Prediction of Four-Year College Student Performance using Cognitive and Noncognitive Predictors and the Impact on Demographic Status of Admitted Students

Neal Schmitt, Abigail Billington, Jessica Keeney

Michigan State University

Frederick L. Oswald

Rice University

Timothy Pleskac, Ruchi Sinha, & Mark Zorzie

Michigan State University

Address correspondence to:

Neal Schmitt
Department of Psychology
Michigan State University
East Lansing, MI 48824-1116
Email: schmitt@msu.edu
Phone: 517-355-9563

Author Note

The financial support of the College Board in the conduct of this project is much appreciated. With the exception of the first author, author order was determined alphabetically reflecting relatively equal contributions to the paper.

Abstract

This study was conducted to determine the validity of noncognitive and cognitive predictors of the performance of college students at the end of their fourth year in college. Results indicate that the primary predictors of cumulative college GPA were SAT/ACT scores and high school GPA (HSGPA) though biographical data and situational judgment measures added incrementally to this prediction. SAT/ACT scores and HSGPA were collected and used in various ways by participating institutions in the admissions process while situational judgment measures and biodata were collected for research purposes only during the first few weeks of the participating students' freshman year. Alternative outcomes such as a self-report of performance on a range of student performance dimensions and a measure of organizational citizenship behavior, as well as class absenteeism were best predicted by noncognitive measures. The racial composition of a student body selected using just cognitive measures or both cognitive and noncognitive measures under various levels of selectivity as well as the performance of students admitted under these scenarios is also reported. We conclude that both the biodata and SJT measures could be useful as a supplement to cognitive indices of student potential in college admissions.

Prediction of Four-Year College Student Performance using Cognitive and Noncognitive Predictors and the Impact on Demographic Status of Admitted Students

As is true when organizations hire employees, colleges and universities seek to admit and recruit the best students. Just as the qualifications that make a good employee vary across organizations or managers, so do the factors underlying notions about excellent student performance. In the educational context, these factors vary as a function of the university or admissions personnel who evaluate student credentials and performance. Traditionally, college admissions personnel use high school grade point averages (HSGPA), standardized tests of cognitive ability in the areas of verbal and mathematical skills (SAT/ACT), and sometimes records of achievement in specific subject matter areas to assess student potential. Each factor provides unique information about the applicant. Letters of recommendation, essays or interviews are being used increasingly by universities to complement these HSGPA and SAT/ACT scores. Schools vary widely in their assessment of the information contained in these supplemental materials. For example, while a reviewer at one school might assign a subjective rating to each component of the application, a reviewer at another school might form ratings of personal qualities (e.g., leadership) based on a holistic review of the materials (Rigol, 2003). Clearly, any systematic and thorough processing of this information, especially when large numbers of applicants must be processed in a short period of time, places a heavy burden on admissions personnel.

Standardized cognitive ability tests or achievement tests (like SAT/ACT) can be administered to large numbers of students efficiently and they provide a standard of comparison across students with differing educational backgrounds. Moreover, research has demonstrated consistently high criterion-related validities (approximately $r = .45$) with cumulative college

GPA, in addition to smaller but practically significant relationships with study habits, persistence, and degree attainment (Hezlett et al., 2001). Higher validities are often observed if the outcomes assessed are more proximal such as first year college GPA (Kuncel, Hezlett, & Ones, 2001, 2004). Sackett, Kuncel, Arneson, Cooper & Waters (2009) in a recent study examined various large datasets and found strong relationships between standardized tests and academic performance ($r = .44$). They found that a vast majority of these relationships were strong even after controlling for factors like socioeconomic status. On the whole both high school grade point average and standardized tests have been shown to have predictive validity in determining a variety of academic performance outcomes (e.g., Bridgeman, McCamley-Jenkins, & Ervin, 2000; Kuncel, Credé, & Thomas, 2007; Kuncel & Hezlett, 2007; Kuncel, Hezlett, & Ones, 2001, 2004).

Some college personnel and researchers, however, have reservations about standardized cognitive ability tests. Researchers point to the fact that even with the relatively high validity of the SAT and ACT college admissions tests and HSGPA, there remains a large portion of unexplained variance in college student performance measures (Breland, 1998; Payne, Rapley, & Wells, 1973). Various stakeholders in admissions testing are also becoming strident in demanding a broader array of selection tools with adequate criterion-related validity, less adverse impact, and greater relevance to a broader conceptualization of college performance. As a result of these demands, universities are already changing the role standardized tests (SAT or ACT) play in the selection process. For example, the University of California has begun to use the SAT-II, an instrument more directly tied to high school curricula, for admission decisions. More recently, in 2008, Wake Forest University became the first top 30 national university to make standardized tests (SAT or ACT) optional (Landau, 2008). More generally, a NACAC

commission (2008) recommended that the role of standardized tests in college admissions be reevaluated and perhaps diminished.

