2005). Jones and Byrnes (2006) found that self-regulation was nearly as potent of a predictor of knowledge growth as prior knowledge. Relatedly, motivation researchers routinely find that interest accounts for 10% of the unique variance in achievement, while self-efficacy accounts for an additional 13–15%. Whereas intelligence has been found to explain as much as 10% unique variance in achievement after controls, Byrnes and Jones (2006) found it was not a significant predictor when prior knowledge and motivation were controlled.

But in addition to these factors that have received a fair amount of attention, there may be others that have not been considered to the same extent. We propose that the O–P framework can be very useful for identifying these additional factors. As noted earlier, opportunity factors pertain to exposure to content (e.g., coursework, homework, emphasis, etc.) and teacher quality (e.g., communication skills, equitable treatment of students, classroom management, etc.). Some of these possibilities were not examined here and could increase the amount of variance explained in future studies. Propensity factors, in contrast, are any factors that relate to the ability or willingness to learn content once it has been exposed or presented. In addition to the propensity factors examined here or studied extensively by others, there are other factors that conceivably could affect the extent to which a high school student pays attention or processes information, such as level of arousal (e.g., low arousal due to lack of sleep or excessive arousal due to stress), emotional states (e.g., distress or depression due to lack of parental attention or problems at home or problems in relationships), and cognitive level (e.g., ability to engage in abstract or multidimensional thinking). These possibilities should be considered in future studies as well.

The second way to increase the amount of explained variance is to develop more precise or sensitive methods of assessment for all of the significant predictors in Tables 2–5 as well as those omitted from the tables such as self-regulation and interest. A basic tenet of measurement theory is that the observed score for any variable is a combination of the "true score" for that variable plus measurement error (Cohen & Swerdlik, 2001; Cronbach, 1995). Measurement error decreases the size of correlations in studies of the test–retest reliability of the same measure, and also decreases the size of correlations among several different measures. For example, if the test-retest reliability correlation for a measure is $r = .90$, 19% of the variance is due to measurement error (i.e., $1 - (.90^2) = .19$). Measurement error can also be approximated by computing the standard error of measurement (which can be used to construct confidence intervals). The goal is to increase precision until measurement error is minimized and correlations are maximized. Jones and Byrnes (2006), for example, found that the correlation between self-regulation and math achievement could be increased from $r = .25$ (the typical correlation found in the literature) to $r = .45$ by switching from using a self-report questionnaire to using teacher ratings of student self-regulation. We suspect that altering the manner of assessment of some of the variables in Tables 2–5 would increase correlations, involving these variables in the same way. Furthermore, had we made strong claims that a factor was not an important predictor because its correlation was not significant, such claims could be open to refutation because increased precision could increase a non-significant correlation to the point that it becomes significant (e.g., $r = .15$–$.25$).

Other ways to improve upon our efforts would be to "unpack" or refine the measurement of certain factors. As noted earlier, for example, opportunity to learn should not be limited to enrollment in classes (Tate, 1995; Taylor et al., 2000). In future studies, it would be important to include additional measures of teacher quality (e.g., equal treatment of students, classroom management skills). In addition, there are multiple cognitive theories that can explain why prior knowledge (cognitive propensity) promotes knowledge growth, but no detailed accounts (to our knowledge) that explain the intervening causal mechanisms that link SES and achievement. What exactly happens in higher SES homes to promote achievement that does not happen in lower SES homes? Do highly educated parents teach their children content and therefore fill in gaps left by inadequate instruction? Do they provide extracurricular experiences that promote knowledge growth? Revealing such differences will go a long way to help explain the amount of variance related to SES. If carried out properly, such studies should demonstrate that all effects of SES are indirect (mediated through other variables). A similar approach could be used to "unpack" the remaining 1–3% of the variance accounted for by gender and race/ethnicity. What other factors could these variables be indexing (e.g., student interest, teacher bias?). Although there are several methodologies that would be useful in this regard, it would seem that conducting detailed, exploratory qualitative studies of high achieving and lower achieving students would be especially useful and enlightening.

4.4. *What implications do the results have for intervention efforts?*

As argued earlier, there is a direct relationship between accurate theories of knowledge growth and effective forms of intervention. Accurate theories describe the complete set of factors that are responsible for knowledge growth, and effective forms of interventions target these factors. At present, theory building in the area of achievement is still a work in progress, so it would be premature to draw definitive conclusions about the implications of the present study for intervention efforts. However, it is worth noting that should additional studies confirm the fact that a specific combination of distal, opportunity, and propensity factors promotes achievement, these findings would have important implications for policy. With few exceptions, most contemporary programs designed to enhance achievement in students (or reduce achievement gaps among demographic subgroups) focus exclusively on opportunity factors. For example, they (a) alter the curriculum in some way (e.g., adopt strategies such as cooperative learning or implement NCTM standards), (b) implement professional development for teachers to improve the quality of their instruction, or (c) increase student access to educational opportunities (e.g., busing programs, voucher programs, magnet schools, etc.). The present findings suggest that these measures will only have an effect on students who are able and willing to take advantage of these new opportunities. To increase the effectiveness of opportunity-focused interventions that are implemented in a particular grade (e.g., 9th-grade), it would be very important to also implement distal-focused and propensity-focused interventions in earlier grades (e.g., middle school). If the latter interventions are effective, students would then bring the skills and motivation they need to fully benefit from enhanced learning opportunities in high school.

