American Educational Research Journal
December 2019, Vol. 56, No. 6, pp. 2077–2115
DOI: 10.3102/0002831219843292
Article reuse guidelines: sagepub.com/journals-permissions

Why High School Grades Are Better Predictors of On-Time College Graduation Than Are Admissions Test Scores: The Roles of Self-Regulation and Cognitive Ability

Brian M. Galla
University of Pittsburgh
Elizabeth P. Shulman
Brock University
Benjamin D. Plummer
University of Michigan
Margo Gardner
Columbia University
Stephen J. Hutt
University of Colorado Boulder
J. Parker Goyer
Stanford University
Sidney K. D'Mello
University of Colorado Boulder
Amy S. Finn
University of Toronto
Angela L. Duckworth
University of Pennsylvania

Compared with admissions test scores, why are high school grades better at predicting college graduation? We argue that success in college requires not only cognitive ability but also self-regulatory competencies that are better indexed by high school grades. In a national sample of 47,303 students who applied to college for the 2009/2010 academic year, Study 1 affirmed that high school grades out-predicted test scores for 4-year college graduation. In a convenience sample of 1,622 high school seniors in the Class of 2013, Study 2 revealed that the incremental predictive validity of high school grades for college graduation was explained by composite measures of self-regulation, whereas the incremental predictive validity of test scores was explained by composite measures of cognitive ability.

KEYWORDS: cognitive ability, college admissions tests, college graduation, report card grades, self-regulation

Since World War II, high school report card grades and standardized college entrance exams—most prominently the SAT and ACT—have been the two gatekeepers of university admissions (National Association for College Admission Counseling, 2008). Longitudinal studies of students who graduated

---

BRIAN M. GALLA is an assistant professor of applied development psychology and research scientist in the Learning Research and Development Center at the University of Pittsburgh, 5940 Wesley W. Posvar Hall, 230 South Bouquet Street, Pittsburgh, PA 15260; e-mail: *gallabri@pitt.edu*. His research examines how mindfulness and self-control contribute to positive development from adolescence to young adulthood.

ELIZABETH P. SHULMAN is an assistant professor of psychology and director of the Youth Development Lab at Brock University. Her work examines psychosocial development during the transition from childhood to adulthood and how it relates to decision making, risk taking, and antisocial behavior.

BENJAMIN D. PLUMMER is a learning experience designer at the University of Michigan's Ross School of Business. His work focuses on applying self-determination theory and principles of gameful design to create learning environments that support student motivation and engagement.

MARGO GARDNER is a senior research scientist at the National Center for Children and Families, Teachers College, Columbia University. Her research examines contextual influences on academic success and well-being during childhood, adolescence, and young adulthood.

STEPHEN J. HUTT is a graduate student in the Department of Computer Science and Institute of Cognitive Science at the University of Colorado Boulder. His research interests include how Big Data and artificial intelligence techniques may be leveraged to improve education.

J. PARKER GOYER is a postdoctoral fellow in the Department of Psychology at Stanford University. Her research examines how social contexts moderate the role of psychological factors on educational outcomes.

SIDNEY K. D'MELLO is an associate professor in the Institute of Cognitive Science and Department of Computer Science at the University of Colorado Boulder. He is interested in the dynamic interplay between cognition and emotion while individuals and groups engage in complex real-world tasks.

AMY S. FINN is an assistant professor of psychology and director of the Learning and Neural Development Lab at the University of Toronto. Her work examines how cognitive and brain development shape learning outcomes and how the environment affects these relationships.

ANGELA L. DUCKWORTH is the founder and CEO of Character Lab, a nonprofit whose mission is to advance the science and practice of character development. She is also the Christopher H. Browne Distinguished Professor of Psychology at the University of Pennsylvania, faculty codirector of the Penn-Wharton Behavior Change for Good Initiative, and faculty codirector of Wharton People Analytics. She studies self-control, grit, and other determinants of achievement.

from U.S. high schools in the late 20th century have shown that both metrics provide unique predictive power for success in college. It is surprising, perhaps, that high school grades are a stronger incremental predictor of college outcomes—including what is arguably the most important outcome of all: graduation. What abilities do report card grades reflect that admissions tests do not? And what do these relationships reveal about what is required to succeed in college? In the current investigation, we test the hypothesis that report card grades differentially reflect self-regulation, whereas standardized admissions test scores reflect cognitive ability—and that the predictive validity of grades can be explained by the relatively greater importance of self-regulation in completing a college degree.

The Surprising Prognostic Power of High School Grades

In the early 20th century, high school teachers tasked with assessing ever greater numbers of students in more diverse and specific subject areas began to replace oral and written narrative descriptions of student progress with percentage grades (Brookhart et al., 2016). Skepticism about the validity of teacher-assigned grades dates to their inception: "Over the past 100 years, teacher-assigned grades have been maligned by researchers and psychometricians alike as subjective and unreliable measures of student academic achievement" (Brookhart et al., 2016, p. 833). More recently, trends toward grade inflation have cast doubt on the relevance of the high school transcript to college admissions (Marcus, 2017).

Although the standards teachers use to assign grades are subjective and heterogeneous (Camara, Kimmel, Scheuneman, & Sawtell, 2003; Cross & Frary, 1999), the predictive validity of report card grades for later academic success is among the sturdiest findings in social science (Brookhart et al., 2016; Thorsen & Cliffordson, 2012). Numerous studies show that grades are associated with all measures of academic achievement, including standardized achievement test scores (Duckworth, Weir, Tsukayama, & Kwok, 2012); on-time grade promotion and high school graduation (Jackson, 2018); enrollment, performance, and persistence in college (Pattison, Grodsky, & Muller, 2013; Robbins et al., 2004; Westrick, Le, Robbins, Radunzel, & Schmidt, 2015); lifetime educational attainment (French, Homer, Popovici, & Robins, 2015); and both prestige of occupation and annual income in adulthood (French et al., 2015; Pattison et al., 2013).

Of special relevance to the current investigation, high school report card grades outperform admissions test scores in predicting college graduation. For example, a study of nearly 80,000 first-time students entering the University of California between 1996 and 1999 showed that high school grades were superior to SAT test scores in predicting 4-year graduation as well as cumulative GPA (Geiser & Santelices, 2007). These findings were replicated in a landmark study of nearly 150,000 first-time college students of the 1999 entering cohort in flagship universities across the country and four statewide systems (Bowen, Chingos, & McPherson, 2009). Most recently, the superior predictive validity of report card grades relative to ACT scores for college grades and retention was demonstrated in a meta-analysis of over 180,000 students at 50 colleges and universities (Westrick et al., 2015).

Why Report Card Grades Depend on Self-Regulation

Over two millennia ago, Aristotle purportedly observed that "the roots of education are bitter, but the fruit is sweet" (Laertius, 1925, p. 461). At the dawn of the 20th century, James (1899) speculated that in all schoolwork there is "a large mass of material that must be dull and unexciting" (pp. 104–105). More recently, experience-sampling data from over 1,000 middle and high school students reveal that studying, completing homework, and other aspects of schoolwork are less enjoyable than playing sports, watching television, socializing with friends, and resting (Duckworth, Taxer, Eskreis-Winkler, Galla, & Gross, 2019). Notably, this pattern is similar across gender, grade level, and even current level of academic achievement (see also, Wong & Csikszentmihalyi, 1991).

What is self-regulation? Self-regulation can be defined as a set of goal-directed motivational and volitional competencies (Carver & Scheier, 1982; Davisson & Hoyle, 2017). Self-regulatory competencies include self-control: the ability to act, think, and feel in ways that are more valuable in the long-run than momentarily more alluring alternatives (Hofmann, Friese, & Strack, 2009; Mischel, Shoda, & Rodriguez, 1989). Over time periods extending to months or years, successful self-regulation also requires grit, the tendency to sustain passion and perseverance toward challenging goals (Duckworth & Gross, 2014; Eskreis-Winkler, Gross, & Duckworth, 2017). Self-control and grit are highly correlated facets of Big Five conscientiousness, a family of individual differences encompassing not only self-regulation but also traditionalism, orderliness, and others (Roberts, Chernyshenko, Stark, & Goldberg, 2005). Self-regulation also overlaps with the construct of self-regulated learning—a set of motivational, volitional, and cognitive competencies and strategies that have been shown to help students become independent learners (Zimmerman, 1990).

Prior research has identified self-regulatory competence as central to academic success at all levels of schooling (Duckworth & Carlson, 2013; Poropat, 2009). For example, in cross-sectional studies of large, nationally representative samples, high school students who earn higher grades report spending more time doing homework and less time "killing time" (e.g., playing videogames or watching television) and more reliably coming to class with their books, school supplies, and completed homework (Willingham, Pollack, & Lewis, 2002). Self-control has also been shown to prospectively predict improvements in report card grades (Duckworth et al., 2012;

Duckworth, Tsukayama, & May, 2010; Tsukayama, Duckworth, & Kim, 2013). A smaller literature has shown grit is associated with higher academic grades (Duckworth, Peterson, Matthews, & Kelly, 2007; Hoyle, Davisson, Daniel, & Komoski, 2018; Muenks, Wigfield, Yang, & O'Neal, 2017).

Why Admissions Test Scores Depend on Cognitive Ability

The Scholastic Aptitude Test (SAT)[1] debuted in 1926 with the explicit aim of leveling the playing field for students aspiring to earn a college diploma (Lemann, 1999). Advocates championed the SAT—adapted from tests of general cognitive ability developed for the U.S. Army—as an unbiased assessment of intelligence and therefore the ideal predictor for meeting the intellectual demands of a college education (Frey & Detterman, 2004). Unlike teacher-assigned grades, standardized tests promised to eliminate subjective judgment from the evaluation of academic ability. The ACT was introduced as a competitor to the SAT in 1959[2] and has since grown in popularity; generally, scores on the SAT and ACT are highly correlated (rs >.8, Coyle & Pillow, 2008).

The relevance of cognitive ability to academic performance is straightforward. Cognitive ability "involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (Gottfredson, 1997, p. 13). Scores on diverse tests of cognitive ability are highly intercorrelated (Jensen, 1998) and, like self-regulation, cognitive ability (also referred to as general intelligence, general mental ability, or simply "*g*") is a higher order construct explaining much of the variance of its components including working memory, processing speed, fluid reasoning, verbal ability, and so on (Carroll, 1992).

Across grade levels, measures of cognitive ability predict academic performance (Deary, Strand, Smith, & Fernandes, 2007; Kuncel & Hezlett, 2007; Kuncel, Ones, & Sackett, 2010; Neisser et al., 1996; Poropat, 2009). Cognitive ability has demonstrated especially high associations with admissions test scores: After correcting for nonlinearity and restriction on range, Frey and Detterman (2004) found that measures of cognitive ability were correlated at .87 (Study 1) and .72 (Study 2) with SAT scores, concluding that the SAT is "mainly a test of *g*" (p. 373). Likewise, Koenig, Frey, and Detterman (2008) estimated corrected correlations of .75 and .77 between ACT scores and measures of cognitive ability, concluding that the "ACT is a measure of general intelligence" (p. 158).

Where Grades and Standardized Admissions Tests Differ

Report card grades and standardized admissions test scores both measure students' academic skills and knowledge. Relations among these metrics are typically strong, with adjusted correlations in the .5 range (Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008; Robbins et al., 2004). Notably, these associations do not approach unity, even when controlling for measurement error, subject matter, and variation across teachers and schools in grading standards (Willingham et al., 2002).

