Trait Complex, Cognitive Ability, and Domain Knowledge Predictors of Baccalaureate
Success, STEM Persistence, and Gender Differences

Phillip L. Ackerman & Ruth Kanfer

Georgia Institute of Technology

Margaret E. Beier

Rice University

Send correspondence to: Phillip L. Ackerman
School of Psychology
Georgia Institute of Technology
654 Cherry Street, MC 0170
Atlanta, GA 30332-0170, USA
FAX: 404-252-3061 or 404-894-6904
e-mail: phillip.ackerman@psych.gatech.edu

Running Head: Predicting STEM Persistence & Academic Achievement

Submitted to: *Journal of Educational Psychology*

Original Submission: May 12, 2012

Revision submitted: October 10, 2012

Second Revision submitted: February 19, 2013

Abstract

Prediction of academic success at post-secondary institutions is an enduring issue for educational psychology. Traditional measures of high-school grade point average and high-stakes entrance examinations are valid predictors, especially of first-year college grades, yet a large amount of individual-differences variance remains unaccounted for. Studies of individual trait measures (e.g., personality, self-concept, motivation) have supported the potential for broad predictors of academic success, but integration across these approaches has been challenging. The current study tracks 589 undergraduates from their first semester through attrition or graduation (up to 8 years beyond their first semester). Based on an integrative trait-complex approach to assessment of cognitive, affective, and conative traits, patterns of facilitative and impeding roles in predicting academic success were predicted. We report on the validity of these broad trait complexes for predicting academic success (grades and attrition rates) in isolation, and in the context of traditional predictors and indicators of domain knowledge (Advanced Placement® exams). We also examine gender differences and trait complex by gender interactions for predicting college success and persistence in STEM fields. Inclusion of trait-complex composite scores and average AP exam scores raised the prediction variance accounted for in college grades to 37%, a marked improvement over traditional prediction measures. Math/Science Self-Concept and Mastery/Organization trait complex profiles were also found to differ between men and women who had initial STEM major intentions, but who left STEM for non-STEM majors. Implications for improving selection and identification of students at-risk for attrition are discussed.

Keywords: Traits, STEM, gender, GPA, attrition

According to recent statistics (National Center for Educational Statistics, 2012), the overall attrition rate in Title IV institutions is 28.1% (Title IV institutions are those that are allowed to participate in Pell Grants and other federal financial assistance programs). That is, more than 1 in 4 first-time degree/certificate seeking undergraduate students fail to complete a degree within 6 years. At public 4-year institutions, the overall attrition rate is lower (21.4%), but still represents a major concern to institutions and stakeholders (e.g., students, parents, state and federal policy-makers). Although the situation may not be as stark as the statistics indicate because reported attrition rates include transfers from one institution to another, the loss of students prior to completing a degree represents a failure in either the selection of students, or a failure to identify students and to provide interventions for students who are at-risk for attrition. It might be pointed out, however, that attritions do also include students who realize that they are not well-matched to an institution, or who find that they are not ultimately interested in completing a degree. Whether these latter situations can be regarded as selection failures is an open question. In addition to the overall attrition issues for colleges and universities, there has recently been substantial discussion in the domains of educational research and public policy about the difficulties in both attracting and retaining students in Science, Technology, Engineering, and Math (STEM) majors (e.g., see Ehrenberg, 2010, President's Council of Advisors on Science and Technology, 2012; Sullivan, 2006, Xie & Killewald, 2012).

This study examines the role of trait complexes and domain knowledge as they influence the selection and retention of talented students and account for gender differences in STEM and non-STEM majors in a selective undergraduate institution. From a signal detection theory perspective, the goal in improving selection is to reduce 'false alarms' -- defined as those

students who leave prior to attaining a degree, and to increase ‘hits’ -- defined as those who complete a degree within a defined time period (e.g., four to five years from matriculation). Once students have matriculated, the goal shifts to identifying students who are ‘at risk’ for attrition, identifying the reasons why they are at risk, and designing interventions to address the problems that are amenable to remediation. At Georgia Institute of Technology (Georgia Tech) and similar institutions with high rates of STEM-oriented students, an additional concern relates to the identification of students who are at-risk for leaving STEM majors prior to degree completion. Furthermore, interest in understanding whether the patterns of STEM-persistence differ for women and men is a focus of much discussion in educational research (Ceci et al., 2009).

At highly selective institutions such as the Georgia Tech, where this study was conducted, concern about student retention rates loom large. For example, in the current admissions period (Fall, 2012), 14,700 applications were received for 2,400 spots (Georgia Institute of Technology, 2012). Despite the use of high school grade point average (GPA), and SAT scores along with other aspects of the application files to predict college performance, only 82% of the freshman students who have matriculated over the past decade persisted until completion of the baccalaureate degree (Georgia Institute of Technology, 2009). Moreover, in a recently completed archival study of students who matriculated to Georgia Tech over a 6-year period (1999-2004), 20.1% of students at Georgia Tech who had an initial STEM¹ major did not complete a degree within 6 years. Of those who did complete a degree, 15.5% changed majors from a STEM field to a non-STEM field (Ackerman, Kanfer, & Calderwood, in press). Clearly, there is room for improvement when it comes to predicting academic success, attrition, and

STEM persistence, which provided the impetus for the current study. As described below, the approach adopted in this study was to use trait complexes as an organizing framework -- ability, domain knowledge and non-ability student characteristics that would, in turn as hypothesized, provide stronger correspondence between predictors and criteria. We then evaluated the impeding and facilitating influences of these complexes and gender interactions for the purpose of improving selection and retention of students at highly selective institutions like Georgia Tech.

Background

Numerous researchers have investigated various predictors of academic performance (e.g., GPA) and attrition in studies of post-secondary education over the past 100 or so years. Early studies and applications before and after World War I focused mainly on the use of aptitude and intelligence tests for predicting success in college/university study (for a review see Ackerman, 1996; see also E. L. Thorndike, 1920). Since the 1920s, some investigators have suggested that aptitude tests should focus on at most two different types of abilities (verbal and mathematical/quantitative) for prediction of college/university academic success (e.g., the early versions of the College Entrance Examination Board tests, which was the precursor of the SAT, see Carroll, 1982). Other investigators have taken a more differential approach, which focuses either on the profiles of student aptitudes/abilities (e.g., Thurstone's Primary Mental Abilities approach, see Bernreuter & Goodman, 1941), or on including other salient abilities/aptitudes/achievements into the prediction equation (e.g., the ACT, see Cole, 1969). In the STEM domain, there has been significant discussion related to the consideration of spatial abilities, including constructs of spatial visualization, but also including measures of mechanical

knowledge/mechanical reasoning (e.g., see Wai, Lubinski, & Benbow, 2009; Webb, Lubinski, & Benbow, 2007). Although there have been several archival and exploratory demonstrations of high correlations between spatial ability measures and post-secondary success, spatial ability measures have not been used in a high-stakes testing environment for educational selection.²

Over the course of the last century, other research efforts were made to incorporate measures of vocational interests, personality traits, self-concept, motivational traits and skills, and numerous other psychological measures for prediction of academic success (for reviews, see Richardson, Abraham, & Bond, 2012; Stein, 1963). On the one hand, because most of these non-ability measures are assessed with self-report instruments, and thus are subject to distorted responding to varying degrees, they tend not to be ideal for use in post-secondary admissions (Ellingson & McFarland, 2011). On the other hand, if such measures have validity for predicting important academic outcomes, they might be used post-matriculation for improving ‘classification.’ That is, such measures could help guide students to select the academic majors for which they are best suited (see French, 1955/1966). They could also be used to identify at-risk students for potential interventions aimed at improving academic performance and retention (e.g., Kahn et al., 2002). Despite the long history of research in this field, there has been a lack of demonstrable progress in the development of robust psychological instruments for the prediction of academic success at the post-secondary level.

Although models have been developed and evaluated to predict post-secondary student attrition (e.g., Tinto, 1975), they tend to focus on the entire “system” involved in undergraduate education (e.g., family background, peer and faculty interactions, institutional commitment). However, *individual attributes* represent a small part of these systems, and attributes other than

intellectual ability have merely been suggested as potential influences.

Recently, three meta-analyses have been conducted that attempt to further explore the role of individual attributes in predicting academic performance (Poropat, 2009; Richardson et al., 2012; Robbins et al., 2004). Robbins et al. examined the role of psychosocial and study skills in predicting both academic performance and retention. Their meta-analysis included a variety of trait measures (e.g., achievement motivation, general self-concept), and determined that these variables accounted for significant variance in both academic performance and retention. Specifically, they found mean observed correlations for Achievement Motivation of $r = .257$ for GPA and $r = .105$ for retention, and the mean observed correlations for General Self-Concept of $r = .037$ and $.059$, for GPA and retention, respectively. Other than “academic self-efficacy” and “academic-related skills,” individual correlations with trait measures and academic performance/retention criteria were relatively modest in magnitude. In the context of traditional predictors of academic success (e.g., SAT, high school GPA), Robbins et al. (2004) found that Achievement Motivation accounted for only an additional 1-2% of variance in college GPA, and including a direct measure of student self-efficacy for academic performance improved the amount of variance accounted for over traditional measures only minimally (from $R^2 = .219$ to $R^2 = .262$, Robbins et al., 2004, Table 11). Note, though, that these estimates were based on correlations corrected for measurement error, and thus represent estimated true-score correlations; they are not the results that could be expected in an operational environment.

The Poropat (2009) meta-analysis examined the relationship between five personality factors and GPA across all levels of education (from primary to post-secondary). He found that three personality traits had small effects (i.e., d value greater than .10, Cohen, 1988) for

predicting academic performance at the post-secondary level: Conscientiousness ($\rho = .23$), Agreeableness ($\rho = .06$), and Openness ($\rho = .07$). These correlations are corrected in a manner similar to that used by Robbins et al. (2004), and thus do not represent correlations found in operational environments. Together, results from both of these studies suggest that personality, self-concept, and related trait variables have potential for predicting post-secondary academic performance and retention/attrition. To date, though, the magnitude of the demonstrated effects is not large, especially in the context of traditional measures of prior academic achievement (high school GPA) and ability (SAT/ACT test scores).

Richardson et al. (2012) provided an updated meta-analysis (based on studies published between 1997 and 2010) that included several of the trait constructs from the previous meta-analyses, but also focused on more specific constructs, such as academic self-efficacy, grade goals, and effort regulation. In general, they found relatively modest correlations between trait measures and academic performance criteria, and larger, but still moderate sized correlations with the specific academic variables of self-efficacy and grade goals. For example, they found that academic self-efficacy, goals for course grades, and locus of control together accounted for 14% of the variance in university grades. Richardson et al. estimated approximately 20% of the variance in university GPA was accounted for by personality and self-efficacy/goals. The largest correlations were found between measures of self-efficacy and GPA.