There are a number of potential benefits to be gained from broadening the selection criteria beyond SAT/ACT and HSGPA, but one important benefit is the potential increase in the diversity of students admitted into colleges. Whereas minority students often score lower on cognitive ability tests such as the SAT/ACT, there are small or no differences between majority and minority groups on many noncognitive assessments of background, interests, and motivation (Hough, 1998; Sackett, Schmitt, Ellingson, & Kabin, 2001). These relative differences in the measures translate into different rates of selection across demographic groups depending on the institution's selectivity and the manner in which the tests are used.

The need to incorporate more than just cognitive factors in the admission process has led to a growing interest in non-cognitive predictors of academic performance. Past studies have examined the role of non-cognitive predictors of academic success such as meta-cognitive skills (e.g., Zeegers, 2001), study attitudes (e.g., W.S. Zimmerman, Parks, Gray, & Michael, 1977), study motivation (e.g., Melancon, 2002) and even personality traits (e.g., Ridgell & Lounsbury, 2004). In a more recent meta-analysis, Crede and Kuncel (2008) found that non-cognitive factors like study habit, skill and study motivation among other attitudinal constructs accounted for incremental variance in academic performance beyond standardized tests and previous grades. A challenge, however, in including these non-cognitive predictors and broadening the selection criteria is how to maintain an objective means of comparing applicants on the basis of not only their cognitive ability but also their noncognitive abilities and profiles (e.g., citizenship, perseverance, adaptability). The latter noncognitive attributes are often thought to be represented in essays and interviews, both of which are labor intensive to score in reliable ways, particularly

in large undergraduate universities. Consistent with this challenge, our research team, with the support of the College Board, has been working for the last several years to develop and validate two noncognitive measures that would help evaluate applicants on twelve different dimensions relevant to college performance (see Oswald, Schmitt, Kim, Ramsay, & Gillespie, 2004; Schmitt et al., 2003; Schmitt et al., 2007) using an objectively scored format. Oswald et al. reported promising validities for biodata and situational judgment measures for a variety of outcomes measured at the end of the first year of college at one large university. Schmitt et al. (2007) reported encouraging validity against college GPA, absenteeism, and several nonacademic criteria for a group of students from 10 different universities at the conclusion of their first year in college. The focus of the latter study was on using biodata, situational judgment, and ability measures to profile students with differing outcome profiles.

In the current article, we report four-year predictive validities for the sample of 2,771 students evaluated in Schmitt et al. (2007) using college GPA, graduation status, class attendance, academic satisfaction, and organizational citizenship behavior as outcomes. In addition, with this sample of students, we examine the consequences of using the biodata and situational judgment measures, SAT/ACT, and HSGPA in a composite to make admissions decisions at varying levels of selectivity. The outcomes with respect to the ethnic diversity of the students admitted and the average GPA under these various hypothetical conditions are reported.

Contributions of this Study

The current four-year longitudinal study provides predictive validities for both cognitive (test scores and high school grades) and non-cognitive predictors (biodata and SJT) for a variety of academic outcomes like cumulative 4 year college GPA, graduation status, class attendance, academic satisfaction, and organizational citizenship behavior. The present study also illustrates

how the use of both cognitive and non-cognitive predictors may influence the ethnic diversity of admitted students at varying levels of selectivity.

As was noted at the beginning of this article, the admissions problems of academic administrators are very similar to those of private and public employers in at least four important ways. As in business organizations, there is a concern for the nature of the performance construct (Binning & Barrett, 1989; Campbell, McCloy, Oppler, & Sager, 1993) and the adequacy of measures of the criterion(a). Recent thinking on the nature of performance has led to a concern for organizational citizenship behavior and counterproductivity (Borman & Motowidlo, 1997; Dalal, 2005). Industrial/organizational psychologists interested in selection and employee performance have long been interested in work attendance and satisfaction as well as turnover (e.g., Ployhart, Schneider, Schmitt, 2006). Second, when interested in these alternative forms of performance, there has been an increasing recognition of the importance of various noncognitive predictors such as personality and background experiences (Hough, Oswald, & Ployhart, 2001; Mumford, & Stokes, 1992; Schmidt & Hunter, 1998). Third, selection researchers have also been forced to consider relatively short term performance outcomes as opposed to more desirable long term alternatives (Ackerman, 1989; Henry & Hulin, 1989). Finally, there is certainly a continuing interest in the development and use of predictors that contribute to appropriate inferences about performance with a lessened degree of adverse impact on members of different minority subgroups that are selected (Bobko, Roth, & Potosky, 1999). The research reported in this paper has been informed and motivated by selection research and we think it also contributes to that work in all four areas described.

Expanding the Criterion Space of College Student Performance

Academic institutions clearly express the desire to admit students who will succeed in the college environment, whether that means succeeding academically, interpersonally, psychologically, or otherwise. If we take seriously what colleges claim they hope to achieve with their students in various promotional materials, it also seems appropriate to reconsider traditional GPA and graduation criteria and to expand the criteria space of college student performance. Although some studies have expanded the nature of the predictors of college student success beyond the traditional predictors of HSGPA and SAT/ACT (e.g., Boyer & Sedlacek, 1988; Harackiewicz, Barron, Tauer, & Elliot, 2002), most have remained focused on the prediction of first year college GPA.