Grades are based largely on students' mastery of the academic content taught in the classroom (Willingham et al., 2002). But when assigning grades, teachers also incorporate a "hodgepodge" of nonachievement factors including attitudes, behavior, and effort (Brookhart, 1991, 1994; Parsons, 1959). For instance, in a survey of 155 high school teachers, 73% said homework had a moderate or strong influence on grades, and 71% said conduct should be taken into consideration (Cross & Frary, 1999).

Admissions test scores, on the other hand, assess general reasoning abilities and knowledge of specific academic topics, including mathematics, reading, and writing. Willingham (2005) observed that

> high-stakes tests represent an external standard; that is, how well teachers and students have done on standards that apply across an educational system. Even under the best of circumstances, teacher standards and system standards are unlikely to be precisely the same. (p. 130)

Thus, in contrast to quizzes and tests created by classroom teachers, admissions tests include at least some problems that have not been introduced as part of formal instruction (Willingham et al., 2002). Likewise, admissions tests do not directly assess factors like classroom conduct or homework completion.

Compared to standardized tests, what aspects of student competence are differentially reflected in teacher-assigned grades? In a study of 8,454 high school students in the National Educational Longitudinal Study of 1988, Willingham et al. (2002) found that the disparity between report card grades and standardized test scores was substantially reduced when controlling for self-reported homework completion, class participation, discipline problems, and a variety of other indicators that Willingham (2005) later described as "*conative skills* like volition, habits of inquiry, and self-regulation" (p. 133). In a study of 264, 7- to 12-year-old students, Valiente, Lemery-Chalfant, Swanson, and Reiser (2008) found that self-control predicted higher report card grades via class participation and social competence. More recently, Duckworth et al. (2012) found in a convenience sample of 510 middle school students that teacher-rated and self-reported self-control predicted increases in report card grades over time, and these gains were mediated by teacher ratings of homework completion and classroom conduct.

Compared with report card grades, what aspects of student competence are differentially reflected in standardized admissions tests? Duckworth et al. (2012) hypothesized that "intelligence helps students learn and solve problems independent of formal instruction" (p. 439) and found evidence that, indeed, measured cognitive ability prospectively predicted increases in standardized achievement test scores in a national longitudinal study of 1,364 children followed by the National Institute of Child Health and Human Development. In a separate sample, teachers judged the mastery of skills and knowledge *not* taught in class and the mastery of skills and knowledge they had directly instructed to be equally relevant to performance on standardized tests (Study 3, Duckworth et al., 2012). By contrast, the mastery of skills and knowledge *not* taught in class were rated by the same teachers as the least relevant factor in determining report card grades.

Current Investigation

The benefits of a college education for later economic and professional success are clear (Faust, 2014), yet many students in the United States do not complete their degree on time. In 2009, for example, 39.9% of students who embarked on a 4-year college program graduated on time (National Center for Education Statistics, 2017b). Although an additional 19.5% of entering students in that cohort earned their degree within 6 years of enrolment, graduating late encumbers students with additional costs, including tuition, board, lost wages, and student debt (Complete College America, 2014).

What does it take to graduate from college? There are many reasons why students may fail to earn a college degree, including, importantly, contextual factors, such as socioeconomic barriers (e.g., Goldrick-Rab, 2006; Goldrick-Rab & Pfeffer, 2009), which are not the focus of the current investigation. We contend, however, that for students of all backgrounds, self-regulation is vital to success in college for all the same reasons self-regulation supports academic achievement at earlier stages of education. In fact, demands on self-regulation are greater in college, as attendance, homework, and studying are less closely monitored by teachers and parents. Not surprisingly, therefore, in a study of an array of self-reported psychosocial factors in a sample of 14,464 students at 48 different colleges, Robbins, Allen, Casillas, Peterson, and Le (2006) found that academic discipline, defined as "the amount of effort a student puts into schoolwork and the degree to which he or she sees himself as hardworking and conscientious," to be the strongest and most consistent predictor of first-year GPA (p. 600), followed by general determination, defined as "extent to which a student strives to follow through on commitments and obligations" (p. 600).

In the current investigation, we compare the predictive validities of high school grades and admissions test scores and what their relative contributions reveal about the nature of completing an undergraduate degree on time. We hypothesize that self-regulation explains the incremental predictive validity of high school grades, whereas cognitive ability explains the incremental predictive validity of admission test scores. In Study 1, we examine the predictive validities of high school grades and college admissions test scores for college graduation in a national sample of 47,303 students who applied to college via the Common Application for the 2009/2010 academic year—a decade later than students studied by Bowen et al. (2009) and Geiser and Santelices (2007). In Study 2, we examine the extent to which these predictive relationships in a demographically diverse convenience sample of 1,622 high school students in the Class of 2013 can be explained by measures of self-regulation and cognitive ability administered during the final year of high school. Collectively, these two prospective longitudinal studies provide a contemporary and novel empirical test of speculative assumptions that "high school grades measure a student's ability to 'get it done' in a more powerful way than do SAT scores" (Bowen et al., 2009, p. 123), whereas standardized college entrance exams primarily assess "cognitive skills that are vital to educational objectives" (Willingham, 2005, p. 133).

Study 1

In Study 1, we sought to replicate prior findings comparing the predictive validities of high school grades and admissions test scores for college graduation (Bowen et al., 2009; Geiser & Santelices, 2007). Replication was not a foregone conclusion for two reasons. First, high school grades have risen steadily over the past decade, with increased bunching in the A range (Hurwitz & Welch, 2018). Second, both the SAT and ACT college entrance exams have undergone substantial structural changes (e.g., both exams added a writing section in 2005) during the same period.

To create a national longitudinal dataset that included reliable and objective measures of predictors and college graduation, we worked with the Common Application (Common App, www.commonapp.org) and the National Student Clearinghouse (NSC, www.studentclearinghouse.org). To protect privacy, Common App contracted a third-party entity to collect, anonymize, and deliver the dataset to our team. For additional details about this data set, please refer to Hutt, Gardner, Kamentz, Duckworth, and D'Mello (2018).

Method

Participants

The sample for Study 1 was drawn from the population of students who completed the Common App during the 2008/2009 academic year for college admission during the 2009/2010 academic year (N = 413,675). From this population, we selected students who had not enrolled in a postsecondary institution prior to 2008 (N = 311,308). This ensured the accuracy of records reflecting time to degree attainment. Our final analytic sample included N = 47,303 students for whom data on high school grade point average (HSGPA) were available electronically.[3]

Slightly more than half of the students in the analytic sample were female (56.1%). Just under half identified as Caucasian (49.1%), while smaller percentages identified as Asian (10.5%), Latino (9.0%), African American (6.1%), or members of multiple or other race/ethnic groups (9.0%). Race/ethnicity was not reported for 16.4% of students. Slightly less than half the sample (46.7%) reported having two parents with a college degree, 24.2% reported having one parent with a college degree, and 29.1% reported that neither parent had a college degree. Nearly half of students attended non-Title I public high schools (44.4%), 29.7% attended private high schools, 19.7% attended Title I public high schools, and 0.6% were homeschooled. Data on type of high school attended were missing for 5.7% of students. As reported in the Supplemental Material available for this article in the online version of the journal, the analytic sample of 47,303 students was largely representative of the full sample of 311,308 students.

Procedure

We obtained students' cumulative high school grade point average, as reported by guidance counselors, on the secondary school report section of the Common App. SAT and ACT scores and demographic characteristics (described below) were obtained from the student section of the Common App. We obtained matched college enrollment and degree attainment data (degree title, date awarded, and college awarding the degree) from the 2015 NSC database.

Measures

High school grade point average (HSGPA). Because guidance counselors reported grades on a variety of different scales, we rescaled them to a common metric by computing a proportion score. This involved dividing each student's cumulative HSGPA from the maximum value of their high school's grading scale (also reported by the guidance counselor). The proportion scores were then screened for outliers, as these may have indicated counselors' entry errors. We removed values below the bottom 1% and above the top 1% of proportion scores. This yielded a normally distributed measure of HSGPA ranging from 0.50 to 1.23 (M = 0.92, SD = 0.12). A proportion score of 0.50 is equivalent to earning a 2.00 on a 4.00 scale, and a proportion score of 1.23 is roughly equivalent to earning a 5.00 on a 4.00 scale (we were not given information about whether the reported HSGPA was unweighted or weighted, so we did not make further adjustments to these scores).

College admissions test scores. Just under half (49.2%) of the analytic sample took only the SAT, 16.1% took only the ACT, and 30.3% took both the SAT and ACT. We computed a total SAT score by summing students' highest reported scores on the mathematics (*M* = 609.4, *SD* = 102.5), writing (*M* = 596.6, *SD* = 99.2), and critical reading (*M* = 599.0, *SD* = 101.1) subtests (each scored from 200 to 800). We computed a composite ACT score by taking the average of students' highest reported scores on the mathematics (*M* = 27.0, *SD* = 5.0), English (*M* = 27.7, *SD* = 5.3), reading (*M* = 28.2, *SD* = 5.4), and science (*M* = 26.1, *SD* = 4.7) subtests (each scored from 1 to 36). The total SAT (*M* = 1804.9, *SD* = 273.0, range = 620 to 2400) and composite ACT scores (*M* = 27.2, *SD* = 4.4, range = 1 to 36) were highly correlated ($r = .90$, $p < .001$). To compute a single admission test score variable, we first converted students' ACT score to an SAT score using published guidelines (ACT, 2013). For students reporting both SAT and ACT scores, we used the higher of the two scores. This resulted in a final score on the SAT with *M* = 1828.6 and *SD* = 273.6, range = 620 to 2400.

Demographic characteristics. From the Common App student section, we obtained students' self-reported gender, race/ethnicity, parent education, and type of high school attended. College attainment is correlated with a number of demographic characteristics and indicators of socioeconomic status (Attewell, Heil, & Reisel, 2011) that are also correlated with high school grades and admissions test scores (Robbins et al., 2004; Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). Thus, all models controlled for gender, race/ethnicity (five dummy codes for African American, Latino, Asian/Pacific Islander, multiple or other races/ethnicities, and no race/ethnicity reported versus Caucasian), parent education (two dummy codes for one or two vs. zero parents with a college degree), and type of high school (two dummy codes for Title I public high school and private school versus non–Title I public high school [and homeschool]). See the Supplemental Material available for this article in the online version of the journal for information about demographic differences on study variables.

On-time college graduation. Students who earned a bachelor's degree within 4 years of applying to college through the Common App were classified as 1 (*graduated within 4 years*), and all other participants were classified as 0 (*did not graduate within 4 years*).

Analytic Plan

We used structural equation modeling in *Mplus* Version 7.2 (Muthén & Muthén, 2012). Approximately 2.0% of the data were treated as missing for HSGPA, and 4.4% of the data for SAT/ACT score and 5.7% of the data for type of high school attended were not available. Missing data were handled using full-information maximum likelihood, which produces less biased and more efficient estimates than other methods, such as listwise deletion (Baraldi & Enders, 2010; Schafer & Graham, 2002). Thus, the sample

Table 1
Results of Structural Model Predicting On-Time College Graduation in Study 1 (*N* = 47,303)

Variables	β	*OR*	LCI	UCI	*p*
HSGPA	0.13	1.28	1.250	1.310	<.001
Admissions test scores	0.06	1.12	1.093	1.149	<.001
Demographic characteristics					
African American	−0.04	0.74	0.681	0.812	<.001
Asian	−0.03	0.84	0.786	0.894	<.001
Latino	−0.04	0.78	0.722	0.837	<.001
Other ethnicities	−0.03	0.84	0.786	0.903	<.001
No race reported	−0.01	0.95	0.898	0.998	.042
Female	0.06	1.24	1.193	1.289	<.001
One parent with college degree	0.04	1.21	1.142	1.273	<.001
Two parents with college degree	0.06	1.26	1.195	1.321	<.001
Title I public school	−0.01	0.94	0.888	0.987	.014
Private school	0.01	1.06	1.008	1.104	.021

Note. HSGPA = senior year high school grade point average; *OR* = odds ratio; LCI/UCI = lower/upper 95% confidence interval of the odds ratio; β = standardized regression coefficient.

size for the main model was 47,303. Both HSGPA and SAT/ACT score were *z*-score standardized during analysis to enable path comparisons.