Brunswik Symmetry

Wittmann and Süß (1999) have argued that one of the main reasons that individual trait measures provide modest predictive validities for academic and employment performance outcomes pertains to the lack of a correspondence between the predictor space and the criterion

space. In particular, Wittmann and Süß proposed that optimal validities can be found when the breadth and the direct correspondence of predictor and criterion measures are matched (termed Brunswik Symmetry). When predictors are broader than the criterion, they will often account for significant variance, but will also have non-overlapping variance. When the predictors are narrower than the criterion, there may be no match between predictors and the criterion (i.e., no validity), or a limited match (i.e., they account for a relatively small portion of criterion variance). Matching the appropriate levels of breadth and correspondence between predictors and criteria is a desirable goal for developing theory and maximizing the utility of instruments developed for selection. In the context of Brunswik Symmetry, one hypothesis for the relatively modest meta-analytically derived correlations between non-ability trait measures and measures of academic performance/retention (Poropat, 2009; Robbins et al., 2004) is that the studies that underlie the meta-analyses focus on trait measures that are narrower in scope than the criteria of post-secondary GPA and retention (Ackerman & Kanfer, 2005).

The Trait-Complex Approach

The trait-complex approach (Ackerman & Heggestad, 1997; Snow, 1963) offers an alternate view to the use of relatively narrow individual traits for predicting academic performance criteria. The origins of the approach were articulated by Snow (1963) in a study of learner characteristics in the context of post-secondary learning from instructional films. Snow suggested the concept of “aptitude complexes” as representing “combinations of levels of some variables which are particularly appropriate or inappropriate for efficient learning” (Snow, 1963, p. 120). For Snow and Cronbach (see Cronbach & Snow, 1977) an “aptitude” referred to any trait construct, not just those in the cognitive ability domain. The underlying idea was that

constellations of traits that were facilitative or impeding of learning should be considered when predicting individual differences in academic performance criteria.

Ackerman (1996) extended Snow's notion of aptitude complexes by proposing the PPIK (intelligence-as-Process, Personality, Interests, and intelligence-as-Knowledge) framework for understanding individual differences in domain knowledge as a function of intelligence. PPIK is an "investment" framework that proposed that the direction and intensity of intelligence-as-process investments are a function of affective (personality), and conative (interests, motivation) trait complexes that influence intellectual development, resulting in individual differences in the depth and breadth of domain knowledge. Although there was little research that directly addressed the aptitude complex idea in subsequent decades, a meta-analysis of intelligence, personality, and interest trait relations by Ackerman and Heggestad (1997) found common elements among these domains (e.g., ability, personality, interests). The constellations of traits were labeled "trait complexes" after Snow's "aptitude complexes" to better describe that they included not only cognitive variables, but also variables in the affect and motivation/volition domains. Ackerman and Heggestad (1997) further proposed that some of these trait complexes (Science/Math and Intellectual/Cultural) would be synergistically facilitative of learning, and others (Clerical/Conventional and Social) would be either uncorrelated with learning, or might be impeding of learning (see also Ackerman, 2003a; Ackerman, Chamorro-Premuzic, & Furnham, 2011).

Subsequent empirical studies have supported the overarching framework of trait complexes across trait domains (e.g., Armstrong, Day, McVay, & Rounds, 2008; Staggs, Larson, & Borgen, 2007; Sullivan & Hansen, 2004). Moreover, the facilitative and impeding nature of

some trait complexes for determining individual differences in the breadth and depth of domain knowledge and for predicting post-secondary academic achievement has also been demonstrated. Ackerman et al. (2001) found two facilitative trait complexes (Science/Math Technology and Verbal/Intellectual) and three impeding trait complexes (Social Potency/Enterprising, Social Closeness/Femininity and Traditionalism/Worry/Emotionality) that were associated with both individual and gender differences in patterns of domain knowledge and academic performance (GPA) in a study of first-year college students. In addition, gender differences in trait complexes were associated with gender differences in domain knowledge. Other studies have examined similar trait complexes in adult learners (e.g., Ackerman & Beier, 2006) and trait-complex influences on test performance (e.g., Ackerman & Kanfer, 2009). In general, it appears that the breadth of the predictor space encompassed by the trait complex approach and the underlying synergy of the underlying traits together yield higher validities for academic performance variables, in comparison to approaches that consider individual trait measures in isolation.

STEM Persistence and Gender Differences

In recent years, especially in the U.S., educators and public policy makers have focused increased attention on selection and retention of students in the STEM fields (e.g., see Baskin, 2012). A report from the National Center for Educational Statistics (NCES; 2009) indicated that for the 1995-96 cohort, 22.8% of entering undergraduates had an initial STEM major intention (32.9% male, 14.5% female). In contrast, the 2003-04 cohort had only a 13.7% rate of initial STEM major intentions.³ Although these statistics do not take into account the overall increase in US post-secondary institution enrollment, they do suggest a decline in the number of STEM enrollments. Further evidence for this decline is provided by analysis of total enrollment in

degree-granting institutions by NCES (National Center for Education Statistics, 2012) that show a decline of a half-million students with initial STEM majors between 1995 (1.9 million students) and the 2003-2004 cohort (1.4 million students).

Recent investigations of this decline have often focused on gender differences and similarities in terms of abilities, personality traits, vocational interests, self-concept, and related traits (e.g., see Lippa, 2005). Nonetheless, there remains controversy about why a lower proportion of women than men pursue education and vocations in STEM fields. A recent review by Ceci, Williams, and Barnett (2009) integrated findings from a variety of sources to identify the major factors that explain gender differences in STEM professions. An extended review of the consideration of biological and sociocultural factors they examined is beyond the scope of this review. However, Ceci et al. identified “motivation/attitudes/interests and activities” as the primary source of influence with “performance on gatekeeper tests” as a secondary source. The literature they consider on gatekeeper tests is limited primarily to tests of general and broad content abilities (i.e., math and verbal abilities), and thus does not consider spatial abilities and mechanical knowledge, two domains that have been shown to have substantial gender differences favoring males (see Voyer, Voyer, & Bryden, 1995) and marked correlations with success in STEM educational pursuits (e.g., see Wai et al., 2009).

Although informative, the Ceci et al. (2009) review does not consider a critical issue of growing importance in academic success; namely gender differences in Advanced Placement (AP) STEM courses that affect both admissions and the coursework taken in college (e.g., see Stumpf & Stanley, 1997). Specifically, the issue concerns both gender differences in the proportions of students taking AP STEM courses and the magnitude of gender differences in

performance on domain-knowledge AP tests, especially in the STEM areas. Such gender differences are largely hidden in large-scale assessments of course-taking, such as the National Assessment of Educational Progress (NAEP), which do not take account of AP enrollment, *per se*. For example, data from NAEP (1999) suggest that course taking patterns are not responsible for mean gender differences on science and math achievement tests, because girls and boys now take an equal *number* of STEM courses in high school. The number of courses may not, however, be as important as course content. Ma and Johnson (2008), for example, proposed that coursework in calculus during high school “is a powerful career filter that critically screens females for prestigious occupations” (p. 75). Those students who do not take calculus in high school are more likely to be female (e.g., see College Board, 2011), and appear to have a significant disadvantage in majoring in STEM areas.

Domain Knowledge

The most significant source of influence of university study lies with development of subject-matter knowledge (Astin, 1993). In their review of educational research, Pascarella and Terenzini (1991) estimated that freshman-to-senior gains in verbal skills averaged about .56 *sd* units, general quantitative skills averaged .24 *sd* units, but specific subject matter knowledge increased .87 *sd* units. Given the importance of transfer-of-knowledge in determining the acquisition of new knowledge, there is a basis for emphasizing the foundation of prior knowledge and skills at entry to a university major.

AP exams provide an increasingly common indicator of domain knowledge and skills prior to university matriculation. Over the past decade, the number of college-bound high school students completing AP exams has exploded. In 2002, 1.5 million exams were completed

(DiYanni, 2009), a number that more than doubled in 2010 (3.2 million exams, completed by 1.8 million students; College Board, 2011). Focusing on domain knowledge in university study and beyond has been fruitful in terms of predicting both individual differences and gender differences in academic achievement (Ackerman, Bowen, Beier, & Kanfer, 2001; Rolfhus & Ackerman, 1999). Empirical research has also supported the proposition that gender differences in domain knowledge are especially important for success in university-level achievement in STEM areas. For example, Taasobshirazi and Carr (2008) identified gender differences in physics expertise and expertise in other STEM areas at matriculation as a major reason why more women than men “derail” from such majors during college.⁴

In support of the importance of domain knowledge, Average AP exam score (as a measure of domain knowledge) has been shown to have substantial predictive validity for post-secondary academic success (in terms of grades and graduation rate; Ackerman et al., 2001, Ackerman et al., in press; Dwyer, 2011; Geiser & Santelices, 2004). Whether the student successfully completed three or more STEM-related AP tests (defined as scores of 4 or 5 on the AP tests), has also been indicated as a significant predictor of STEM persistence (Ackerman et al., in press). In sum, domain knowledge is a potentially important contributing factor in predicting success and retention in STEM fields, and gender differences in domain knowledge at matriculation likely lead to gender differences in STEM success.

Overview of Current Study

Two enduring issues that confront selective undergraduate institutions are improving the selection and improving the retention of talented students. From a signal detection theory perspective, the goal is to reduce ‘false alarms’ -- defined as those students who leave prior to

attaining a degree, and to increase ‘hits’ -- defined as those who complete a degree within a defined time period (e.g., four to five years from matriculation). Once students have matriculated, the goal becomes identifying students who are ‘at risk’ for attrition, identifying the reasons why they are at risk, and designing interventions to address the problems that are amenable to remediation. At Georgia Tech and similar institutions with high rates of STEM-oriented students, an additional concern relates to the identification of students who are at-risk for leaving STEM majors prior to degree completion. Furthermore, interest in understanding whether the patterns of STEM-persistence differ for women and men is a focus of much discussion in educational research (Ceci et al., 2009).

Georgia Tech provides an especially interesting institution for the study of STEM persistence because it attracts a substantial majority of its applicants to STEM majors, especially in Engineering, which results in a high percentage of students who declare an initial STEM major (87.8% in 2002-03). In the Fall 2002 entering class of 2,274 first-time college students, women made up 28% of cohort group. In addition, over the course the undergraduate program, a significant percentage of initial STEM-major students (15.1%) switch out of STEM and complete a degree in another subject, which represents a significant sample for the study of STEM-persisters and STEM-leavers.