References

Alexander, P. A., & Winne, P. (2006). *Handbook of educational psychology* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Alexander, P. A., Sperl, C. T., Buehl, M. M., Fives, H., & Chiu, S. (2004). Modeling domain learning: Profiles from the field of special education. *Journal of Educational Psychology, 96*, 545–557.

Berry, J. M., & West, R. L. (1993). Cognitive self-efficacy in relation to personal mastery and goal setting across the life span. *International Journal of Behavioral Development, 16*, 351–379.

Byrnes, J. P. (2003). Factors predictive of ethnic differences in mathematics proficiency in White, Black, and Hispanic 12th graders. *Journal of Educational Psychology, 95*, 316–326.

Cohen, J., Albright, L., Cohen, D., Hollender, D., Jiang, T., & Chan, T. (2003). *AM statistical software*. Washington, DC: Jon Cohen and the American Institutes for Research.

Cohen, R. J., & Swerdlik, M. (2001). *Psychological testing and assessment: An introduction to tests and measurement*. New York: McGraw-Hill.

Corno, L., Cronbach, L. J., Kupermintz, H., Lohman, D. F., Mandinach, E. B., Porteus, A. W., et al. (2002). *Remaking the concept of aptitude: Extending the legacy of Richard E. Snow*. Mahwah, NJ: Lawrence Erlbaum.

Cronbach, L. J. (1995). *A valedictory: Reflections on 60 years in educational testing*. Washington, DC: National Academies Press.

Curtin, T. R., Ingels, S. J., Wu, S., & Heuer, R. (2002). *NELS:88 base-year to fourth follow-up data file user's manual* (NCES publication No. 2002–323). Washington, DC: US Department of Education, National Center for Education Statistics.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon (series Ed.), N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 1018–1095).

Hoff, E. (2001). *Language development* (2nd ed.). Belmont, CA: Wadsworth/Thomson Learning.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Ingels, S. J., Dowd, K. L., Baldridge, J. D., Stipe, J. L., Bartot, V. H., & Frankel, M. R. (1994). *National education longitudinal study of 1988 second follow-up: Student component data file user's manual* (NCES publication No. 94–374). Washington, DC: US Department of Education, National Center for Education Statistics.

Jacobs, J. K., Hiebert, J., Giwin, K. B., Hollingsworth, H., Garnier, H., & Wearne, D. (2006). Does eighth grade mathematics teaching in the United States align with the NCTM standards? Results from the TIMSS 1995 and 1999 video studies. *Journal of Research in Mathematics Education, 37*, 5–32.

Johnson, E. G. (1989). Considerations and techniques for the analysis of NAEP data. *Journal of Educational Statistics, 14*, 303–334.

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

Kuhn, D. (1992). Piaget's child as scientist. In H. Beilin & P. Pufall (Eds.), *Piaget's theory: Prospects and possibilities*. Mahwah, NJ: Lawrence Erlbaum.

Leder, G. C. (1992). Mathematics and gender: Changing perspectives. In D. A. Grouws (Ed.), *Handbook of research on mathematical teaching and learning* (pp. 597–624). New York: Macmillan.

Linn, R. L., & Dunbar, S. B. (1992). Issues in the design and reporting of the National Assessment of Educational Progress. *Journal of Educational Measurement, 29*, 177–194.

Marsh, H. W., & O'Neill, R. (1984). Self-description questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21*, 153–174.

Miller, D. C., & Byrnes, J. P. (2001). To achieve or not to achieve: A self-regulation perspective on adolescents' academic decision making. *Journal of Educational Psychology, 93*, 677–685.

Muthén, L. K., & Muthén, B. O. (2006). *Mplus statistical analysis with latent variables: User's guide*. Los Angeles, CA: Muthén & Muthén (www.statmodel.com).

Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.

Opdenakker, M., Van Damme, J., De Fraine, B., Van Landeghem, G., & Onghena, P. (2002). The effects of schools and classes on mathematics achievement. *School Effectiveness and School Improvement, 13*, 399–427.

Owings, J. A. (1995). *Psychometric report for the NELS:88 base year through second follow-up* (NCES publication No. 95–382). Washington, DC: US Department of Education, National Center for Education Statistics.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekarts & P. R. Pintrich (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic Press.

Pressley, M., Wharton-McDonald, R., & Rafael, L. M. (2002). Exemplary first-grade teaching. In B. Taylor & P. D. Pearson (Eds.), *Teaching reading: Effective schools accomplished teachers* (pp. 73–88). Mahwah, NJ: Lawrence Earlbaum.

Reichenbach, H. (1938). *Experience and prediction*. Chicago: University of Chicago Press.

Roscigno, V. J. (2000). Family/school inequality and African-American/Hispanic achievement. *Social Problems, 47*, 266–290.

Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly, 47*, 1–41.

Tate, W. F. (1995). School mathematics and African American students: Thinking seriously about opportunity-to-learn standards. *Educational Administration Quarterly, 31*, 424–448.

Taylor, B., Pressley, M., & Pearson, D. (2000). *Effective teachers and schools: Trends across recent studies*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement.

Zimmerman, B. J., & Kitsantas, A. (2005). The hidden dimension of personal competence: Self-regulated learning and practice. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 509–526). New York, NY: Guilford.