We regressed college graduation (a binary variable) on HSGPA and SAT/ACT score. We also included demographic characteristics (gender, race/ethnicity, parent education, type of high school attended) as predictors of college graduation, HSGPA, and SAT/ACT score. HSGPA and SAT/ACT score were allowed to covary. We used the Wald test (Cox & Hinkley, 1974) to compare the strength of the paths leading from HSGPA and SAT/ACT score to college graduation. The Wald test assesses how much worse the model fit would be if the compared paths were constrained to be equal, with a significant ($p < .05$) test indicating unequal path strengths.[4]

Results and Discussion

Consistent with 2009 national averages (National Center for Education Statistics, 2017b), 39.4% of students in our sample earned a bachelor's degree within four years of applying to college through the Common App. Also consistent with prior research (Robbins et al., 2004), HSGPA and admissions test scores were positively correlated with each other ($r = .50$, $p < .001$).

Did HSGPA predict college graduation more powerfully than admissions test scores? It did. Adjusting for demographic characteristics, the incremental predictive validity of HSGPA for college graduation ($OR = 1.28$, $p < .001$, 95% CI [1.25, 1.31]) was stronger than the incremental predictive validity of SAT/

ACT score (OR = 1.12, $p < .001$, 95% CI [1.09, 1.15]), Wald(1) = 38.38, $p < .001$.[5] See Table 1 for full results of this model.

Study 2

In Study 2, we used data from a prospective longitudinal study of high school seniors to test the main hypothesis that the predictive validity of high school report card grades for 4-year college graduation can be differentially explained by self-regulation. Support for this hypothesis depends on showing that self-regulation plays a role in both report card grades and college graduation, and furthermore, that accounting for this shared variance reduces the association between report card grades and graduation to nonsignificance. We also tested the complementary hypothesis that the predictive validity of admission test scores for college graduation can be differentially explained by cognitive ability. Support for this hypothesis depends on showing that the relationship between admission test scores and college graduation drops to nonsignificance when accounting for the shared variance that cognitive ability explains in both variables.

To increase reliability and validity, we used multiple measures of both self-regulation and cognitive ability (Rushton, Brainerd, & Pressley, 1983). To measure self-regulation, we combined students' self-reported ratings of their own grit and self-control with ratings from classroom teachers on each of these same students' grit and self-control. This strategy aligns with a large precedent in the psychological literature for combining ratings from multiple sources in the assessment of self-regulation (Duckworth & Seligman, 2005; Moffitt et al., 2011; Sorensen, Dodge, & Conduct Problems Prevention Research, 2016; Valiente et al., 2008; Zimmerman & Kitsantas, 2014). Likewise, given the multidimensional and hierarchical organization of cognitive ability (Carroll, 1993; Lubinski, 2004; Neisser et al., 1996), we combined tests of verbal ability, fluid reasoning, working memory, and processing speed.

Method

Participants

The final analytic sample included N = 1,622 high school seniors (M_{age} = 17.94 years, SD = 0.54) from four public schools in the northeastern United States. According to school records, 35.8% of the students were Caucasian, 28.7% were African American, 18.3% were Asian, 13.4% were Latino, and 2.8% mixed race. Race/ethnicity data were missing for 1% of the sample. Just over half (58.1%) were from low-income households as indicated by their eligibility for free or reduced-price lunch. Half (49.7%) of the sample was female. The analytic sample represents about 74% of the high school seniors across the four schools. The demographic characteristics of students in the analytic sample were generally comparable with the aggregate demographic characteristics across schools based on public information (National Center for Education Statistics, 2017a); see Supplemental Material available in the online version of the journal for more information.

Procedure

Students were recruited during the fall semester of their senior year (2012/2013 academic year) into a study examining predictors of college persistence. An informational letter from each school's principal along with an opt-out parent consent form was sent to each student's home. Students whose parents did not opt them out of participation completed assent forms prior to data collection. There was no stopping rule for data collection; the sample size reflects the maximum number of students we were able to recruit within the allotted time provided by the schools.

Students completed measures of self-regulation and cognitive ability on school computers between the months of November 2012 and April 2013. The testing procedure was modified slightly across schools to accommodate different class schedules. Separately, up to two classroom teachers completed measures assessing their students' self-regulation. We obtained senior year report card grades and SAT scores from high school records. In summer 2017, we obtained college enrollment and graduation data from the NSC database.

Measures[6]

Self-regulation. Students completed four items assessing grit (e.g., "I am diligent. I never give up," α = .78) adapted from Duckworth et al. (2007). Students also completed four items from Park, Tsukayama, Goodwin, Patrick, and Duckworth (2017) assessing self-control in the domain of schoolwork (e.g., "I pay attention and resist distractions in class," α = .71) and four items assessing self-control in the domain of interpersonal relationships (e.g., "I can remain calm even when criticized or otherwise provoked," α= .68). The 5-point Likert-type response scale for all items ranged from 1 = *not at all like me* to 5 = *very much like me*. Versions of these scales have been used in prior research and have demonstrated acceptable internal reliability consistency, convergent validity with theoretically related constructs (e.g., conscientiousness, academic diligence), and predictive validity with report card grades (Park et al., 2017; West et al., 2016).

Separately, up to two classroom teachers completed measures of grit and self-control for each of their current students. To minimize burden, teachers provided a single overall rating for each construct (for an identical approach to assessing teacher-reported grit and self-control for middle school students, see Park et al., 2017; Park et al., 2018). This measurement approach follows prior research showing that single-item measures of personality traits demonstrate adequate levels of convergent and discriminant validity, test-retest reliability, and convergence between self- and observer-ratings (Gosling, Rentfrow, & Swann, 2003; Rammstedt & John, 2007). For example, teachers viewed all four grit items simultaneously and then provided an overall grit rating for each student using a 5-point Likert-type scale ranging from 1 = *not at all like my student* to 5 = *very much like my student*. See Supplemental Material available for this article in the online version of the journal for further information about the reliability and validity of these teacher ratings.

Cognitive ability. Students completed measures of verbal ability, fluid reasoning, working memory, and processing speed (described below). These four computer-based performance measures were adapted to accommodate group administration in classrooms (a similar approach to measuring fluid cognitive ability was used in Finn et al., 2014). Pilot testing prior to administration with our target sample revealed that scores on the modified cognitive ability measures converged with scores on the original measures (see Supplemental Material available for this article in the online version of the journal for more information).

Verbal ability was assessed using the Mill Hill Vocabulary Test, Junior version set A (Raven, Raven, & Court, 1998). For this untimed multiple-choice test, participants were shown a target word and then had to select a synonym for the target word from among six response options. Performance on the Mill Hill was calculated as the total number of correct responses, with a maximum score of 33.

Fluid reasoning was assessed using the matrix reasoning subtest of the Kaufmann Brief Intelligence Test (Kaufman & Kaufman, 1990). Students were shown a series of matrices in which one portion of the pattern was missing and asked to select the shape that completed the pattern from a set of response options. The modified test included a total of 36 problems. The task ended after four consecutive incorrect responses or completion of all problems. The number of correct answers before a ceiling of four incorrect trials in a row constituted the raw score, which was converted to an age-normed scaled score. Cases with scaled scores of 40 (n = 49; lowest score possible), believed to be a sign of students' misunderstanding, were treated as missing data.

Working memory was assessed using the Symmetry Span task (Unsworth, Heitz, Schrock, & Engle, 2005). The Symmetry Span combines a memory task (i.e., memorize the position of red squares on a grid) with a decision task (i.e., determine the vertical symmetry of geometric patterns). The task consisted of 12 test blocks, with each block containing between 2 and 5 red squares to memorize. During each trial, students first determined whether a shape was symmetrical (*yes* or *no*) and then had to memorize the position of a red square on a grid. At the end of each block, students were asked to recall the sequence of the red squares in the correct order. Performance was calculated as the percentage of correct responses at each memory load (i.e., 2 squares, 3 squares, 4 squares, 5 squares), which were then averaged to obtain an overall score.

Processing speed was assessed using the Digit Symbol Substitution and Symbol Search tasks from the Wechsler Intelligence Scale for Children (Wechsler, 2003). For the Digit Symbol task, students were shown a key in which nine geometric symbols corresponded to the numbers 1 through 9. Below the key, students were presented with a grid containing symbols and blank spaces. Students had to translate the symbols into numbers with the help of the provided key. During the Symbol Search task, students were shown two geometric symbols on the left portion of the interface for which they had to find a match among five options on the right side of the interface (participants could also select "*no*" to indicate no match). Students were given two minutes per task to answer as many problems as possible. Scaled scores, based on age, were generated from the number of correct responses for each task, and then averaged to form a composite score.

HSGPA. As in Study 1, high schools varied in their grading scales. But unlike Study 1, we were able to consult high school administrators and the schools' handbooks to more precisely rescale grades to a common metric. Instead of computing proportion scores, we used additional information provided by the College Board (2018) to convert HSGPA to a 100-point scale. These conversions had the effect of putting students' senior year report card grades on a common metric without losing information about differences in academic performance and grading standards across schools (which proportion scores did not allow). Slight modifications to the conversion tables were made to accommodate weighted HSGPA provided by two of the four high schools (allowing for a maximum score of 105).

SAT score. As in Study 1, we calculated SAT total scores by summing students' highest reported scores on the mathematics (M = 499.1, SD = 111.8), writing (M = 454.6, SD = 101.1), and critical reading (M = 462.1, SD = 99.3) subtests. The high schools in our sample did not provide ACT score.

College graduation. Students who earned a bachelor's degree within 4 years of high school were classified as 1 = *graduated within 4 years*; all other participants were classified as 0 = *did not graduate within 4 years/not enrolled*.

Demographic characteristics. From high school records, we obtained information about students' gender, race/ethnicity, eligibility for free or reduced-price lunch, and date of birth (to calculate age). Similar to Study 1, all models controlled for gender, race/ethnicity (four dummy codes for African American, Latino, Asian, and multiple or other races/ethnicities versus Caucasian), socioeconomic status (dummy code for eligibility for free or reduced-price lunch vs. not eligible), high school attended (three dummy

Table 2
Descriptive Statistics for Study 2 (*N* = 1,622)

Variables	*M*	*SD*	*N*
Self-regulation			
Self-reported grit	3.81	0.74	1,477
Self-reported schoolwork self-control	3.61	0.71	1,478
Self-reported interpersonal self-control	3.85	0.70	1,477
Teacher 1 ratings of grit	3.45	1.19	1,400
Teacher 1 ratings of schoolwork self-control	3.52	1.25	1,400
Teacher 1 ratings of interpersonal self-control	4.09	1.09	1,392
Teacher 2 ratings of grit	3.38	1.22	1,316
Teacher 2 ratings of schoolwork self-control	3.49	1.24	1,316
Teacher 2 ratings of interpersonal self-control	4.03	1.08	1,316
Cognitive ability			
Fluid reasoning	94.70	15.17	1,340
Working memory	0.46	0.22	1,461
Processing speed	85.53	11.56	1,426
Verbal ability	21.47	4.18	1,459
Academic outcomes			
HSGPA	83.13	9.72	1,615
SAT score	1415.9	284.5	1,005
4-year college graduation	10.6%		1,622

Note. HSGPA = senior year high school grade point average.

codes for each school versus the high school with the largest sample), and age. See Supplementary Material in the online version of the journal for information about demographic differences on study variables.