The current study investigates the selection and retention of STEM students and gender differences in STEM persistence. In the Fall of 2002, first year students enrolled in a college life skills (“freshman experience”) course for incoming students at Georgia Tech were assessed using the trait-complex perspective to determine whether a small number of broad “facilitative” or “impeding” families of traits could be used to enhance the prediction of academic success and

STEM persistence.

To further evaluate the role of individual and gender differences in domain knowledge as predictors of academic achievement and STEM-persistence, we obtained assessments of domain knowledge measured in high school, as indexed by AP exam scores, college transcripts, and college admissions records. This information, along with the trait complex assessment, was used to determine academic success and to evaluate the trait complex predictors of college achievement (e.g., GPA), degree completion, and persistence in STEM majors, both in isolation and in conjunction with traditional predictors of academic success (i.e., high school GPA and SAT scores). We had three major hypotheses for this investigation, as follows:

Hypothesis 1: Trait complexes would represent independently significant and incrementally significant predictors of academic achievement (beyond traditional predictors) and STEM persistence

Hypothesis 2: Individual differences in AP exam domain knowledge indicators would also be independent and incremental predictors of academic achievement and STEM persistence; and

Hypothesis 3: Gender differences in trait complexes and AP exam patterns would be associated with STEM persistence.

Method

Sample

Psychology 1000 is a one-credit elective course at Georgia Tech for freshmen undergraduate students (renamed GT 1000 in 2005). Course topics are designed to aid the students' transition to college, and include for example, "time management, learning skills,

career planning, psychological hardiness, teamwork, and leadership” (Hagearty, 2003, p. 2). The class meets in small groups of between 15 and 25 students per section led by Georgia Tech staff serving as volunteer instructors. The Fall 2002 cohort of first-year freshmen (transfer students were excluded from the analysis) consisted of 2,274 students. There were 1,196 students (52.6%) from this cohort enrolled in Psychology 1000. Questionnaires were distributed to approximately 1,100 of the students and responses were obtained from 592 students, a 54% response rate. Data from three students were excluded because they failed to follow instructions. The final sample consisted of $N = 589$ students.

Descriptive statistics and comparisons between the sample and the population of first-year students at Georgia Tech are provided in **Table 1**. As can be seen in the table, the sample was representative of the population.⁵ With the large samples in each group, the power to detect small differences is high, and significant differences were found for the two groups on high school GPA and Cumulative GPA at Georgia Tech, both favoring the sample over the population. In both cases, the effect sizes of the differences are less than small in magnitude (i.e., small effects are between $d = .20$ and $.49$, Cohen, 1988). Significant, but small differences were also found for the percentage of men participating in the study compared to the population (66.9% sample vs. 72.0% population), and for the percentage of initial STEM majors in the study (85.7% sample vs. 87.8% population).

Self-Report Measures

Five broad trait complexes were assessed with a battery of personality, motivation, self-regulation, self-concept, self-estimates of ability and related measures. Trait complexes were: science/math, mastery/approach-achievement motivation, verbal/intellectual, avoidance in

achievement contexts, and social/extroversion. The first three have been identified as facilitative -- being positively related to academic performance and acquisition of domain knowledge, and the latter two have been identified as impeding -- being negatively related to both academic performance and the acquisition of domain knowledge (Ackerman, 2003a, 2003b; Ackerman & Beier, 2006; Ackerman et al., 2001; Rolfhus & Ackerman, 1999). A list of scales, the number of items in each scale, along with means, standard deviations, and internal consistency estimates (Cronbach, 1951) are provided in **Table 2**. The measures administered were as follows:

Motivational Traits. Four scales from the Motivational Trait Questionnaire short form (Kanfer & Ackerman, 2000). Included were two approach-oriented scales (Desire to Learn and Mastery) and two avoidance-oriented scales (Worry and Emotionality in achievement contexts).

Personality. Three scales from the NEO-Five-Factor Inventory (FFI; Costa & McCrae, 1992) and two scales from the Multidimensional Personality Questionnaire (MPQ; Tellegen, 1982). Included were scales of Extroversion, Openness to Experience, Conscientiousness from the NEO-FFI; Social Potency and Social Closeness from the MPQ.

Motivated Strategies for Learning Questionnaire (MSLQ). The MSLQ was designed to measure “motivation and use of learning strategies by college students” (Pintrich, Smith, Garcia, & McKeachie, 1993, p. 801). Included were nine scales from the MSLQ: Test Anxiety, Intrinsic Goal Orientation, Peer Learning, Metacognitive Self-Regulation, Time and Study Environmental Management; Effort Regulation, Critical Thinking, Organization, and Rehearsal.

Self-Concept. Four scales of different domains of academic self-concept (e.g., see Ackerman, Kanfer, & Goff, 1995; Ackerman & Wolman, 2007) were administered, including Verbal, Math, Spatial, and Science.

Self-Estimates of Abilities and Skills. Four scales of self-estimates were included. These scales ask the participants to rate their abilities in comparison to the population at-large (see Ackerman & Wolman, 2007 for an extensive discussion of these measures). Included were scales of Organizational Skills, Spatial and Science Abilities, Verbal Abilities and Math Abilities.

Numerical Preferences. This is a shortened version of Viswanathan's (1993) "Preference for Numerical Information" scale. Questions on this scale ask about the participant's comfort and preferences for using mathematics (e.g., "I enjoy work that requires the use of numbers").

Life Goals. Two scales of life goals were included, based on a measure developed by Roberts and Robins (2000). One subscale asked participants about their desire for obtaining a high status in their intended profession or community (e.g., "Having an influential and prestigious occupation"); the other scale asked participants about their desire to have a life relatively free of difficulties (e.g., "Avoiding hard work").

Additional scales were administered in the questionnaire packet, but are not discussed here for space considerations.

Cognitive Ability and Domain Knowledge Measures

Academic Records. Admissions and Georgia Tech transcript records were obtained from the Office of Institutional Research and Planning at Georgia Tech. The following key measures were obtained as predictor and criterion variables:

High School Grade Point Average. This cumulative GPA variable is based on grades obtained by the students, with bonuses provided for enrollment in honors and Advanced

Placement courses.

SAT Scores. The version of the SAT for the 2002 entering class only included Verbal and Quantitative (Math) sections. Thus, separate measures were obtained for Verbal and Math sections, respectively.

Advanced Placement® Exam Scores. The type of AP exam completed and the scores for each exam were obtained from institutional records. Two indicators of domain knowledge were derived from these scores: the average of all AP test scores completed by the student, and a classification of whether the student successfully completed three or more STEM AP tests (including Biology, Calculus, Chemistry, Computer Science, Environmental Science, and Physics).

Criterion Measures

Grade Point Average. Yearly (Years 1 - 4) and cumulative GPA were obtained from Georgia Tech transcripts.

Major. Both initial major intention (prior to or at matriculation) and final major (based on degree earned or terminal major, in the case of students who left Georgia Tech prior to graduation) were obtained from transcripts. The majors classified as STEM in this study included Physical and Biological Science, Technology, Engineering, and Math. To be consistent with NCES designations (e.g., NCES, 2009), Social Sciences (e.g., Psychology, Economics, Management, International Affairs) and Foreign Languages were categorized as non-STEM majors. STEM "persisters" were those with initial and final STEM majors, who completed a degree. STEM "leavers" were those with an initial STEM major intention, but who completed a non-STEM degree.

Attrition. Students who did not complete a degree at Georgia Tech by Fall, 2010 (eight years after matriculation) were identified as having attritted. Specific reasons for student attrition were not available, but include voluntary transfer to another post-secondary institution and involuntary attrition (e.g., due to insufficient GPA).

Procedure

Instructors of Psychology 1000 distributed consent forms, academic release forms, and the paper questionnaire packet to freshmen students during the first three weeks of their first semester at Georgia Tech (September, 2002). Students were told that completion of the questionnaires was optional and they were asked to complete the questionnaires on their own and return them within two weeks. Final academic records (including admissions records and complete Georgia Tech transcripts) were obtained in the Fall of 2010.

Results

The analysis begins with general descriptive statistics comparing attrition and STEM persistence factors in the sample to the population. The next stages of analysis are divided into four sections. The first stage pertains to the analysis of the trait measures and the derivation of five broad trait complex composites. In this section, we also describe gender differences in the trait complexes. In the second section, we evaluate the predictive validity of the trait complex and Average AP exam scores for GPA across the years of enrollment and overall cumulative GPA, in isolation, and in the context of testing incremental validity for these measures after traditional predictors (high school GPA and SAT scores) are entered into the regression equations. For the third section, we focus on the three sets of variables for prediction of STEM persistence, and we examine interactions among the trait complex variables and gender for

STEM persisters and STEM leavers. In the final section, we examine the predictors of graduation/attrition.

Descriptive Information

Of the 589 participants, 482 completed a Bachelors degree within the 8 years from questionnaire completion to the transcript download, yielding a completion rate of 81.8%. Although 503 students indicated an initial intention to major in a STEM area (85.4%), by graduation there were only 352 STEM majors (73% of those students who completed a Bachelors degree, including 5 students who completed a degree in a STEM area, but who declared an initial non-STEM intention). These are classified as “STEM persisters.” A total of 76 students who completed a degree but changed majors from STEM to a non-STEM domain were identified as “STEM leavers.” Of these, the vast majority of students changed their major to Management (75%), with the remaining to a variety of other non-STEM majors. For the 5 of 59 students who declared an initial non-STEM major and completed a STEM degree, they all transferred to Engineering and related majors.

Trait Complexes

The 29 trait measures were subjected to an exploratory factor analysis as the first step to derivation of trait complex estimates. Five factors were extracted using iterated communalities and principal factor extraction. The factor matrix was then rotated to an orthogonal varimax criterion. The five broad trait-complex factors are generally similar to those that were anticipated, though the scales that have the highest loadings on any particular factor tend to vary somewhat from one study to the next. Following from earlier studies (Ackerman et al., 2001), the trait complexes were named on the basis of the scales that had the highest factor loadings, as:

I. Math/Science Self-Concept; II. Mastery/Organization; III. Openness and Verbal Self-Concept; IV. Anxiety in Achievement Contexts; and V. Extroversion. The first three factors were hypothesized to represent facilitative trait complexes and the remaining two were hypothesized to represent impeding trait complexes (e.g., see Ackerman et al., 2001).