Analytic Plan

As in Study 1, we tested our hypotheses using *Mplus* Version 7.2 (Muthén & Muthén, 2012) and used full-information maximum likelihood to retain cases with missing data. Data on main study variables were available for between 62% and 100% of observations (see Table 2). The sample size for all analyses (with the exception of descriptive statistics) was 1,622. Except where noted, demographic characteristics (gender, race/ethnicity, eligibility for free or reduced-price lunch, high school attended, age) were included as covariates in the prediction of HSGPA, SAT score, and college graduation.

In the first step, we ran a series of confirmatory factor analyses to assess the measurement of our latent factors, self-regulation and cognitive ability. We used conventional indices and their respective cutoffs to assess fit. Values of .90 or higher for the comparative fit index (CFI) indicate acceptable fit to the data, and values of .95 or higher indicate excellent fit (Bentler &

Bonett, 1980; Schumacker & Lomax, 2010). Root mean square error of approximation (RMSEA) values of .08 or less indicate acceptable fit, and values of .05 or less indicate excellent fit (Browne & Cudeck, 1993; Schumacker & Lomax, 2010). Standardized root mean square residual (SRMR) values of .08 or less indicate good fit to the data (Hu & Bentler, 1999). In these measurement models, we specified the aforementioned demographic characteristics as auxiliary variables (missing data correlates).

In the second step, we fit a series of structural models to test the strength of relationships among key variables. In Model 1, we regressed college graduation on HSGPA and SAT score. In Model 2, we regressed college graduation on self-regulation and cognitive ability. Then, in Model 3, we regressed HSGPA and SAT score on self-regulation and cognitive ability. In each of the three models, we used Wald tests to compare the strengths of different paths on the dependent variable(s). HSGPA and SAT score were allowed to covary, and self-regulation and cognitive ability were also allowed to covary.

Finally, we tested the main hypothesis that the predictive validity of HSGPA for college graduation is explained by self-regulation, whereas the predictive validity of SAT score for college graduation is explained by cognitive ability. In Model 4, we regressed college graduation on HSGPA, SAT score, and self-regulation. Our aim in adding self-regulation as another predictor of college graduation was not to examine its unique contribution to graduation over and above HSGPA and SAT score, but rather to demonstrate that including it in the model would differentially affect the predictive relationship between HSGPA and college graduation compared to SAT score and graduation. Specifically, we hypothesized that controlling for self-regulation would attenuate the significant predictive relation between HSGPA (but not SAT score) and college graduation. In Model 5, we regressed college graduation on HSGPA, SAT score, and cognitive ability. Our reasoning was that, if the relation between SAT score and college graduation is explained by the role that cognitive ability plays in both variables, then controlling for cognitive ability should attenuate the significant predictive relation between SAT score (but not HSGPA) and college graduation.

Results and Discussion

Descriptive Statistics

Approximately 11% of students in our sample had obtained a bachelor's degree within 4 years. The lower graduation rate in comparison to Study 1 is likely attributable to the fact that this sample involved high school students (not Common App applicants), the majority of whom were from low-income households as indicated by eligibility for free and reduced-price lunch. Students from lower income households enroll in and graduate from college at lower rates than students from higher income households (National

Center for Education Statistics, 2011, 2014b, 2015). See Table 2 for descriptive statistics on the other study variables.

Table 3 provides bivariate correlations among study variables. As in Study 1, HSGPA and SAT score were strongly positively correlated (r = .49, p < .001).

Measurement Models

To create a latent self-regulation factor, we specified three first-order latent factors representing the shared variance among the manifest grit and self-control scales—one for self-report and two for teacher-report measures. These, in turn, served as indicators for a second-order latent factor, self-regulation. The initial self-regulation factor provided acceptable-to-good fit to the data: $\chi^2(24) = 160.92$, $p < .001$, CFI = .976, RMSEA = .059, SRMR = .040. For increased parsimony and interpretability, we constrained the loadings on each indicator across the two teacher factors to equality; we also constrained the two latent teacher factor loadings on the second-order factor to equality. The constrained model also provided acceptable-to-good fit to the data: $\chi^2(27) = 166.34$, $p < .001$, CFI = .976, RMSEA = .056, SRMR = .044. The constrained model did not worsen model fit compared to the baseline model, $\Delta\chi^2(3) = 5.42$, $p = .143$, so we retained the constrained model in subsequent analysis. The variance explained for each of the first-order indicators ranged from 37% to 92% (representing standardized loadings ranging from .61 to .96, ps < .001).

Similarly, we created a latent cognitive ability factor. In this model, verbal ability, fluid reasoning, working memory, and processing speed were specified as indicators of a latent factor. The initial cognitive ability factor provided somewhat acceptable fit to the data, $\chi^2(2) = 41.81$, $p < .001$, CFI = .953, RMSEA = .111, SRMR = .033, but the RMSEA suggested room for improvement. We constrained loadings of the three indicators of fluid abilities (fluid reasoning, processing speed, working memory; Finn et al., 2014) to equality, which resulted in acceptable-to-good fit to the data: $\chi^2(4) = 43.32$, $p < .001$, CFI = .953, RMSEA = .078, SRMR = .036. The constrained model did not worsen model fit compared to the baseline model, $\Delta\chi^2(2) = 1.52$, $p = .469$, so we retained the constrained model in subsequent analysis. The variance explained for each of the four individual indicators ranged from 26% to 40% (representing standardized loadings ranging from .51 to .63, ps < .001). See the Supplementary Material available in the online version of the journal for a graphical depiction of the latent factor measurement models for both self-regulation and cognitive ability.

We also assessed the fit of a joint measurement model consisting of both latent factors (self-regulation and cognitive ability), with these factors permitted to covary. This measurement model provided acceptable fit to the

Table 3

Bivariate Correlations in Study 2 (N = 1,622)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Self-reported grit															
2. Self-reported schoolwork self-control	.61***														
3. Self-reported interpersonal self-control	.48***	.47***													
4. Teacher 1 ratings of grit	.11***	.20***	.15***												
5. Teacher 1 ratings of schoolwork self-control	.10**	.19***	.14***	.79***											
6. Teacher 1 ratings of interpersonal self-control	.06*	.14***	.19***	.54***	.64***										
7. Teacher 2 ratings of grit	.10**	.19***	.05	.38***	.39***	.30***									
8. Teacher 2 ratings of schoolwork self-control	.07*	.18***	.07*	.38***	.41***	.36***	.82***								
9. Teacher 2 ratings of interpersonal self-control	.01	.10**	.10**	.28***	.35***	.38***	.55***	.63***							
10. Fluid reasoning	−.04	−.04	.10***	.11***	.14***	.17***	.17***	.20***	.18***						
11. Working memory	−.05*	−.12***	.01	.00	.02	.03	.02	.04	.05	.40***					
12. Processing speed	−.06*	−.07**	−.02	.09**	.11***	.13***	.11***	.14***	.19***	.35***	.41***				
13. Verbal ability	−.05	−.11***	.12***	.14***	.17***	.17***	.15***	.17***	.20***	.39***	.23***	.31***			
14. HSGPA	.08**	.15***	.11***	.47***	.52***	.39***	.51***	.50***	.41***	.31***	.15***	.26***	.28***		
15. SAT score	−.06	−.07*	.02	.24***	.27***	.18***	.23***	.23***	.26***	.44***	.28***	.36***	.59***	.49***	
16. 4-year college graduation	.07**	.07*	.08**	.19***	.22***	.16***	.19***	.19***	.16***	.15***	.06*	.13***	.17***	.27***	.27***

Note. HSGPA = senior year high school grade point average. Pairwise deletion resulted in sample sizes from 840 to 1,615.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

data: $\chi^2(66) = 424.38$, $p < .001$, CFI = .948, RMSEA = .058, SRMR = .063. The two factors were positively correlated ($r = .30$, $p < .001$).

Assessing Strength of Relationships Among Key Variables

In Models 1 through 3, we assessed the strength of relationships among self-regulation, cognitive ability, HSGPA, SAT score, and college graduation. Full results of Models 1 and 2 are provided in Table 4; results of Model 3 are reported in Table 5.

Model 1: Predictive validity of HSGPA and SAT score for college graduation. Adjusting for demographic characteristics, HSGPA (*OR* = 2.79, $p < .001$, 95% CI [2.05, 3.80]) and SAT score (*OR* = 1.64, $p < .001$, 95% CI [1.25, 2.14]) each incrementally predicted college graduation, but as in Study 1, HSGPA was a stronger predictor, Wald(1) = 4.33, $p = .037$.

Model 2: Predictive validity of self-regulation and cognitive ability for college graduation. Adjusting for demographic characteristics, self-regulation (*OR* = 3.62, $p < .001$, 95% CI [2.33, 5.61]) and cognitive ability (*OR* = 1.68, $p = .002$, 95% CI [1.21, 2.34]) each incrementally predicted college graduation, but self-regulation was a stronger predictor, Wald(1) = 6.34, $p = .012$.

Model 3: Predictive validity of self-regulation and cognitive ability for HSGPA and SAT score. Adjusting for demographic characteristics, self-regulation was a stronger predictor of HSGPA ($\beta = .77$, $p < .001$, 95% CI [.70, .83]) than of SAT score ($\beta = .19$, $p < .001$, 95% CI [.12, .27]), Wald(1) = 145.86, $p < .001$. Cognitive ability, on the other hand, was a stronger predictor of SAT score ($\beta = .74$, $p < .001$, 95% CI [.68, .81]) than of HSGPA ($\beta = .22$, $p < .001$, 95% CI [.15, .29]), Wald(1) = 141.32, $p < .001$.

To partition the unique and shared variance that self-regulation and cognitive ability explained in HSGPA and SAT score, we fit a series of follow-up structural models. Figures 1 and 2 show that self-regulation explained 41.4% of the variance in HSGPA and 2.6% of the variance in SAT score. Cognitive ability explained 39.6% of the variance in SAT score and 3.4% of the variance in HSGPA. Notably, only 7.1% of the explained variance in SAT score and 8.3% of the explained variance in HSGPA was shared by self-regulation and cognitive ability. Full results of these additional structural models are provided in the Supplemental Material available for this article in the online version of the journal.

Main Analysis

In Models 4 and 5, we tested the hypothesis that the predictive validity of HSGPA for college graduation is explained by self-regulation, whereas the predictive validity of SAT scores for college graduation is explained by cognitive ability. Full results of Models 4 and 5 are provided in Table 6.

Table 4
Results of Structural Models Predicting On-Time College Graduation in Study 2 (N = 1,622)

	Model 1					Model 2				
Variables	β	*OR*	LCI	UCI	*p*	β	*OR*	LCI	UCI	*p*
HSGPA	0.43	2.79	2.049	3.796	<.001					
SAT score	0.21	1.64	1.251	2.141	<.001					
Self-regulation						0.53	3.62	2.329	5.611	<.001
Cognitive ability						0.22	1.68	1.212	2.340	.002
Demographic characteristics										
School Dummy 1	0.15	2.14	1.355	3.369	.001	0.17	2.52	1.515	4.198	<.001
School Dummy 2	0.11	2.31	1.383	3.858	.001	0.08	1.94	1.114	3.374	.019
School Dummy 3	0.01	1.04	0.584	1.841	.902	0.14	2.28	1.177	4.422	.015
African American	−0.02	0.93	0.564	1.522	.763	−0.06	0.72	0.421	1.237	.235
Latino	−0.07	0.61	0.282	1.337	.219	−0.09	0.51	0.225	1.178	.116
Asian	−0.02	0.90	0.557	1.461	.676	−0.05	0.71	0.411	1.239	.231
Other ethnicities	−0.06	0.44	0.119	1.591	.208	−0.09	0.29	0.070	1.173	.082
Female	0.11	1.66	1.144	2.395	.008	0.02	1.11	0.711	1.738	.642
Free/reduced-price lunch	−0.06	0.76	0.518	1.119	.165	−0.03	0.85	0.555	1.290	.438
Age	−0.01	0.99	0.807	1.211	.912	0.00	1.01	0.810	1.259	.932

Note. OR = odds ratio. LCI/UCI = lower/upper 95% confidence interval of the odds ratio. β = standardized regression coefficient. HSGPA = senior year high school grade point average.

Table 5

Results of Structural Model Predicting HSGPA and SAT Score in Study 2 (*N* = 1,622)

	Model 3: HSGPA				Model 3: SAT Score			
Variables	β	LCI	UCI	*p*	β	LCI	UCI	*p*
Self-regulation	0.77	0.701	0.831	<.001	0.19	0.116	0.267	<.001
Cognitive ability	0.22	0.146	0.291	<.001	0.74	0.678	0.807	<.001
Demographic characteristics								
School Dummy 1	−0.07	−0.122	−0.022	.005	0.00	−0.049	0.056	.890
School Dummy 2	−0.09	−0.138	−0.047	<.001	0.13	0.083	0.171	<.001
School Dummy 3	0.18	0.118	0.235	<.001	0.12	0.062	0.174	<.001
African American	−0.08	−0.133	−0.028	.003	−0.07	−0.122	−0.011	.019
Latino	−0.05	−0.102	0.003	.063	−0.02	−0.081	0.035	.434
Asian	−0.01	−0.067	0.039	.615	0.01	−0.047	0.057	.850
Other ethnicities	−0.04	−0.079	0.008	.112	−0.06	−0.108	−0.020	.004
Female	−0.03	−0.084	0.019	.219	−0.02	−0.074	0.025	.342
Free/reduced-price lunch	0.06	0.014	0.107	.010	0.03	−0.016	0.083	.184
Age	–0.01	–0.059	0.035	.615	–0.03	–0.086	0.021	.231

Note. β = standardized regression coefficient; LCI/UCI = lower/ 95% confidence interval of the standardized regression coefficient; HSGPA = senior year high school grade point average.

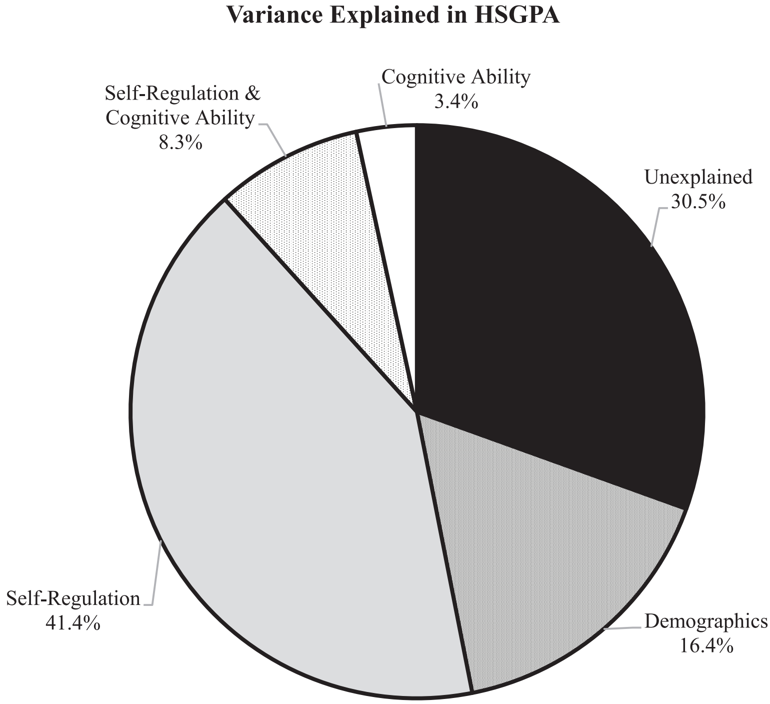

Figure 1. **Variance explained in high school grade point average (HSGPA) in Study 2 (*N* = 1,622).**

Note. This figure depicts the estimated unique and shared variance in high school GPA explained by self-regulation and cognitive ability after partialling variance explained by demographic factors. It is based on the models described in Table S14. Demographic characteristics included gender, race/ethnicity, eligibility for free or reduced-price lunch, high school attended, and age.

Model 4: Self-regulation explains the relationship between HSGPA and college graduation. In this model, as in prior models, HSGPA and SAT score were specified as predictors of college graduation. Critically, self-regulation and demographic characteristics, but not cognitive ability, were also specified as predictors of HSGPA, SAT score, and college graduation. After partialling the variance in HSGPA, SAT score, and college graduation due to self-regulation, HSGPA was no longer a significant predictor of college graduation (OR = 1.46, p = .242, 95% CI [0.78, 2.74]), but SAT score remained significantly associated with college graduation (OR = 1.72, p < .001, 95% CI [1.29, 2.29]). Figure 3 (upper panel) graphically depicts the results of Model 4.

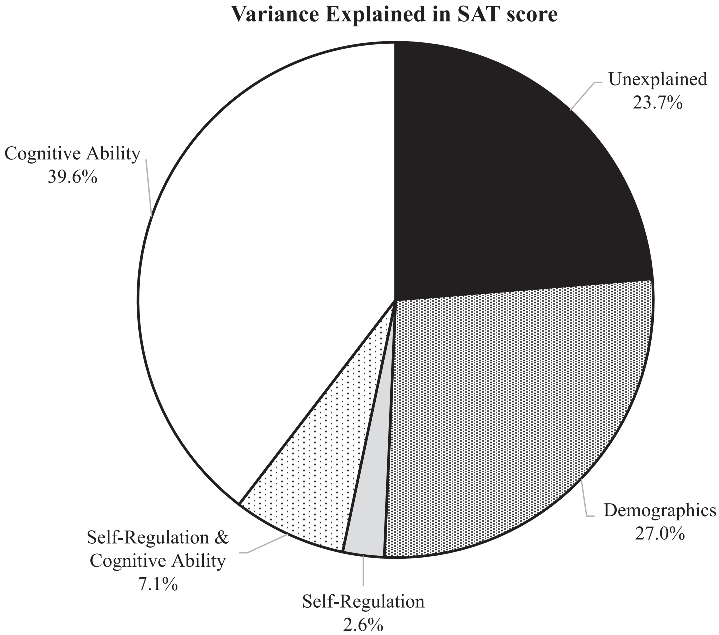

Figure 2. **Variance explained in SAT score in Study 2 (*N* = 1,622).**

Note. This figure depicts the estimated unique and shared variance in SAT score explained by self-regulation and cognitive ability after partialling variance explained by demographic factors. It is based on the models described in Table S15. Demographic characteristics included gender, race/ethnicity, eligibility for free or reduced-price lunch, high school attended, and age.

Model 5: Cognitive ability explains the relationship between SAT score and college graduation. In this model, HSGPA and SAT score were again specified as predictors of college graduation. Critically, cognitive ability and demographic characteristics, but not self-regulation, were also specified as predictors of HSGPA, SAT score, and college graduation. After partialling the variance in HSGPA, SAT score, and college graduation due to cognitive ability, SAT score was no longer statistically significantly associated with college graduation (OR = 1.67, p = .080, 95% CI [0.94, 2.96]) whereas HSGPA remained a significant predictor of college graduation (OR = 2.65, p < .001, 95% CI [1.95, 3.61]). Figure 3 (lower panel) graphically depicts the results of Model 5.

General Discussion

In two prospective longitudinal studies, we unpacked the differential predictive validities of high school report card grades and standardized

Table 6

Results of Structural Models Predicting On-Time College Graduation From Self-Regulation, Cognitive Ability, HSGPA, and SAT Score in Study 2 (*N* = 1,622)

	Model 4					Model 5				
Variables	β	*OR*	LCI	UCI	*p*	β	*OR*	LCI	UCI	*p*
HSGPA	0.16	1.46	0.775	2.742	.242	0.41	2.65	1.951	3.610	<.001
SAT score	0.23	1.72	1.289	2.287	<.001	0.23	1.67	0.941	2.959	.080
Self-regulation	0.35	2.32	1.115	4.824	.024					
Cognitive ability						0.02	1.05	0.544	2.042	.877
Demographic characteristics										
School Dummy 1	0.14	2.15	1.343	3.454	.001	0.15	2.17	1.370	3.447	.001
School Dummy 2	0.07	1.76	0.981	3.143	.058	0.10	2.22	1.313	3.756	.003
School Dummy 3	0.07	1.54	0.771	3.076	.222	0.01	1.06	0.578	1.956	.843
African American	−0.04	0.82	0.485	1.384	.456	−0.01	0.95	0.576	1.558	.831
Latino	−0.08	0.56	0.247	1.246	.153	−0.06	0.66	0.300	1.432	.289
Asian	−0.04	0.77	0.458	1.300	.330	−0.01	0.92	0.568	1.490	.734
Other ethnicities	−0.07	0.38	0.101	1.460	.160	−0.06	0.44	0.119	1.657	.227
Female	0.06	1.30	0.849	2.004	.225	0.11	1.69	1.159	2.457	.006
Free/reduced-price lunch	−0.04	0.82	0.544	1.221	.321	−0.05	0.77	0.516	1.151	.203
Age	−0.01	0.98	0.790	1.204	.818	0.00	1.01	0.816	1.239	.960

Note. OR = odds ratio; LCI/UCI = lower/upper 95% confidence interval of the odds ratio; β = standardized regression coefficient; HSGPA = senior year high school grade point average.