Unit-weighted z -score composites for the trait complexes were formed from the scales that had salient loadings on the respective factors (based on principles of aggregation as described by Cohen, 1988, and R. L. Thorndike, 1986). The component scales that comprise the trait complex composites are provided in **Table 3**. Given that the trait complexes represent relatively broad constructs, we did not expect them to have extremely high levels of homogeneity. Reliability estimates for each trait complex were calculated using the formula provided by Nunnally (1978) for calculating the reliability of a linear composite. Reliability estimates ranged from .89 to .92 indicating relatively high levels of homogeneity for the trait complex composites. Correlations among the trait complex measures are also shown in the table. The trait complex correlation matrix indicates positive manifold among the three ‘facilitative’ trait complexes: math/science self-concept, approach-oriented motivational traits, and openness to experience and verbal self-concept (I, II, and III). Also, there were negligible correlations among the ‘impeding’ trait complex of anxiety in achievement contexts and the trait complex of extroversion (IV and V). For comparison purposes, extroversion measures are generally found to be orthogonal to achievement measures, but they are sometimes observed to have small negative correlations with achievement (e.g., see Poropat, 2009).

Gender Differences in Trait Complexes. Because there are gender differences in many of the component scales that comprise the trait complexes, it was expected that gender

differences would also be manifest in trait complex composite scores. From the extant literature on similar selective college-student samples (Ackerman et al., 2001), we expected that men would have higher mean scores on the math/science self-concept trait complex, and that women would have higher mean scores on the others. Means, standard deviations, *t*-tests for the differences between means, and Cohen's *d* for the five trait complexes are shown in **Table 4**. Significant differences in the predicted direction were found for four of the five trait complexes. For the trait complexes with significant gender differences, only mastery/organization (II) reached a medium magnitude (larger than $d = .50$); the others were of small magnitudes ($d = .20$ to $.49$).

Prediction of Academic Grades

The traditional prediction of academic success at many selective colleges and universities is drawn largely on the basis of a regression equation that includes cumulative high-school GPA (at least cumulative to the first semester of the applicant's senior year of high school), and a high-stakes exam (typically SAT or ACT). Thus, in order to evaluate the influence of domain knowledge assessment (average AP exam scores) and trait complex variables, we computed separate regressions for these later predictors in isolation, and then in conjunction with the traditional predictor measures to determine whether the AP average and trait complex scores provide incremental predictive validity for grades across the years of enrollment at Georgia Tech. It should be noted however, that similar to other selective institutions, there is a substantial restriction of range in talent among the students who matriculated at Georgia Tech, (i.e., mean high school GPA was 3.85 and mean SAT score, Verbal + Quantitative, was 1323 for this sample). Although it is possible to 'correct' the correlations between the predictors and the

criteria, on the basis of explicit selection to get an estimate of the true-score validity of the predictors, we instead report raw multiple correlations because these provide an accurate indicator of the actual utility of including the measures in an operational educational environment.

The results of the separate and hierarchical regressions of the predictors and cumulative GPA at Georgia Tech are presented in **Table 5**. The traditional predictors (high school GPA and SAT) account for 23% of the variance in first-year grades. As with most studies predicting academic success, there were declining validities for later years (e.g., see Beier & Ackerman, 2012; Humphreys, 1968). In isolation, a single average AP exam score provides a similar degree of validity for predicting GPA, accounting for 21% of the variance in first-year grades. Although there is common variance among the traditional predictors and the Average AP exam scores in the hierarchical regressions, Average AP exam scores provide significant incremental predictive validity for grades across all years. The greatest increment in predictive validity was found for first-year GPA -- a 9% increment to the overall amount of variance accounted for. These results confirm Hypothesis 2, that is, AP exam domain knowledge indicators provided independent and incremental prediction of academic achievement.

The trait complexes in isolation also have significant validity for predicting cumulative college GPA. Similar to the other measures, the highest predictive validity was found for the first year GPA ($R^2 = .14$), but the decline in validity was less pronounced in later years (dropping to $R^2 = .10$ for Year 3). Even after the traditional predictors and average AP scores were entered into the regression equation, trait complex scores accounted for significant incremental predictive validity (an additional 5 – 8% of variance accounted for). In comparison to the

predictive validity for first-year GPA of the traditional measures ($R^2 = .23$), the addition of average AP exam scores and trait complex scores substantially raised the total percent of variance accounted for (total $R^2 = .40$). These results confirm Hypothesis 1, that is, AP exam domain knowledge indicators provided independent and incremental prediction of academic achievement. Average AP exam scores were also positively correlated with the three facilitative trait complexes ($r = .37, .10$, and $.24$ with math/science self-concept, mastery/organization and openness and verbal self-concept complexes respectively, all $p < .05$), negatively related to one of the impeding trait complexes ($r = -.11, p < .05$), and not significantly correlated with the anxiety in achievement contexts complex ($r = -.06$), though the trait complex scores were obtained *after* completion of the AP exams. Because of the temporal order of the assessments, AP exam scores could have conceivably influenced the trait complex scores.

Prediction of STEM Persistence

Three different logistic regressions for STEM persistence were performed, highlighting the effects of: (a) traditional predictors of academic performance; (b) average AP exam scores, (c) trait complexes, gender, and first-year GPA; and (d) a stepwise regression with the variables from (a) - (c) entered as predictors. The results of these analyses are shown in **Table 6**. The first model, with traditional predictor measures, indicated an overall significant prediction of STEM persistence, though SAT-Verbal scores did not account for significant variance. Similarly, in the second model, average AP exam scores also showed a significant prediction of STEM persistence. For the third model, we initially entered all five of the trait complexes, along with gender and interaction terms for the trait complexes and gender, and first-year GPA. The reason for including first-year GPA is that it has been shown to be a very powerful predictor of STEM

persistence (Ackerman et al., in press), and we wanted to ascertain whether the trait complexes and gender variables remained important predictors in the context of predictions made on the basis of first-year GPA. Although as expected, first-year GPA was a significant predictor of STEM persistence, only two of the trait complexes and their gender interactions remained significant. So, trait complexes III, IV and V (openness and verbal self-concept, anxiety in achievement contexts, and extroversion) were removed from the equation, and the regression re-computed. The remaining model results are shown in Table 6. Although gender as a separate variable did not significantly contribute to the prediction of STEM persistence, the effects of the math/science self-concept trait complex (Complex I) and the mastery/organization trait complex (Complex II), and their interactions with gender were all significant predictors of STEM persistence.

Because the gender and trait complex interactions appeared to be especially salient, we plotted the respective means for the trait complexes for both genders and for STEM-persisters and STEM-leavers, along with the respective means for students who had both initial and final majors in non-STEM areas for reference purposes. These means are shown in **Figure 1** and **Figure 2**. Before discussing these figures, it may be useful to recall (from Table 4) that there were baseline gender differences on these two trait complexes across the entire sample -- men had higher scores on the math/science self-concept (Complex I) than women (a small effect) and women had higher scores on mastery/organization (Complex II) than men (a medium effect). These baseline differences are reflected in the two figures across the subsamples. However, the interaction effects between the trait complexes and gender for STEM persistence present a very different picture for men and women. On average, women who were STEM-leavers had lower

math/science self-concept than men, and had substantially lower math/science self-concept than women who were STEM-persisters. Women who were STEM-leavers had math/science self-concept scores that were, in fact, much closer to the scores of women who had initial and final majors in non-STEM areas. Men who were STEM-leavers had mean math/science self-concept scores that were similar to men who were STEM-persisters, but were quite dissimilar to men who were non-STEM initial and final majors.

In contrast, the means shown for mastery/organization (Complex II) indicate a very different pattern. For this complex, the women who were STEM-leavers had similar scores to those who were STEM-persisters and non-STEM majors. For the men, those who were STEM-leavers had lower mean scores on mastery/organization than either the STEM-persisters or the non-STEM majors. These results provided support for Hypothesis 3, that is, that gender differences in trait complexes and AP exam patterns would be associated with STEM major intentions and STEM persistence.

Taken together these results indicate that the underlying profiles of men and women who were STEM-leavers are very different. That is, women who left STEM majors tended to have lower math/science self-concept, but relatively high levels of the mastery/organization trait complex, while men who left STEM majors tended to have lower levels of mastery/organization trait complex. Interestingly, of the STEM-leavers that remained at Georgia Tech, 75% of both the men and women completed their BS degrees the School of Management. It is tempting to suggest that, on average, women are more likely to leave the STEM area if they have low self-concept for pursuit of the STEM content, and men are more likely to leave the STEM area if they have less of a desire to engage in the challenging curriculum associated with STEM majors.

Many commentators focus almost exclusively on STEM-leaving as an aversion-driven reaction, but this is not the only psychologically plausible explanation. That is, it is possible that at least some of the STEM-leavers choose different majors, not because they feel unsatisfied with STEM domains or they feel incapable of completing the STEM curriculum, but rather because they are attracted to majors in other fields, such as business or management. At the very least, these represent hypotheses that may be suitable for future investigations.

It should also be noted that there are ubiquitous gender differences in GPA in both secondary and post-secondary studies (e.g., see Ceci et al., 2009). In this sample, gender differences on first-year GPA was substantial favoring women ($GPA_{men} = 2.89$; $GPA_{women} = 3.14$, $t(540) = -4.29$; $p < .01$; $d = -.40$). The first-year GPA of STEM-leavers was markedly lower than the first-year GPA of STEM persisters, yet the gender differences in first-year GPA for STEM leavers were *twice* the magnitude than those found for the overall sample (STEM-leavers: $GPA_{men} = 2.35$; $GPA_{women} = 2.81$, $t(74) = -3.33$; $p < .01$; $d = -.80$).

Three or more STEM AP Exams. One variable that was derived from an archival study of AP predictors of academic success and STEM persistence (Ackerman et al., in press) is whether students successfully completed (defined as scores of 4 or 5 on the AP exams) three or more AP exams in the STEM areas (e.g., chemistry, physics, calculus). The students in the current sample represent a small portion of the overall population in the archival study. Nonetheless, the results with this group are illuminating. Only 35 students in the sample successfully completed three or more AP exams prior to matriculation. However, at the end of the first year at Georgia Tech, the mean GPA of this subgroup was 3.50, compared to the remaining sample mean GPA of 2.94 (a significant difference with a large effect size, $t(540)$, $p <$

.01; $d = -.90$). All 35 of these students had initial major intentions in STEM and only one left Georgia Tech prior to degree completion. Of the remaining 34 students in this group, all completed degrees in STEM. These students also were significantly higher, on average, on the math/science self-concept trait complex ($M = .80$, $t(548) = -4.96$, $p < .01$, $d = -.97$), significantly higher on the openness and verbal trait complex ($M = .66$, $t(547) = -3.89$, $p < .01$; $d = -.71$), and significantly lower on the anxiety in achievement contexts trait complex ($M = -.45$, $t(567) = 2.72$, $p < .01$, $d = .52$). The effect size difference for the math/science self-concept trait complex was large; the others were medium. Although men and women participants in this study completed an equivalent number of AP exams ($M_{\text{men}} = 2.78$, $M_{\text{women}} = 2.97$; $t(585) = -.85$, ns), only 7 of the 35 students who were in the three or more STEM AP category were women (20%).

Although only small subsample of students completed 3 or more AP exams in STEM, these results indicate that the combination of ability, motivation, personality, self-concept and other characteristics that lead students to successfully develop domain knowledge in the STEM areas in high school are, in turn, strongly related to initial STEM major intentions and STEM persistence in post-secondary study.

Prediction of Graduation/Attrition

Of the original 589 students in the study, 105 did not graduate within the eight-year window of the available transcripts. Students transfer from one institution to another, or leave post-secondary study for a variety reasons. Dominant reasons for students switching from STEM majors include a loss of interest in STEM, greater interest in non-STEM majors, dissatisfaction with STEM curricula, and so on (see Seymour & Hewitt, 1997). At Georgia Tech, many students leave because of poor academic performance, as indexed by multiple

instances of academic probation or academic warning. The mean cumulative GPA for those who graduated was 3.17, and 2.24 for those who did not graduate ($t(581) = -14.96, p < .01, d = -1.36$). To determine the role of the various predictors examined in this study for predicting graduation/attrition, we conducted a set of logistic regressions that largely paralleled those conducted for predicting STEM persistence. The results of these regressions are shown in **Table 7**. The first model included traditional predictors of academic performance (High School GPA, SAT Verbal and SAT Math). Similar to the earlier analysis of STEM persistence, only SAT Verbal scores did not reach a significance criterion for predicting graduation/attrition. The second model included Average AP Exam scores, and this variable was also a significant predictor of graduation/attrition. For the third model, gender and the individual trait complex scores were entered. Gender was not a significant predictor of graduation/attrition, and only the mastery/organization and openness/verbal self-concept trait complexes were significant predictors of graduation/attrition. A fourth model was created with a stepwise entry procedure to determine the most substantial predictors for graduation/attrition, when all variables were considered. The results indicated that the highest magnitude predictors of graduation/attrition were openness/verbal self-concept, average AP exam score, and mastery/organization. Although these results provide insight into the possibility of improving identification of at-risk students prior to (in the case of the AP Exam scores) or early in their freshman year (in the case of the trait complex scores), it is again important to note that prior selection on the basis of High School GPA and SAT scores results in an underestimation of their relative contribution to an overall prediction equation for graduation/attrition.

Discussion/Conclusions

The current study results indicate that a focus on both broad trait complexes (composed of personality, motivation, self-concept and related trait measures) and measures of domain knowledge assessed during high school together account for significant variance in key post-secondary criterion variables of GPA, attrition, and STEM major persistence. In addition, interactions found between trait complex scores and gender on STEM persistence criteria support the proposition that women who leave STEM majors, on average, have different profiles than men who leave STEM majors. In particular, the prominent characteristics of men who left STEM majors for non-STEM majors included substantially lower scores on the mastery/organization trait complex and somewhat lower scores on the anxiety trait complex. In contrast, the prominent characteristics of women who left STEM majors for non-STEM majors included substantially lower scores on math/science self-concept and somewhat higher scores on the anxiety trait complex.

The three major hypotheses identified in this study were supported, specifically: Trait complexes represent independently significant and incrementally significant predictors of academic achievement (beyond traditional predictors) and STEM persistence (Hypothesis 1); Individual differences in AP exam domain knowledge indicators were independent and incremental predictors of academic achievement and STEM persistence (Hypothesis 2); and Gender differences in trait complexes and AP exam patterns were associated with STEM persistence.

Over the past three decades, meta-analytic techniques have been applied to the overarching issue of prediction of academic success in a variety of domains (Poropat, 2009; Richardson et al., 2012; Robbins et al., 2004). Meta-analytic results have provided an important

perspective for evaluating the utility of psychological and other related measures for the higher-education environment, where there are typically many applicants for restricted educational opportunities. Nonetheless, these meta-analyses have three major shortcomings, as follows. First they are not typically capable of aggregating relatively narrow trait measures into larger, synergistic trait complexes. Second, the meta-analyses do not typically consider investigation of interaction effects, with respect to trait-trait interactions, and especially with respect to trait by gender interactions. The third shortcoming is a general issue about meta-analyses. That is, they tend to determine the “average” effect in the literature; depending on the adequacy of the literature (in terms of measure development, sample characteristics, methods of assessment, and so on), the average effect may under-represent what is possible in the field (for an excellent discussion of “best-evidence synthesis,” see Slavin, 1990). Statistical corrections for factor such as measurement error only address one of the key issues, and in doing so, tend to overestimate the effects that are reasonable to expect from a real-world study. Thus, we see these meta-analyses as a starting point, rather than the final word on the issues at hand.

Prediction of Grade Point Average

The Robbins et al. (2004) meta-analysis took account of the traditional predictors of academic performance, some overlapping variables (e.g., achievement motivation), and several other measures (e.g., academic goals) to predict overall GPA. Their meta-analytic correlations were first corrected for measurement error, and thus they noted that the results of their estimates of predictive validity were “optimal situation where there is no measurement error” (p. 275). The model that used only traditional measures had an R^2 of .219 with GPA, for “psychosocial and study still factors” alone, the R^2 with GPA was .164, and the estimated combined model

yielded a validity of $R^2 = .262$.

In contrast, our results, which did *not* correct for measurement error, but rather provide observed validities, indicated slightly higher validity for the traditional predictors of cumulative GPA ($R^2 = .26$). Trait complexes, in isolation had a validity of $R^2 = .15$. In addition, domain knowledge, as indexed by average AP exam scores had a validity of $R^2 = .11$. Together, these three sets of variables yielded a total combined R^2 of .37, which is markedly higher than the values indicated by Robbins et al. for the set of variables they examined, *especially keeping in mind that correcting our results for measurement error, would result in even higher estimated validity coefficients for the measures we examined.*

Poropat's (2009) meta-analysis of five-factor model personality measures and academic performance was not limited to post-secondary study in the way that the Robbins et al. (2004) study was, so the generalizations from his results are less clear. However, the estimated true-score correlations of only one factor (Openness) had a "95% credibility interval" that did not exceed zero, and only Openness ($p = .12$) and Conscientiousness ($p = .22$) exceeded Cohen's criterion for a "small" effect. In contrast, two of our five trait complex variables had *observed correlations* in the range of Poropat's estimated true-score correlations (mastery/organization $r = .28$ and math/science self-concept, $r = .18$). Again, correcting our trait complex correlations for measurement error would increase the variance accounted for beyond these observed correlations.

The Richardson et al. (2012) meta-analysis considered a larger set of variables than the earlier reviews. Several of the component traits in our trait-complex composites were included in their analysis (e.g., personality, motivation, learning strategies), as well as measures of academic

self-efficacy and grade goals, which are more specific to the particular educational environment, and less well identified as traits per se. After the various corrections (e.g., sampling error, publication bias, outliers) were made and the inclusion of the traditional predictors (high school GPA, SAT/ACT), the final meta-analytic model accounted for “28% of the variance in GPA” (p. 371). In comparison, our model (which included the traditional measures, average AP exam scores, and the five trait-complex composites) accounted for 37% of the variance in GPA with observed rather than corrected correlations. Of course, this means that 63% of the variance remains to be accounted for.

The Trait Complex Approach

The results of the current study point to two principal advantages of the trait complex approach to assessment and prediction of educationally-relevant criteria. The first is that the trait complex approach provides for a small number of predictor variables, from a vast array of potential cognitive, affective, and conative traits that are potentially relevant to educational success. The trait complex approach allows for aggregation of predictor variables that essentially cancel-out unrelated and specific variance, in favor of a small set of facilitative and impeding composites with respect to the acquisition of domain knowledge, overall academic success, and STEM persistence. The trait composites appear to operate at an appropriate level of construct-breadth and a good match to educational criteria (i.e., Brunswik Symmetry). The second principal advantage of the trait complex approach is that it allows for the investigation of interactions (such as with gender); something that is generally missing from the meta-analytically derived correlations reported in the literature. The main disadvantage of the trait-complex approach, however, is that the gains in predictive power that come from aggregation

and/or synergy may come at the cost of pinpointing specific trait-behavior linkages.

Advanced Placement Exam Scores

Our findings show the potential value of using AP exam scores in the prediction of future academic performance and STEM major persistence. In our sample of a 2002 student cohort, the average number of AP exams completed was 2.84. By 2009, the average number of AP exams completed by matriculating students was 4.14, with nearly 5% of the students completing 10 or more AP exams. Historically, selective colleges and universities have not taken AP exam performance into account when making admissions decisions for two primary reasons: First, AP exams in the early years of the existence of the AP program (when there were only 10 exams) were almost exclusively completed by high school seniors shortly before graduation and well after the college/university admissions decisions were made. The second reason has a more practical basis -- the College Board does not regularly send out AP exam scores to the applicants' full list of institutions.

Rather than use AP tests for admission decisions, many colleges/universities acknowledge student enrollment in AP (and AP-type courses) by providing bonus credit to the students in the form of a GPA boost for high school AP courses. From a public policy perspective, such a procedure makes sense because it encourages students to enroll in challenging courses in high school. The problem, from a psychometric perspective, is that there is little evidence that such *adjustments* to the GPA have any incremental validity for predicting academic success or STEM persistence in later post-secondary study (e.g., see Adelman, 2006; Geiser & Santelices, 2004; Klopfenstein & Thomas, 2010). In contrast, average AP scores have been shown to provide substantial incremental predictive validity over traditional measures for

predicting academic success in college and STEM persistence. Students who complete few AP exams with high scores tend to do better, compared to students who complete a greater number of AP exams but perform less well on average (Ackerman et al., in press). The underlying reasons for the validity of this measure are likely to be complex, but we think that future research will show that students who accurately gauge their interests, motivation, and ability when deciding to enroll in an AP course, and students who devote the necessary effort toward domain knowledge acquisition in high school, are the students who succeed on AP exams and later post-secondary study. Students who overreach and/or who do not study rigorously for the AP classes and exams appear to be those who may be at-risk for poor performance in the challenging post-secondary environment.

We found that the successful completion of multiple AP exams in STEM was a key indicator of initial STEM persistence. These findings indicate that overall, for STEM persistence, and specifically for girls and women who want to pursue STEM majors, decisions made early in high school may have significant impact on later success. Even decisions made in middle school could significantly impact success in STEM. That is, if a student does not complete an algebra course by the 8th grade, it will be exceedingly difficult to complete an AP Calculus course in high school, which will limit access one of the primary STEM AP courses (Sadler & Tai, 2007; Wang & Goldschmidt, 2003).

In terms of the practical consideration of AP scores being available to inform admission decisions, it is important to note that in recent years the timing of AP exams has changed in conjunction with the proliferation of new AP exams. In the most recent round of testing, just over half of the AP exams were completed by students earlier than their senior year of high

school, meaning that in theory, these exam results could be available for consideration by admissions personnel. Whether having about half of the students' total AP exam data is sufficient for improving admissions decisions is an open question, and one that clearly merits future research.