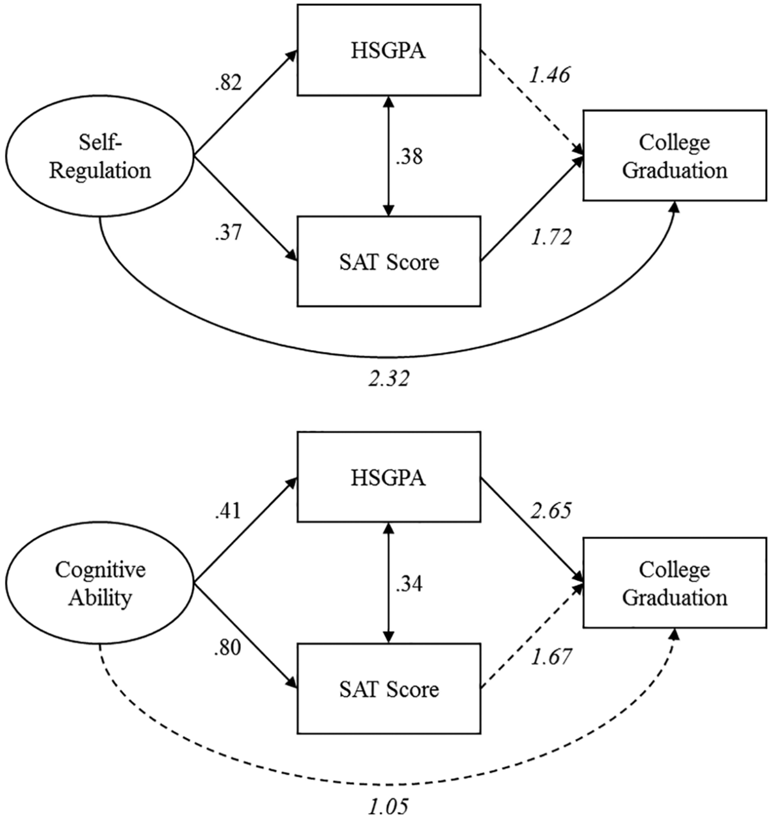

Figure 3. **Results of structural models predicting on-time college graduation from self-regulation, cognitive ability, HSGPA, and SAT score in Study 2 (*N* = 1,622).**

Note. This figure illustrates the relationship between self-regulation, senior year high school grade point average (HSGPA), SAT score, and college graduation (upper panel), and cognitive ability, HSGPA, SAT score, and college graduation (lower panel). Path estimates to college graduation are odds ratios (italicized), whereas all other paths are standardized regression coefficients (or correlations). Ovals represent latent factors and rectangles represent manifest variables. Solid lines indicate parameter estimates that are statistically significant at $p < .05$; dashed lines are not statistically significantly different from zero. HSGPA, SAT score, and college graduation were also regressed on demographic characteristics (gender, race/ethnicity, eligibility for free or reduced-price lunch, high school attended, age). Upper Panel (Model 4): Though HSGPA predicts college graduation, this relationship falls below statistical significance when self-regulation is included as an additional variable in the model. Lower Panel (Model 5): Though SAT score predicts college graduation, this relationship falls below statistical significance when cognitive ability is included as an additional variable in the model.

admissions test scores for college graduation observed in earlier cohorts. In Study 1, report card grades out-predicted admissions test scores on 4-year graduation outcomes in a national sample of 47,303 students who applied to college via the Common App for the 2009/2010 academic year. In Study 2, we followed 1,622 seniors from four urban high schools in the graduating Class of 2013. We found that self-regulation, derived from self-report and teacher ratings administered during the students' senior year in high school, explained more variance in report card grades than in SAT scores, whereas a battery of cognitive ability tests explained more variance in SAT scores than in report card grades. Moreover, the incremental predictive validity of high school grades for college graduation was explained by self-regulation, whereas the incremental predictive validity of SAT scores for college graduation was explained by cognitive ability.

Our findings suggest that report card grades provide information about self-regulation not captured by admissions test scores. In particular, the grades assigned to students by their high school teachers reveal quite a bit about their capacity to resist momentary temptations, regulate emotions, and sustain effort across days, months, and years in pursuit of important goals. Relating grades to self-regulation helps explain why, among more than 10,000 adolescents in the National Longitudinal Survey of Adolescent to Adult Health sample, high school report card grades provided incremental predictive validity over and above measured cognitive ability for lifetime educational attainment and income in adulthood (French et al., 2015). Though there is certainly room for improvement in how teachers assign report card grades (Brookhart, 1991, 1994), they are a powerful indicator of the capacity to pursue goals despite temptations, distractions, and other obstacles.

Implications for Theory and Practice

The current investigation contributes to evidence that competencies other than cognitive ability matter for predicting success in life. For example, in a cohort of a thousand children followed from birth to age 32, composite measures of childhood self-regulation and intelligence emerged as distinct predictors of financial security, physical health, criminal offenses, and substance dependence (Moffitt et al., 2011). As in the current investigation, these findings held true when controlling for family socioeconomic status. Of course, it is not only self-regulation that matters. More generally, an array of personal qualities—variously referred to as character skills (Heckman & Kautz, 2013), social-emotional competencies (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011), personality traits (Poropat, 2009; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), or 21st-century skills (Pellegrino & Hilton, 2012)—have emerged in recent decades as surprisingly powerful predictors of success in school and beyond.

What are the implications of the current findings for policy and practice? Specific prescriptive recommendations are beyond the scope of this investigation. However, the discovery that high report card grades signal strong student self-regulation suggest, at a minimum, that grades be accorded careful consideration in college admissions. Currently, *U.S. News & World Report* weights admissions test scores more than three times as heavily as high school class rank in their college rankings list (Morse & Brooks, 2018). In contrast, recent surveys of college admissions officers suggest increasing appreciation for grades: In 2006, a higher percentage of 4-year colleges gave "considerable importance" to standardized admission test scores (60%) than high school grades in all courses (51%; National Association for College Admission Counseling, 2008). By 2017, these were reversed: the percentage of colleges attributing "considerable importance" to high school grades in all courses was 77.1% and to admission test scores 54.3% (National Association for College Admission Counseling, 2017).

Affirmation of the relevance of teacher-assigned grades, however, is not an indictment of standardized admissions tests. In our investigation, test scores added unique predictive power, over and above grades and demographic characteristics, for college graduation. The incremental predictive validity of the SAT was largely explained by cognitive ability. More generally, a massive literature has demonstrated the predictive validity of both standardized achievement tests and tests of cognitive ability (Berry & Sackett, 2009; Deary et al., 2007; Kuncel & Hezlett, 2007; Sackett et al., 2009). College admissions tests including the SAT and ACT represent the only standardized element of the admissions profile and have definite value. Accordingly, after demonstrating the superior predictive validity of report card grades for performance in college, Bowen et al. (2009) concluded: "In general, it is a judicious combination of high school grades and achievement tests (especially AP tests, including the writing component) that we regard as especially promising" (pp. 132–133).

Moreover, our findings do not speak to possible unintended effects of policy changes—particularly in equilibrium as students, their families, and educators respond to new incentives and disincentives. For example, emphasizing the importance of report card grades might inadvertently contribute to some students torturing themselves over an A-minus (Grant, 2018) and the unintended neglect of other aspects of healthy development, including civic and moral character (Berkowitz, 2012; Brooks, 2015), prosocial purpose (Damon, 2008), happiness (Harker & Keltner, 2001), and intellectual curiosity about subjects not taught in school (Ackerman, 1994). Assigning more weight to grades in the admissions process may also exacerbate recent trends toward grade inflation, particularly in high schools serving students from higher socioeconomic backgrounds (Hurwitz & Lee, 2018). Although grade inflation does not necessarily undermine predictive validity (Pattison et al., 2013), it is common sense that students stand to benefit from honest and accurate feedback (EdNavigator, 2018).

In our view, the limitations of any single metric of student competence—for there is no such thing as a perfectly unbiased, error-free, and complete measure (Duckworth & Yeager, 2015)—recommend a more holistic approach to college admissions. Teacher recommendations, for example, can provide a valuable perspective, particularly when open-ended responses complement validated rating scales such as those we employed in our measure of self-regulation (Kuncel, Kochevar, & Ones, 2014; Walters, Kyllonen, & Plante, 2006). Likewise, information about student abilities and interests are revealed in the pattern of their participation in extracurricular activities—including sports, arts and music, student government, community service, and paid work (Gardner, Roth, & Brooks-Gunn, 2008; Hutt et al., 2018; Troutman & Dufur, 2007; Willingham, 1985). Though our findings affirm the importance of self-regulation to academic success, we caution against the use of self-report questionnaires for admissions decisions, since answers can be easily faked (Duckworth, 2016).

Limitations and Future Directions

This investigation represents an initial step in describing the personal qualities that can explain differences in the predictive validity of the two most important college admissions criteria: high school grades and admissions test scores. There are, however, limitations of the current investigation that suggest promising directions for future research. For example, in Study 2, to reduce burden on classroom teachers, we relied on single-item teacher ratings of grit and self-control. Though this procedure has been used in other research (Park et al., 2017; Park et al., 2018) and these measures demonstrated evidence of reliability and construct validity in our data (see Supplemental Material available for this article in the online version of the journal), future research should include more extensive teacher-report measures as well as parent-report and behavioral measures of self-regulation.

A second limitation of the current investigation is our exclusive focus on self-regulation. Future studies might examine how other personal qualities, including study skills (Credé & Kuncel, 2008), relate to performance in high school and college. Additional research is also needed to understand how self-regulation and cognitive ability result from and interact with aspects of students' school and home environments. Relatedly, the present investigation did not attempt a full accounting of situational factors influencing pathways through college (Goldrick-Rab, 2006; Goldrick-Rab & Pfeffer, 2009).

Though the participants in Study 2 represented a diverse convenience sample of high school seniors, it was not nationally representative. For example, our sample included more African American (29%) and fewer Caucasian (36%) students than the national population of high school graduates in 2013 (16% and 50%, respectively, National Center for Education Statistics, 2014a). The low rates of college graduation in Study 2 also reduced variance in the outcome to explain. Likewise, Study 1 was based on students in the Common Application population, which favors more selective colleges. The external validity of the present findings, therefore, cannot be assumed, and we encourage replication studies, ideally including students from representative samples and even students in different countries.

Another limitation is the potential for selection bias. Because students apply and are admitted to college based on the strength of their test scores and high school grades, range restriction may have attenuated relationships with graduation outcomes. Observed standard deviations for the SAT (Study 1: *SD* = 99.2–102.5; Study 2: *SD* = 99.3–111.8) and ACT section scores (Study 1: *SD* = 4.7–5.4) in our data were somewhat smaller than population norms for the SAT in 2009 and 2013 (*SD* = 111–118; College Board, 2009, 2013) and ACT in 2008 and 2009 (*SD* = 5.0–5.1; National Center for Education Statistics, 2009). We do not know the degree to which the variance in report card grades may have been restricted in our samples. We note that for Study 1, SAT/ACT means and standard deviations in the subsample of 47,303 students were almost identical to the full sample of 311,308 students who completed the Common Application for admission to the 2009/2010 academic year. Thus, test scores in our analytic subsample appeared representative of the population of students who completed the Common Application.

Finally, more research is needed to make causal inferences about the roles of self-regulation and cognitive ability, respectively, on student performance outcomes. An improvement on the design we used here would be to administer measures repeatedly and examine how changes in self-regulation predict subsequent changes in outcomes (Duckworth et al., 2010). Another possibility is to capitalize on natural experiments. For example, the SAT was redesigned in 2016 to better reflect material learned in the classroom (College Board, 2015)—a move recommended by Atkinson and Geiser (2009) and Bowen et al. (2009). If the redesigned SAT fulfills its intent, then we would predict stronger correlations between SAT scores and students' self-regulation than those observed in the current investigation. Such changes might also have the effect of reducing the magnitude of the correlation between cognitive ability and SAT scores. More research is needed to test these possibilities, although emerging work on curriculum-sampling tests provides indirect support (Niessen, Meijer, & Tendeiro, 2018).

Conclusion

The present investigation suggests that report card grades earned in high school are better predictors of on-time college graduation than are admissions test scores in part because grades carry information about self-regulation. Crossing the finish line of college requires not only cognitive ability, which is differentially reflected in college admissions tests, but also the self-regulatory competencies evinced in high school report card grades. Since education is by definition more concerned with developing students than in selecting them, we hope these findings spur interest in interventions aimed at increasing aspects of self-regulation (Eskreis-Winkler et al., 2016; Oettingen, 2014; Yeager et al., 2014), not only in adolescence (Dahl, Allen, Wilbrecht, & Suleiman, 2018; Yeager, Dahl, & Dweck, 2018) but also earlier in life (Blair & Diamond, 2008). And, finally, since grades represent the fulfillment of a "local contract" between students and teachers (Willingham et al., 2002, p. 28), it is only fitting that we end this article by affirming the responsibility of adults for helping young people succeed. Indeed, teachers have enduring influence on students (Kraft, 2019), not only in the cognitive competencies indexed by standardized tests but also the self-regulatory competencies captured in report card grades (Jackson, 2018).