Implications for Educational Practices and Future Research

The findings from this research indicate two sets of significant and substantial indicators for post-secondary academic success and STEM persistence: (a) a small set of trait complexes that have facilitative and impeding properties and (b) individual differences in domain knowledge prior to post-secondary matriculation. The implications of the results regarding domain knowledge are straightforward. Unlike traditional high-stakes ability or aptitude tests, which are norm-referenced, AP-type examinations are content-referenced.⁶ Students who hope to perform well on the tests have clear and well-articulated curricula that represent the content of the exams. Engaging the student with a program of study is perhaps the main strategy towards achieving success on these examinations. Review of exam results could indicate gaps in the students' knowledge, and might be used for the purpose of planning programs of remediation. Furthermore, given that average AP exam scores are predictive of post-secondary success, one implication is that students should only attempt those courses of study for which they have inherent interest and a commitment for high levels of performance. Schools have diverse policies regarding whether access to AP and similar programs is "open" (where any student may enroll, if he/she has the appropriate prerequisite courses), or "restricted" (where enrollment decisions are made on the basis of prior course grades or teacher recommendations). Consideration of the association between AP exam scores and future post-secondary educational

success should perhaps prompt a review of these different policies for enrollment of AP courses at the high-school level.

We recognize that there are competing goals of various stakeholders when it comes to AP and similar courses (e.g., honors courses). For example, some researchers have argued that the AP program should be limited to only those highly intellectually talented students who are most likely to succeed on the exams, while others have adopted a more inclusive approach (for a review and discussion, see Bleske-Rechek, Lubinski, & Benbow, 2004). Although it is tempting to speculate, it is impossible to scientifically evaluate whether AP courses provide a benefit or harm to students who enroll in them but do not obtain ‘passing’ scores on the AP exam, given the absence of a control/treatment study where students would be randomly assigned to AP and non-AP classes, and then outcomes assessed for both. If, however, average AP exam scores were used as part of the post-secondary selection process, the answer would be that enrolling in an AP course but obtaining a low AP Exam score would result in a poorer outcome (in terms of probability of being selected by a college/university). Similarly, although students who are at the extremely high end of the ability scale (e.g., top 1%) may find AP courses to lack intellectual challenge because of attempts to make the AP program accessible to a greater portion of the population (see Bleske-Recheck, et al., 2004), multiple stakeholders will have different opinions as to whether a greater educational good is obtained by making AP programs more or less restrictive, and/or providing different avenues of educational advancement for students of different levels of ability and preparation. Such issues mainly go beyond scientific evaluation, and are much better placed in the educational policy domain.

With respect to STEM majors in general, and gender issues in particular, several related

constructs and factors have been proposed and evaluated as contributing variables to the decisions of individuals to pursue STEM courses and careers. Ceci and his colleagues have identified variables of motivation, interests, and activities as the most critical psychological factors associated with the underrepresentation of women in science fields (Ceci et al., 2009). Other investigators have focused on ability and knowledge differences, especially in the spatial ability and mechanical knowledge domains (e.g., Wai et al., 2009). In the aggregate, although extant research suggests a great many variables exert significant influences, it remains unclear which variables are the *most* important determinants of gender differences in pursuit of STEM careers. Identifying the influences that are most important is a critical issue when it comes to forming action plans to increase the level of STEM participation of women.

Given that STEM domain-knowledge differences between individuals (and genders) are already evident *prior* to university matriculation, the most effective means to address the issue of female underrepresentation in STEM areas will likely be in high school or perhaps earlier, such as when selection for Algebra classes is done in middle school. One possible avenue for future research is to determine, in middle school or early in high school, the most relevant cognitive, affective, and conative variables for predicting success in AP-type exams, and then to provide that information to various stakeholders to aid in decisions regarding selection (which students are allowed/encouraged to enroll in the courses) and classification (which courses are most well-matched to the student's characteristics) decisions.

The implications from the trait-complex results are multifaceted, partly because the general expectation regarding the component traits is that they are relatively stable, and not generally amenable to modification. Thus, the main uses of these measures are for prediction,

either in terms of selection or in terms of identifying students who may be at-risk. Because the underlying self-report measures are not suitable for high-risk assessments (because they can be ‘faked’), the best uses of trait-complex measurements in the domain of selection will be with self-selection; that is, students could complete the measures to gain insight into the most suitable majors for them. Once students matriculate, though, the trait complex measures might be used in a program that focuses on interventions for students who are at risk for attrition from post-secondary study or for leaving STEM majors. Again, the focus would be on situations where the students would complete the measures for insight and counseling purposes. Finally, understanding that gender differences in key trait complexes interact with important academic criteria implies that the key personal characteristics related to success in STEM fields may be different for men and women. If future research confirms this proposition, then it may be effective to consider gender differences in designing educational curricula in a way to maximize STEM major retention.

References

- Ackerman, P. L. (1996). A theory of adult intellectual development: process, personality, interests, and knowledge. *Intelligence*, 22, 229-259. doi: [10.1016/S0160-2896\(96\)90016-1](https://doi.org/10.1016/S0160-2896(96)90016-1)
- Ackerman, P. L. (2003a). Aptitude complexes and trait complexes. *Educational Psychologist*, 38, 85-93. doi: [10.1207/S15326985EP3802_3](https://doi.org/10.1207/S15326985EP3802_3)
- Ackerman, P. L. (2003b). Cognitive ability and non-ability trait determinants of expertise. *Educational Researcher*, 32(8), 15-20.
- Ackerman, P. L., & Beier, M. E. (2006). Determinants of domain knowledge and independent study learning in an adult sample. *Journal of Educational Psychology*, 98, 366-381. doi: [10.1037/0022-0663.98.2.366](https://doi.org/10.1037/0022-0663.98.2.366)
- Ackerman, P. L., Bowen, K. R., Beier, M. B., & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology*, 93, 797-825. doi: [10.1037/0022-0663.93.4.797](https://doi.org/10.1037/0022-0663.93.4.797)
- Ackerman, P. L., Chamorro-Premuzic, T., & Furnham, A. (2011). Trait complexes and academic achievement: Old and new ways of examining personality in educational contexts. *British Journal of Educational Psychology*, 81, 27-40. doi: [10.1348/000709910X522564](https://doi.org/10.1348/000709910X522564)
- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121, 219-245. doi: [10.1037/0033-2909.121.2.219](https://doi.org/10.1037/0033-2909.121.2.219)
- Ackerman, P. L., & Kanfer, R. (2005). Trait complexes, learning, and Brunswik Symmetry.

- Chapter in A. Beuducel, B. Biehl, M. Bosnjak, W. Conrad, G. Schönberger, & D. Wagener (Eds.), pp. 21-38. *Multivariate Research Strategies: Festschrift in honor of Werner W. Wittmann*. Aachen, Germany: Shaker Verlag.
- Ackerman, P. L., & Kanfer, R. (2009). Test length and cognitive fatigue: an empirical examination of performance effects and examinee reactions. *Journal of Experimental Psychology: Applied*, 15, 163-181. doi: [10.1037/a0015719](https://doi.org/10.1037/a0015719)
- Ackerman, P. L., Kanfer, R., & Calderwood, C. (in press). High school Advanced Placement and student performance in college: STEM majors, non-STEM majors, and gender differences. To appear in *Teachers College Record*.
- Ackerman, P. L., Kanfer, R., & Goff, M. (1995). Cognitive and noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied*, 1, 270-304. doi: [10.1037/1076-898X.1.4.270](https://doi.org/10.1037/1076-898X.1.4.270)
- Ackerman, P. L., & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *Journal of Experimental Psychology: Applied*, 13, 57-78. doi: [10.1037/1076-898X.13.2.57](https://doi.org/10.1037/1076-898X.13.2.57)
- Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.
- Armstrong, P. I., Day, S. X., McVay, J. P., & Rounds, J. (2008). Holland's RIASEC Model as an integrative framework for individual differences. *Journal of Counseling Psychology*, 55, 1-18. doi: [10.1037/0022-0167.55.1.1](https://doi.org/10.1037/0022-0167.55.1.1)
- Astin, A. W. (1993). *What matters in college? Four critical years revisited*. San Francisco: Jossey-Bass.

- Baskin, P. (2012). White House and universities pledge greater effort to retain science students. *The Chronicle of Higher Education* (February 7, 2012). Online Edition:
<http://chronicle.com>
- Beier, M. E., & Ackerman, P. L. (2012). Time in personnel selection. In N. Schmitt (Ed.), *The Oxford handbook of personnel selection and assessment* (pp. 721-739). New York, NY: Oxford University Press.
- Benbow, C., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, 210, 1262-1264.
- Bernreuter, R. G., & Goodman, C. H. (1941). A study of Thurstone's Primary Mental Abilities Tests applied to freshman engineering students. *Journal of Educational Psychology*, 32, 55-60. doi: [10.1037/h0062956](https://doi.org/10.1037/h0062956)
- Bethell-Fox, C. E., & Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 12-23. doi: [10.1037/0096-1523.14.1.12](https://doi.org/10.1037/0096-1523.14.1.12)
- Bleske-Recheck, A., Lubinski, D., & Benbow, C. F. (2004). Meeting the educational needs of special populations: Advanced Placement's role in developing exceptional human capital. *Psychological Science*, 15, 217-224. doi: 10.1111/j.0956-7976.2004.15:217
- Carroll, J. B. (1982). The measurement of intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence*. Cambridge: Cambridge University Press. 29-120.
- Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, 135, 218-261. doi: [10.1037/a0014412](https://doi.org/10.1037/a0014412)

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cole, N. S. (1969). Differential validity in the ACT tests. *ACT Research Report, 30*. Iowa City: American College Testing Program.
- College Board (2011). AP Data 2011. <http://professionals.collegeboard.com/data-reports-research/ap/data>
- College Board (2012). Advances in AP: Course and Exam Redesign. <http://aphighered.collegeboard.org/exams/course-exam-revision>.
- Costa, P. T., Jr. & McCrae, R. R. (1992). *Revised NEO Personality Inventory and Five-Factor Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334. doi: [10.1007/BF02310555](https://doi.org/10.1007/BF02310555)
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington Publishers.
- DiYanni, R. (2006). *The story of AP: The Advanced Placement Program*. College Board. <http://apcentral.collegeboard.com/program>
- Dwyer, A. C. (2011). *An investigation of the potential use of advanced placement exam scores in the college admission of transfer students*. Unpublished doctoral dissertation. Lincoln, NE: University of Nebraska.
- Ehrenberg, R. G. (2010). Analyzing the factors that influence persistence rates in STEM field, majors: Introduction to the symposium. *Economics of Education Review, 29*, 888-891.
- Ellingson, J. E., & McFarland, L. A. (2011). Understanding faking behavior through the lens of motivation: An application of VIE theory. *Human Performance, 24*, 322-337. doi:

[10.1080/08959285.2011.597477](https://doi.org/10.1080/08959285.2011.597477)

French, J. W. (1955/1966). The logic of and assumptions underlying differential testing. In A. Anastasi (Ed.), *Testing problems in perspective* (pp. 321-330). American Council on Education.

Geiser, S., & Santelices, V. (2004). *The role of Advanced Placement and honors courses in college admissions*. Berkeley, CA: Center for Studies in Higher Education, UC Berkeley. Retrieved August 20, 2011 from <http://escholarship.org/uc/item/3ft1g8rz>.

Georgia Institute of Technology (2009). *Annual First-Time Freshman Retention Study*. Atlanta, GA: Office of Institutional Research and Planning.

Georgia Institute of Technology (2012). www.gatech.edu/newsroom/release.html?nid=11761, retrieved 4/28/12.

Hagearty, M. (March 24, 2003). Introductory course helps ease adjustment to college life. *The Whistle* (p. 2). Atlanta, GA: Georgia Institute of Technology.

Hedges, L. V., & Nowell, A. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 270, 41-45.

Humphreys, L. G. (1968). The fleeting nature of the prediction of college academic success. *Journal of Educational Psychology*, 59, 375-380. doi: [10.1037/h0026234](https://doi.org/10.1037/h0026234)

Kahn, J. H., Nauta, M. M., Gailbreath, R. D., Tipps, J., & Chartrand, J. M. (2002). The utility of career and personality assessment in predicting academic progress. *Journal of Career Assessment*, 10, 3-23. doi: [10.1177/1069072702010001001](https://doi.org/10.1177/1069072702010001001)

Kanfer, R., & Ackerman, P. L. (2000) Individual differences in work motivation: Further explorations of a trait framework. *Applied Psychology: An International Review*, 49 (3),

469-481. doi: [10.1111/1464-0597.00026](https://doi.org/10.1111/1464-0597.00026)

- Klopfenstein, K., & Thomas, M. K. (2010). Advanced placement participation: Evaluating the policies of states and colleges. In Sadler, P. M., Sonnert, G., Tai, R. H., & Klopfenstein, K. (Eds). *AP: A critical examination of the Advanced Placement program* (pp. 189-218). Cambridge, MA: Harvard Educational Press.
- Lippa, R. (2005). *Gender, nature, and nurture*. Mahwah, NJ: Erlbaum.
- Lohman, D. F., Pellegrino, J. W., Alderton, D. L., & Regian, J. W. (1987). Dimensions and components of individual differences in spatial abilities. In S. H. Irvine & S. E. Newstead (Eds.). *Intelligence and cognition: Contemporary frames of reference. NATO ASI series D: Behavioral and social sciences, No. 38*, (pp. 253-312). Dordrecht, Netherlands: Martinus Nijhoff Publishing.
- Ma, X., & Johnson, W. (2008). Mathematics as the critical filter: Curricular effects on gendered career choices. In H. M. G. Watt & J. S. Eccles (Eds). *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 55-84). Washington, DC. American Psychological Association. doi: [10.1037/11706-002](https://doi.org/10.1037/11706-002)
- NAEP 1999. *[National Assessment of Educational Progress] Trends in Academic Progress: Three Decades of Student Performance.* National Center for Education Statistics (NCES) 2000-469. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics (2009). Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education. *Stats in Brief* NCES 2009-161. Washington, DC: IES. National Center for Education Statistics.

National Center for Educational Statistics (2012). CES datalab tables.

www.nces.ed.gov/datalab/tableslibrary/viewtable.aspx?tableid=7506. Downloaded April 28, 2012.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw Hill.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students*. San Francisco: Jossey-Bass.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53, 801 - 813. doi: [10.1177/0013164493053003024](https://doi.org/10.1177/0013164493053003024)

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

President's Council of Advisors on Science and Technology (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: Executive Office of the President of the United States. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138, 353-387. doi: [10.1037/a0026838](https://doi.org/10.1037/a0026838)

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](https://doi.org/10.1037/0033-2909.130.2.261)
- Roberts, B. W., & Robins, B. W. (2000). Broad dispositions, broad aspirations: The intersection of personality traits and major life goals. *Personality and Social Psychology Bulletin*, 26, 1284-1296. doi: [10.1177/0146167200262009](https://doi.org/10.1177/0146167200262009)
- Rolfhus, E. L., & Ackerman, P. L. (1999). Assessing individual differences in knowledge: Knowledge structures and traits. *Journal of Educational Psychology*, 91, 511-526. doi: [10.1037/0022-0663.91.3.511](https://doi.org/10.1037/0022-0663.91.3.511)
- Sadler, P. M., & Tai, R. H. (2007). Advanced Placement exam scores as a predictor of performance in introductory college biology, chemistry, and physics classes. *Science Educator*, 16(2), 1-19.
- Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.
- Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research*, 60, 471-499.
- Snow, R. E. (1963). *Effects of learner characteristics in learning from instructional films*. Unpublished doctoral thesis. Purdue University. University Microfilms #6404928.
- Staggs, G. D., Larson, L. M., & Borgen, F. H. (2007). Convergence of personality and interests: Meta-analysis of the Multidimensional Personality Questionnaire and the Strong Interest Inventory. *Journal of Career Assessment*, 15, 423-445. doi: [10.1177/1069072707305760](https://doi.org/10.1177/1069072707305760)
- Stein, M. I. (1963). *Personality measures in admissions: Antecedent and personality factors as predictors of college success*. New York: College Entrance Examination Board (CEEB).
- Stumpf, H., & Stanley, J. C. (1996). Gender-related differences on the College Board's

- Advanced Placement and Achievement Tests, 1982-1992. *Journal of Educational Psychology*, 88, 353-364. doi: [10.1037/0022-0663.88.2.353](https://doi.org/10.1037/0022-0663.88.2.353)
- Sullivan, B. A., & Hansen, J-I. C. (2004). Mapping associations between interests and personality: Toward a conceptual understanding of individual differences in vocational behavior. *Journal of Counseling Psychology*, 51, 287-298. doi: [10.1037/0022-0167.51.3.287](https://doi.org/10.1037/0022-0167.51.3.287)
- Sullivan, J. F. (2006). Broadening engineering's participation -- A call for K-16 engineering education. *The Bridge*, 36, No. 2. Washington, DC: National Academies Press.
- Taasoobshirazi, G., and Carr, M. (2008). Gender differences in science: An expertise perspective. *Educational Psychology Review*, 20, 149-169. doi: [10.1007/s10648-007-9067-y](https://doi.org/10.1007/s10648-007-9067-y)
- Tellegen, A. (1982). *Brief manual for the Multidimensional Personality Questionnaire (MPQ)*. Minneapolis, MN: Author.
- Thorndike, E. L. (1920). Intelligence examinations for college entrance. *Journal of Educational Research*, 1, 329-337.
- Thorndike, R. L. (1986). The role of general ability in prediction. *Journal of Vocational Behavior*, 29, 332-339. doi: [10.1207/s15327906mbr2003_1](https://doi.org/10.1207/s15327906mbr2003_1)
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 43, 89-125.
- Viswanathan, M. (1993). Measurement of individual differences in preference for numerical information. *Journal of Applied Psychology*, 78, 741-752. doi: [10.1037/0021-9010.78.5.741](https://doi.org/10.1037/0021-9010.78.5.741)

- Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117, 250-270. doi: [10.1037/0033-2909.117.2.250](https://doi.org/10.1037/0033-2909.117.2.250)
- Wang, J., & Goldschmidt, P. (2003). Importance of middle school mathematics on high school students' mathematics achievement. *The Journal of Educational Research*, 97, 3-19. doi: [10.1080/00220670309596624](https://doi.org/10.1080/00220670309596624)
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over fifty years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101, 817-835. doi: [10.1037/a0016127](https://doi.org/10.1037/a0016127)
- Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology*, 99, 397-420. doi: [10.1037/0022-0663.99.2.397](https://doi.org/10.1037/0022-0663.99.2.397)
- Wittmann, W. W., & Süß, H.-M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik Symmetry. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 77 - 108). Washington, D.C.: American Psychological Association. doi: [10.1037/10315-004](https://doi.org/10.1037/10315-004)
- Xie, Y. & Killewald, A. A. (2012). *Is American science in decline?* Cambridge, MA: Harvard University Press.

Author Note

Correspondence concerning this article should be addressed to Phillip L. Ackerman, School of Psychology, Georgia Institute of Technology, 654 Cherry Street, MC 0170, Atlanta, GA 30332-0170. E-mail: phillip.ackerman@psych.gatech.edu

Table 1. Population and Sample Demographics

	Population		Sample		<i>t</i>	<i>d</i>
	<i>Mean</i>	<i>sd</i>	<i>Mean</i>	<i>sd</i>		
High School GPA	3.81	.28	3.85	.28	-3.10**	-.14
SAT Verbal	636.00	69.15	635.55	70.33	.14	.01
SAT Math	689.17	62.18	687.50	65.81	.57	.03
Cumulative GPA at Georgia Tech	2.91	.73	3.01	.67	-3.01**	-.14
	%		%		$\chi^2(1df)$	
Gender (% Male)	72.0%		66.9%		7.62**	
Initial STEM Majors (%)	87.8%		85.7%		4.00*	

Notes: Population: $N = 2,274$; Sample: $N = 589$; t -test $df = 2,861$; χ^2 $df = 1$

* $p < .05$; ** $p < .01$. GPA = Grade Point Average. STEM = Science, Technology, Engineering, and Math; d = Cohen's d

Table 2. Descriptive statistics for the trait measures.