Notes

The research and writing of the article was made possible by the National Institute on Aging R24-AG048081-01, the Bill and Melinda Gates Foundation (OPP1046454), the John Templeton Foundation (54929), and the National Science Foundation (DRL 1235958 and IIS 1523091). Any opinions, findings and conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the funding agencies. The authors are grateful to the students and teachers who participated in this research and to the school administrators who helped facilitate it.

Supplemental material is available for this article in the online version of the journal.

[1]In 1990, the College Board changed the name of the SAT to the Scholastic Assessment Test and then, in 1996, dropped the name altogether—the SAT is no longer an acronym insofar as, officially, it stands for nothing. In 2005, the SAT was substantially revised (e.g., dropping verbal analogies, incorporation of a writing section). In 2016, after the data collection in the current investigation, the SAT was again revised (e.g., making the writing section optional, reduced emphasis on general reasoning, see https://collegereadiness.collegeboard.org/sat/inside-the-test/compare-old-new-specifications).

[2]The creator of the ACT, E. F. Lindquist (1958) emphasized that "the examination must make (the student) feel that he has earned the right to go to college by his own efforts, not that he is entitled to college because of his innate abilities or aptitudes. . . . In other words, the examination must be regarded by him as an *achievement* test" (p. 109). The ACT aimed to adhere more closely to what students were taught in the classroom and included, in addition to math and verbal sections, social studies, and science. As mentioned in the main text, the SAT and ACT scores are nevertheless extremely highly correlated.

[3]The data in this study were limited to applications completed online (rather than on paper). Of the 311,308 students, we had online secondary school reports for 47,303 students, which contained critical information about high school grade point average. We assume that the majority of remaining secondary school reports were submitted on paper, and therefore were not available to our team.

[4]The Supplemental Material available for this article in the online version of the journal contains results from a series of follow-up analyses testing the robustness of the main findings reported in both Study 1 and Study 2.

[5]We fit another model predicting on-time college graduation from HSGPA and SAT/ACT score, but without covarying demographic characteristics. The incremental predictive validity of HSGPA for college graduation (OR = 1.28, $p < .001$, 95% CI [1.25, 1.31]) was stronger than the incremental predictive validity of SAT/ACT score (OR = 1.18, $p < .001$, 95% CI [1.16, 1.21]), Wald(1) = 15.15, $p < .001$.

[6]Items for all scales used in this study are reported in the Supplementary Material available in the online version of the journal.

References

Ackerman, P. L. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D. K. Detterman (Ed.), *Current topics in human intelligence* (Vol. 4: Theories of Intelligence, pp. 1–27). Norwood, NJ: Ablex.

ACT. (2013). *ACT–SAT concordance: A tool for comparing scores*. Retrieved from https://www.act.org/content/dam/act/unsecured/documents/reference.pdf

Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher*, *38*, 665–676. doi:10.3102/0013189X09351981

Attewell, P., Heil, S., & Reisel, L. (2011). Competing explanations of undergraduate noncompletion. *American Educational Research Journal*, *48*, 536–559. doi:10.3102/0002831210392018

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, *48*, 5–37. doi:10.1016/j.jsp.2009.10.001

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606. doi:10.1037/0033-2909.88.3.588

Berkowitz, M. W. (2012). Moral and character education. In K. R. Harris, S. Graham, T. Urdan, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook: Vol. 2. Individual differences and cultural and contextual factors* (pp. 247–264). Washington, DC: American Psychological Association.

Berry, C. M., & Sackett, P. R. (2009). Individual differences in course choice result in underestimation of the validity of college admissions systems. *Psychological Science*, *20*, 822–830. doi:10.1111/j.1467-9280.2009.02368.x

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology*, *20*, 899–911. doi:10.1017/S0954579408000436

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton, NJ: Princeton University Press.

Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement*, *10*, 35–36. doi:10.1111/j.1745-3992.1991.tb00182.x

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, *7*, 279–301. doi:10.1207/s15324818ame0704_2

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., . . . Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, *86*, 803–848. doi:10.3102/0034654316672069

Brooks, D. (2015). *The road to character*. New York, NY: Random House.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural models* (pp. 136–162). Newbury Park, CA: Sage.

Camara, W., Kimmel, E., Scheuneman, J., & Sawtell, E. A. (2003). *Whose grades are inflated?* Retrieved from https://files.eric.ed.gov/fulltext/ED563203.pdf

Carroll, J. B. (1992). Cognitive abilities: The state of the art. *Psychological Science*, *3*, 266–270. doi:10.1111/j.1467-9280.1992.tb00669.x

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press.

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin*, *92*, 111–135. doi:10.1037/0033-2909.92.1.111

College Board. (2009). *2009 College-bound seniors: Total group profile report*. Retrieved from http://media.collegeboard.com/digitalServices/pdf/research/cbs-2009-national-TOTAL-GROUP.pdf

College Board. (2013). *2013 College-bound seniors: Total group profile report*. Retrieved from http://media.collegeboard.com/digitalServices/pdf/research/2013/TotalGroup-2013.pdf

College Board. (2015). *Test specifications for the redesigned SAT*. Retrieved from https://collegereadiness.collegeboard.org/pdf/test-specifications-redesigned-sat-1.pdf

College Board. (2018). *How to convert your GPA to a 4.0 scale*. Retrieved from https://pages.collegeboard.org/how-to-convert-gpa-4.0-scale

Complete College America. (2014). *Four-year myth*. Retrieved from https://completecollege.org/wp-content/uploads/2017/05/4-Year-Myth.pdf

Cox, D. R., & Hinkley, D. V. (1974). *Theoretical statistics*. New York, NY: Wiley.

Coyle, T. R., & Pillow, D. R. (2008). SAT and ACT predict college GPA after removing g. *Intelligence*, *36*, 719–729. doi:10.1016/j.intell.2008.05.001

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, *3*, 425–453. doi:10.1111/j.1745-6924.2008.00089.x

Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, *12*, 53–72. doi:10.1207/s15324818ame1201_4

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature*, *554*, 441–450. doi:10.1038/nature25770

Damon, W. (2008). *The path to purpose: How young people find their calling in life*. New York, NY: Free Press.

Davisson, E. K., & Hoyle, R. H. (2017). The social-psychological perspective on self-regulation. In T. Egner (Ed.), *The Wiley handbook of cognitive control* (pp. 440–454). Hoboken, NJ: Wiley.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13–21. doi:10.1016/j.intell.2006.02.001

Duckworth, A., & Gross, J. J. (2014). Self-control and grit: Related but separable determinants of success. *Current Directions in Psychological Science*, *23*, 319–325. doi:10.1177/0963721414541462

Duckworth, A. L. (March 26, 2016). Don't grade schools on grit. *The New York Times*. Retrieved from https://www.nytimes.com/2016/03/27/opinion/sunday/dont-grade-schools-on-grit.html

Duckworth, A. L., & Carlson, S. M. (2013). Self-regulation and school success. In B. W. Sokol, F. M. E. Grouzet, & U. Müller (Eds.), *Self-regulation and autonomy: Social and developmental dimensions of human conduct* (pp. 208–230). New York, NY: Cambridge University Press.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*, 1087–1100. doi:10.1037/0022-3514.92.6.1087

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*, 939–944. doi:10.1111/j.1467-9280.2005.01641.x

Duckworth, A. L., Taxer, J., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-control and academic achievement. *Annual Review of Psychology*, *70*, 373–399.

Duckworth, A. L., Tsukayama, E., & May, H. (2010). Establishing causality using longitudinal hierarchical linear modeling: An illustration predicting achievement from self-control. *Social Psychology and Personality Science*, *1*, 311–317. doi:10.1177/1948550609359707

Duckworth, A. L., Weir, D., Tsukayama, E., & Kwok, D. (2012). Who does well in life? Conscientious adults excel in both objective and subjective success. *Frontiers in Personality Science and Individual Differences*, *3*, 356.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*, 237–251. doi:10.3102/0013189x15584327

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432. doi:10.1111/j.1467-8624.2010.01564.x

EdNavigator. (December 11, 2018). *Muddled: How confusing information from schools is failing American families*. Retrieved from https://medium.com/ednavigator/muddled-78ce70407fda

Eskreis-Winkler, L., Gross, J. J., & Duckworth, A. L. (2017). Grit: Sustained self-regulation in the service of superordinate goals. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory and applications* (3rd ed., pp. 380–395). New York, NY: Guilford Press.

Eskreis-Winkler, L., Shulman, E. P., Young, V., Tsukayama, E., Brunwasser, S. M., & Duckworth, A. (2016). Using wise interventions to motivate deliberate practice. *Journal of Personality and Social Psychology*, *111*, 728–744. doi:10.1037/pspp0000074

Faust, D. (October 23, 2014). College helps students dream of more than a salary. *USA Today*. Retrieved from https://www.usatoday.com/story/opinion/2014/10/23/harvard-liberal-arts-college-university-cardinal-newman-thinking-wisdom-column/17779447/

Finn, A. S., Kraft, M. A., West, M. R., Leonard, J. A., Bish, C. E., Martin, R. E., . . . Gabrieli, J. D. E. (2014). Cognitive skills, student achievement tests, and schools. *Psychological Science*, *25*, 736–744. doi:10.1177/0956797613516008

French, M. T., Homer, J. F., Popovici, I., & Robins, P. K. (2015). What you do in high school matters: High school GPA, educational attainment, and labor market earnings as a young adult. *41*, 370–386. doi:10.1057/eej.2014.22

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*, 373–378. doi:10.1111/j.0956-7976.2004.00687.x

Gardner, M., Roth, J., & Brooks-Gunn, J. (2008). Adolescents' participation in organized activities and developmental success 2 and 8 years after high school: Do sponsorship, duration, and intensity matter? *Developmental Psychology*, *44*, 814–830. doi:10.1037/0012-1649.44.3.814

Geiser, S., & Santelices, M. V. (2007). *Validity of high school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes*. Retrieved from https://cshe.berkeley.edu/sites/default/files/publications/rops.geiser._sat_6.13.07.pdf

Goldrick-Rab, S. (2006). Following their every move: An investigation of social-class differences in college pathways. *Sociology of Education*, *79*, 67–79. doi:10.1177/003804070607900104

Goldrick-Rab, S., & Pfeffer, F. T. (2009). Beyond access: Explaining socioeconomic differences in college transfer. *Sociology of Education*, *82*, 101–125. doi:10.1177/003804070908200201

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*, 504–528. doi:10.1016/S0092-6566(03)00046-1

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence*, *24*, 13–23. doi:10.1016/S0160-2896(97)90011-8

Grant, A. M. (December 8, 2018). What straight-A students get wrong. *The New York Times*. Retrieved from https://www.nytimes.com/2018/12/08/opinion/college-gpa-career-success.html

Harker, L., & Keltner, D. (2001). Expressions of positive emotion in women's college yearbook pictures and their relationship to personality and life outcomes across adulthood. *Journal of Personality and Social Psychology*, *80*, 112–124. doi:10.1037/0022-3514.80.1.112

Heckman, J., & Kautz, T. (2013). Achievement tests and the role of character in American life. In J. Heckman, J. E. Humphries, & T. Kautz (Eds.), *The myth of achievement tests: The GED and the role of character in American life* (pp. 1–71). Chicago, IL: University of Chicago Press.