<i>Scale</i>	<i>#items</i>	<i>Mean</i>	<i>sd</i>	<i>α</i>
1. Desire to Learn (MTQ)	8	34.74	5.36	.80
2. Mastery (MTQ)	8	34.33	5.47	.79
3. Worry (MTQ)	10	37.60	8.25	.84
4. Emotionality (MTQ)	9	28.36	7.47	.82
5. Extraversion (NEO-FFI)	12	46.83	7.97	.79
6. Openness (NEO-FFI)	12	42.10	7.36	.60
7. Conscientiousness (NEO-FFI)	12	46.74	7.43	.76
8. Social Potency (MPQ)	26	11.85	5.98	.88
9. Social Closeness (MPQ)	22	15.34	5.07	.87
10. Test Anxiety (MSLQ)	5	16.80	5.02	.80
11. Intrinsic Goal Orientation (MSLQ)	4	16.14	3.22	.70
12. Peer Learning (MSLQ)	3	10.34	2.45	.48
13. Metacognitive Self-Regulation (MSLQ)	12	44.53	6.99	.73
14. Time and Study Environmental Management (MSLQ)	8	32.64	5.86	.76
15. Effort Regulation (MSLQ)	4	16.46	3.17	.62
16. Critical Thinking (MSLQ)	5	17.79	4.36	.78
17. Organization (MSLQ)	4	13.57	3.71	.69
18. Rehearsal (MSLQ)	4	15.30	3.45	.65
19. Verbal Self-Concept	6	26.15	5.53	.82
20. Math Self-Concept	6	28.37	5.35	.87
21. Spatial Self-Concept	6	27.20	5.24	.84
22. Science Self-Concept	6	26.28	5.68	.89
23. Self-estimate of Organizational Skills	5	18.55	4.59	.76
24. Self-estimate of Spatial and Science Abilities	2	9.63	2.10	.65
25. Self-estimate of Verbal Abilities	4	18.55	4.59	.86
26. Self-estimate of Math Abilities	4	21.72	4.22	.92
27. Numerical Preferences	11	49.12	9.30	.89
28. Life goals - status	6	19.02	5.05	.78
29. Life goals - easy life	5	17.48	3.20	.60

Notes: NEO-FFI = NEO Five-Factor Inventory; MPQ = Multidimensional Personality Questionnaire, MTQ = Motivational Trait Questionnaire, MSLQ = Motivated Strategies for Learning Questionnaire. α = Cronbach's alpha internal consistency statistic.

Table 3. *Trait complexes and their intercorrelations*Factor/Trait Complex**I. Math/Science Self-Concept**

Scales: Math Self-Concept, Spatial Self-Concept, Science Self-Concept, Self Estimates of Math Ability, Numerical Preferences

Number of scales = 6, $\alpha = .87$

II. Mastery/Organization

Scales: Desire to Learn, Mastery, Conscientiousness, Metacognitive Self-Regulation, Time and Study Environmental Management, Effort Regulation, Organization, Rehearsal, Self-estimates of Organizational Skills

Number of scales = 9, $\alpha = .88$

III. Openness and Verbal Self-Concept

Scales: Openness, Intrinsic Goals, Critical Thinking, Verbal Self-Concept, Self-Estimate of Verbal Abilities

Number of scales = 5, $\alpha = .70$

IV. Anxiety in Achievement Contexts

Scales: Worry, Emotionality, Test Anxiety

Number of scales: = 3, $\alpha = .85$

V. Extroversion

Scales: Extroversion, Social Potency, Social Closeness, Peer Learning, Life Goals (Status), Life Goals (Easy Life)

Number of scales = 6, $\alpha = .71$

Correlations among trait complexes

	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
1. Math/Science Self-Concept				
2. Mastery/Organization	.36**			
3. Openness/Verbal Self-Concept	.38**	.45*		
4. Anxiety in Achievement Contexts	-.12*	.03	-.10*	
5. Extroversion	-.03	.22**	.04	.09*

Notes: * $p < .05$; ** $p < .01$

Table 4. Trait Complexes by Gender.

	Men	sdMen	Women	sdWomen	<i>t</i>	<i>d</i>
1. Math/Science SC	.080	1.00	-.162	.99	2.75**	.24
2. Mastery/Organization	-.201	.91	.387	1.06	-6.82**	-.60
3. Openness/Verbal Self-Concept	-.012	1.02	.024	.95	-.41ns	.04
4. Anxiety in Achievement	-.102	.98	.205	1.01	-3.49**	-.31
5. Extroversion	-.070	1.02	.139	.85	-2.24*	-.22

N(men) = 394

N(women) = 195

Notes: SC= Self Concept; ns = not significant; d = Cohen's d, *p < .05; **p < .01

Table 5. Prediction of yearly cumulative and overall cumulative GPA from traditional predictors, average AP exam scores and Trait complex scores.

Regression Step	Cumulative Grade Point Average				
	Year 1	Year 2	Year 3	Year 4	Cumulative
1. HS GPA, SAT-V,SATM					
R^2 in Isolation	.23	.20	.21	.15	.26
R^2 to Add	.23	.20	.21	.15	.26
Total R^2	.23	.20	.21	.15	.26
2. Average AP Exam Scores					
R^2 in Isolation	.21	.20	.18	.12	.15
R^2 to Add	.09	.09	.07	.07	.05
Total R^2	.32	.30	.28	.22	.31
3. Complexes 1 - 5					
R^2 in Isolation	.14	.12	.10	.12	.11
R^2 to Add	.08	.07	.05	.05	.06
Total R^2	.40	.37	.34	.27	.37

Notes: Degrees of freedom in numerator: Step 1 = 3, Step 2 = 4, Step 3 = 9. Degrees of freedom for denominator vary, based on sample with GPA at each year (e.g., Year 1 = 542, Year 2 = 517, Year 3 = 494, Year 4 = 307) and students with AP exam scores (N = 396). Cumulative GPA was the terminal mean GPA for all participants, regardless of attrition or graduation status. Regressions were computed with pairwise inclusion procedures. All correlations are significantly different from zero, $p < .01$.

Table 6.

Logistic regression analyses for STEM persistence.

Predictors	B	S.E.	Wald	<i>p</i>
Model 1 - Traditional measures				
High School GPA	1.45	.53	7.36	.007
SAT Verbal	.00008	.002	.0018	ns
SAT Math	.0076	.002	9.41	.002
Model 2 - Average AP exam score				
Average AP score	.65	.19	11.81	<.001
Model 3 - Gender, Trait complexes, Interactions, and First-Year GPA				
Gender	-.04	.24	.02	ns
Trait Complex 1	-1.80	.60	9.15	.002
Trait Complex 1 X Gender	1.59	.47	11.48	.001
Trait Complex 2	1.51	.62	6.00	.01
Trait Complex 2 X Gender	-1.15	.45	6.63	.01
First-Year GPA	-3.58	.84	18.18	<.001
Model 4 Stepwise with Measures from Models 1 - 3				
Trait Complex 1	-1.21	.50	5.88	.02
Trait Complex 1 X Gender	1.10	.38	8.61	.01
First-Year GPA	1.95	.28	49.91	<.001
Model Fit Statistics				
	<i>df</i>	X ²	-2 LL	C&S R ²
Model 1	3	21.10**	361.03	.05
Model 2	1	12.34**	249.70	.04
Model 3	6	90.64**	271.81	.21
Model 4	3	79.18**	267.42	.19

Notes: ns = not significant; ***p* < .01; -2 LL = -2 log likelihood; C & S = Cox & Snell.

Table 7.

Logistic regression analyses for Graduation/Attrition.

Predictors	B	S.E.	Wald	<i>p</i>
Model 1 - Traditional measures				
High School GPA	1.36	.36	14.23	< .001
SAT Verbal	-.0005	.0017	.09	ns
SAT Math	.0038	.0018	4.18	.04
Model 2 - Average AP Exam Score				
Average AP Exam Score	.56	.16	11.84	< .001
Model 3 - Gender and Trait complexes				
Gender	-.06	.28	.05	ns
Trait Complex 1	.25	.14	3.22	ns
Trait Complex 2	.34	.16	4.88	.03
Trait Complex 3	-.49	.15	10.67	.001
Trait Complex 4	.07	.12	.33	ns
Trait Complex 5	-.06	.12	.24	ns
Model 4 Stepwise with Traditional Measures, Average AP Exam Score and Trait Complexes				
Trait Complex 2	.57	.18	9.90	.002
Trait Complex 3	-.86	.20	17.72	<.001
Average AP Exam Score	.75	.19	15.00	<.001
Model Fit Statistics				
	<i>df</i>	X^2	-2LL	C&S R^2
Model 1	3	22.74**	515.57	.04
Model 2	1	12.21**	321.18	.03
Model 3	6	15.61*	434.20	.03
Model 4	3	31.78**	267.09	.08

Notes: ns = not significant; * $p < .05$; ** $p < .01$; ; -2 LL = -2 log likelihood; C & S = Cox & Snell.

Footnotes

-
1. For the current study, we have defined STEM majors as those within the following areas: Physical Sciences (e.g., Physics, Chemistry), Biology, Engineering (all types), Architecture, and Math. The designation of STEM areas is an evolving issue (e.g., whether or not Psychology is considered to be a STEM area), thus when comparing across different studies, one should take account of differing specifications for which major domains are considered STEM or non-STEM.
 2. One possible reason for the lack of high-stakes testing evidence is that spatial/mechanical abilities tend to be influenced much more by drill and practice effects than are tests of verbal or quantitative abilities (e.g., see Bethell-Fox & Shepard, 1988; Lohman, Pellegrino, Alderton & Regian, 1987). The large practice effects on such tests make them somewhat unsuitable in an environment where applicants expend significant resources on coaching or training programs to obtain any possible benefits in entrance test scores. Thus, although there is good evidence for the idea of that spatial abilities are important predictors for STEM success, there remains much work to be done in order to develop such measures in a way that they can be used for high-stakes selection purposes.
 3. NCES used a similar definition of STEM areas to those used in the current investigation (e.g., they excluded social and behavioral sciences).
 - ⁴. Of course, there are gender differences in the antecedent variables (e.g., interests, personality, self-concept, self-efficacy) that lead to an orientation towards or away from acquisition of domain knowledge in STEM areas and non-STEM areas, even during adolescence (e.g., for a

review see Ceci et al., 2009).

⁵. Several studies of gender differences in math abilities that focus on those with extremely high ability levels (e.g., top 1%, see Benbow & Stanley, 1980; Hedges & Nowell, 1995) have noted greater between-individual variances for males, compared to females. Although the Georgia Tech population is certainly above average compared to the SAT-taking population, there were only modest magnitude differences in respective sd's for men and women on the SAT scores in this sample (Verbal = 77.99 and 65.88; Math = 70.12 and 57.63, respectively). The F-ratios for a test of the equality of two variances were $F(382,189) = 1.40$ for SAT Verbal, and 1.48 for SAT Math, which are both significant, $p < .05$. In more highly-selected samples, one might expect greater gender differences in these indicators of interindividual differences.

6. The AP exam program is currently undergoing both course and exam redesign (College Board, 2012). Because this is an ongoing process, it is not clear if changes made to the exams will have an effect on the predictive validity of exam scores for post-secondary educational success.

Figure Captions

Figure 1. Math/Science Self-Concept Trait Complex (Complex I) means by gender and initial/final major. Only degree completers are included. Number of students in each group is indicated on the graph.

Figure 2. Mastery/Organization Trait Complex (Complex II) means by gender and initial/final major. Only degree completers are included. Number of students in each group is indicated for on the graph.

Figure 1.

Figure 2.