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science*, *4*, 162–176. doi:10.1111/j.1745-6924.2009.01116.x

Hoyle, R. H., Davisson, E. K., Daniel, K. E., & Komoski, S. E. (2018). *A model-based research synthesis of the grit literature*. Manuscript in preparation.

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. doi:10.1080/10705519909540118

Hurwitz, M., & Lee, J. (2018). Grade inflation and the role of standardized testing. In J. Buckley, L. Letukas, & B. Wildavsky (Eds.), *Measuring success: Testing, grades, and the future of college admissions* (pp. 64–93). Baltimore, MD: Johns Hopkins University Press.

Hutt, S., Gardner, M., Kamentz, D., Duckworth, A. L., & D'Mello, S. K. (2018, March). *Prospectively predicting 4-year college graduation from student applications*. Paper presented at the 8th International Learning Analytics and Knowledge Conference, New York, NY.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non-test score outcomes. *Journal of Political Economy*, *126*, 2072–2107. doi:10.1086/699018

James, W. (1899). *Talks to teachers on psychology: And to students on some of life's ideals*. New York, NY: Henry Holt.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger/Greenwood.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test (KBIT)*. Circle Pines, MN: American Guidance Service.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average*. Retrieved from https://research.collegeboard.org/sites/default/files/publicat ions/2012/7/researchreport-2008-5-validity-sat-predicting-first-year-college-grade-point-average.pdf

Koenig, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence*, *36*, 153–160. doi:10.1016/j.intell.2007.03.005

Kraft, M. A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, *54*, 1–36. doi:10.3368/jhr.54.1.0916.8265R3

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science*, *315*, 1080–1081. doi:10.1126/science.1136618

Kuncel, N. R., Kochevar, R. J., & Ones, D. S. (2014). A meta-analysis of letters of recommendation in college and graduate admissions: Reasons for hope. *International Journal of Selection and Assessment*, *22*, 101–107. doi:10.1111/ijsa.12060

Kuncel, N. R., Ones, D. S., & Sackett, P. R. (2010). Individual differences as predictors of work, educational, and broad life outcomes. *Personality and Individual Differences*, *49*, 331–336.

Laertius, D. (1925). *Lives of eminent philosophers* (R. D. Hicks, Trans. Vol. 1). Cambridge, MA: Harvard University Press.

Lemann, N. (1999). *The big test: The secret history of the American meritocracy*. New York, NY: Farrar, Straus & Giroux.

Lindquist, E. F. (1958, November). *The nature of the problem of improving scholarship and college entrance examinations*. Paper presented at the Invitational Conference on Testing Problems, Princeton, NJ.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "'General Intelligence,' Objectively Determined and Measured". *Journal of Personality and Social Psychology*, *86*, 96–111. doi:10.1037/0022-3514.86.1.96

Marcus, J. (August 16, 2017). Why suburban schools are inflating kids' grades. *The Hechinger Report*. Retrieved from https://hechingerreport.org/newest-advantage-rich-america-higher-grades/

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*, 933–938. doi:10.1126/science.2658056

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*, 2693–2698. doi:10.1073/pnas.1010076108

Morse, R., & Brooks, E. (September 9, 2018). Best colleges ranking criteria and weights. *U.S. News and World Report*. Retrieved from https://www.usnews.com/education/best-colleges/articles/ranking-criteria-and-weights

Muenks, K., Wigfield, A., Yang, J. S., & O'Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology*, *109*, 599–650. doi:10.1037/edu0000153

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

National Association for College Admission Counseling. (2008). *Report of the commission on the use of standardized tests in undergraduate admission*. Retrieved from https://files.eric.ed.gov/fulltext/ED502721.pdf

National Association for College Admission Counseling. (2017). *State of college admission*. Retrieved from https://www.nacacnet.org/globalassets/documents/publications/research/soca17final.pdf

National Center for Education Statistics. (2009). *ACT score averages and standard deviations, by sex and race/ethnicity, and percentage of ACT test takers, by* *selected composite score ranges and planned fields of study: Selected years, 1995 through 2009*. Retrieved from https://nces.ed.gov/programs/digest/d09/tables/dt09_147.asp

National Center for Education Statistics. (2011). *Percentage distribution of first-time postsecondary students starting at 2- and 4-year institutions during the 2003-04 academic year, by highest degree attained, enrollment status, and selected characteristics: Spring 2009*. Retrieved from https://nces.ed.gov/programs/digest/d14/tables/dt14_326.40.asp

National Center for Education Statistics. (2014a). *Enrollment and percentage distribution of enrollment in public elementary and secondary schools, by race/ethnicity and level of education: Fall 1998 through fall 2023*. Retrieved from http://nces.ed.gov/programs/digest/d13/tables/dt13_203.60.asp

National Center for Education Statistics. (2014b). *Graduation rates of first-time, full-time bachelor's degree-seeking students at 4-year postsecondary institutions, by race/ethnicity, time to completion, sex, and control of institution: Selected cohort entry years, 1996 through 2006*. Retrieved from http://nces.ed.gov/programs/digest/d13/tables/dt13_326.10.asp

National Center for Education Statistics. (2015). *Number and percentage distribution of spring 2002 high school sophomores, by highest level of education completed, and socioeconomic status and selected student characteristics while in high school: 2012*. Retrieved from https://nces.ed.gov/programs/digest/d14/tables/dt14_104.91.asp

National Center for Education Statistics. (2017a). *Common Core of Data (CCD): Public school data 2015-16, 2016-17 school years*. Retrieved from https://nces.ed.gov/ccd/schoolsearch/

National Center for Education Statistics. (2017b). *Graduation rate from first institution attended for first-time, full-time bachelor's degree-seeking students at 4-year postsecondary institutions, by race/ethnicity, time to completion, sex, control of institution, and acceptance rate: Selected cohort entry years, 1996 through 2010*. Retrieved from https://nces.ed.gov/programs/digest/d17/tables/dt17_326.10.asp

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101. doi:10.1037/0003-066X.51.2.77

Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2018). Admission testing for higher education: A multi-cohort study on the validity of high-fidelity curriculum-sampling tests. *PLoS One*, *13*(6), e0198746. doi:10.1371/journal.pone.0198746

Oettingen, G. (2014). *Rethinking positive thinking: Inside the new science of motivation*. New York, NY: Current.

Park, D., Tsukayama, E., Goodwin, G. P., Patrick, S., & Duckworth, A. L. (2017). A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual competencies in children. *Contemporary Educational Psychology*, *48*, 16–27. doi:10.1016/j.cedpsych.2016.08.001

Park, D., Yu, A., Metz, S. E., Tsukayama, E., Crum, A. J., & Duckworth, A. L. (2018). Beliefs about stress attenuate the relation among adverse life events, perceived distress, and self-control. *Child Development*, *89*, 2059–2069. doi:10.1111/cdev.12946

Parsons, T. (1959). The school class as a social system: Some of its functions in American society. *Harvard Educational Review*, *29*, 297–318.

Pattison, E., Grodsky, E., & Muller, C. (2013). Is the sky falling? Grade inflation and the signaling power of grades. *Educational Researcher*, *42*, 259–265. doi:10.3102/0013189X13481382

Pellegrino, J. W., & Hilton, M. L. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington DC: National Academy of Sciences.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*, 322–338. doi:10.1037/a0014996

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*, 203–212. doi:10.1016/j.jrp.2006.02.001

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales. Section 5: The Mill Hill vocabulary scale*. San Antonio, TX: Harcourt Assessment.

Robbins, S. B., Allen, J., Casillas, A., Peterson, C. H., & Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management measures from traditional predictors of college outcomes. *Journal of Educational Psychology*, *98*, 598–616. doi:10.1037/0022-0663.98.3.598

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*, 261–288. doi:10.1037/0033-2909.130.2.261

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, *58*, 103–139. doi:10.1111/j.1744-6570.2005.00301.x

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*, 313–345. doi:10.1111/j.1745-6916.2007.00047.x

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin*, *94*, 18–38. doi:10.1037/0033-2909.94.1.18

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin*, *135*, 1–22. doi:10.1037/a0013978

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177. doi:10.1037//1082-989X.7.2.147

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Routledge.

Sorensen, L. C., Dodge, K. A., & Conduct Problems Prevention Research, G. (2016). How does the Fast Track intervention prevent adverse outcomes in young adulthood? *Child Development*, *87*, 429–445. doi:10.1111/cdev.12467

Thorsen, C., & Cliffordson, C. (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades. *Educational Research and Evaluation*, *18*, 153–172. doi:10.1080/13803611.2012.659929

Troutman, K. P., & Dufur, M. J. (2007). From high school jocks to college grads: Assessing the long-term effects of high school sport participation on females' educational attainment. *Youth & Society*, *38*, 443–462. doi:10.1177/0044118x06290651

Tsukayama, E., Duckworth, A. L., & Kim, B. (2013). Domain-specific impulsivity in school-age children. *Developmental Science*, *16*, 879–893. doi:10.1111/desc.12067

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*, 498–505. doi:10.3758/BF03192720

Valiente, C., Lemery-Chalfant, K., Swanson, J., & Reiser, M. (2008). Prediction of children's academic competence from their effortful control, relationships, and classroom participation. *Journal of Educational Psychology*, *100*, 67–77. doi:10.1037/0022-0663.100.1.67

Walters, A. M., Kyllonen, P., & Plante, J. W. (2006). Developing a standardized letter of recommendation. *Journal of College Admission*, *191*, 8–17.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children—Fourth edition*. San Antonio, TX: Psychological Corporation.

West, M. R., Kraft, M. A., Finn, A. S., Martin, A. R., Duckworth, A. L., Gabrieli, C. F. O., & Gabrieli, J. D. E. (2016). Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis*, *38*, 148–170. doi:10.3102/0162373715597298

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. R., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT scores, high school grades, and SES. *Educational Assessment*, *20*, 23–45. doi:10.1080/10627197.2015.997614

Willingham, W. W. (1985). *Success in college: The role of personal qualities and academic ability*. New York, NY: College Entrance Examination Board.

Willingham, W. W. (2005). Prospects for improving grades for use in admissions. In W. Camara & E. W. Kimmel (Eds.), *Choosing students: Higher education admissions tools for the 21st century* (pp. 127–140). New York, NY: Routledge.

Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement*, *39*, 1–37. doi:10.1111/j.1745-3984.2002.tb01133.x

Wong, M. M., & Csikszentmihalyi, M. (1991). Motivation and academic achievement: The effects of personality traits and the quality of experience. *Journal of Personality*, *59*, 539–574. doi:10.1111/j.1467-6494.1991.tb00259.x

Yeager, D. S., Dahl, R. E., & Dweck, C. S. (2018). Why interventions to influence adolescent behavior often fail but could succeed. *Perspectives on Psychological Science*, *13*, 101–122. doi:10.1177/1745691617722620

Yeager, D. S., Henderson, M., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology*, *107*, 559–580. doi:10.1037/a0037637

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*, 3–17. doi:10.1207/s15326985ep2501_2

Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology*, *39*, 145–155. doi:10.1016/j.cedpsych.2014.03.004

Manuscript received December 29, 2016
Final revision received February 8, 2019
Accepted February 26, 2019