Another recent exploration of predictors of college student success is the meta-analytic effort by Crede & Kuncel (2008). Crede and Kuncel examined the predictive validity of the study habits, skills, and attitudes (SHSA) of college students in predicting GPA. They found that measures of SHSA were strong predictors of college GPA and measures of SHSA provided incremental validity in predicting freshman GPA over and above HSGPA and SAT/ACT scores. They also found that SHSA constructs were only weakly related to measures of cognitive ability and prior performance in high school. Although their findings about SHSA are valuable as possible areas for intervention for at-risk or low-performing college students, the authors caution the use of such measures in admissions contexts. First, the authors point out the vulnerability of such measures to socially desirable responding. Second, there was a near-zero relationship between high school academic performance and SHSA. This finding suggests that the SHSA that best predict college student performance are distinct from those that best predict high school student performance. Because the subjects of the studies aggregated in the Crede and Kuncel

meta-analysis were college students, it is possible that the high-performing students developed the relevant SHSA after they were in college.

Another project with similar goals to ours is The Rainbow Project (Sternberg et. al., 2006). The Rainbow measures assess analytical, creative, and practical abilities—the three components in Sternberg's conceptualization of successful intelligence. Each of these components has verbal, analytical, and figural subscales. In their sample, the triarchic intelligence factors accounted for nearly 9% incremental variance beyond the traditional cognitive factors of high school GPA and SAT scores. These measures also significantly reduced between-group differences on race for Latino and Black students; however, the measures on which they observed the largest reduction in subgroup differences were often not predictive of performance. There were very large amounts of missing data on most predictors, so it was very difficult to ascertain what impact their measures might have in a battery of tests that included both traditional cognitive predictors (SAT and HSGPA) as well as more novel measures. In addition, scoring of many of the Sternberg et al. measures was time-consuming and not very reliable.

A slightly different approach, one we take, is to expand the set of relevant criteria. As described in Oswald et al. (2004), we sought to identify the nature and number of dimensions of college student performance based on a content analysis of material published on college web pages. Based on this sample of mission statements and stated educational objectives from a range of colleges and universities, we identified 12 major dimensions named and defined in Table 1. They deal with intellectual behaviors (Knowledge, Learning, Artistic), interpersonal behaviors (Multicultural, Leadership, Interpersonal, Citizenship), and intrapersonal behaviors (Health, Career, Adaptability, Perseverance, and Ethics). These dimensions served as the basis of our

development of biodata and situational judgment predictor measures as well as some of our outcome variables.

Insert Table 1 about here

Early Development and Validation of the Noncognitive Measures

Although ability tests and high school grades shed important light on the potential for a college applicant to become a successful undergraduate student, a number of nonability measures, such as personality, motivation, and experience measures, also may predict whether or not students will be successful in their academic career (Pantages & Creedon, 1978; Robbins et al., 2004). The first of the new noncognitive measures we developed is a biographical data measure (biodata; cf. Clifton, Mumford, & Baughman, 1999; Mael, 1991; Nickels, 1994). The biodata measure assesses a student's background, interests, hobbies and typical behaviors in a wide variety of academic and life situations. The second is a situational judgment test (SJT; cf. Motowidlo, Dunnette, & Carter 1990). The SJT asks students to judge what behavioral option they would select in a series of hypothetical performance-related academic and social situations that they are likely to experience in a collegiate environment.

Biographical data have been used previously in at least one program of research involving college students. Owens and his colleagues (e.g., Owens, 1976; Owens & Schoenfeldt, 1979; and Mumford & Owens, 1987; Mumford & Stokes, 1992) have reported extensive efforts to use background data, in the form of biodata items, to subgroup samples of individuals with similar biographical profiles. To do so they developed biodata items that captured important behaviors and experiences related to student development. Using scores derived from principal

components analyses of responses to the biodata items, 2000 freshmen were clustered into subgroups with similar profiles producing 23 male and 15 female subgroups (Owens & Schoenfeldt, 1979). To evaluate this subgroup approach, the degree to which subgroup membership was associated with external performance criteria was assessed. Subgroup status was related to a variety of educational outcomes including over and underachievement, college GPA, academic probations and dismissals, and a number of course withdrawals in a series of master's theses and dissertations. Mumford, Connelly, and Clifton (1990) report that subgroups identified in this manner also predicted motivational criteria such as person-job fit and situational choice.

The programmatic research described in the previous paragraph was preceded by at least two studies involving college students. Anastasi, Meade & Schneiders (1960) administered to students a biographical inventory containing items relating to high school involvement, hobbies, and vocational goals. A scoring key was developed that differentiated between high, average and below average groups of students on noncognitive criteria such as faculty ratings, ROTC behaviors, and participation in extracurricular activities. In another study, Richards, Holland & Lutz (1967) used an extracurricular achievement record that captured high school achievement in art, music, literature, dramatic arts, leadership, and science to predict the same outcomes in college. Results indicated that high school accomplishment in one area predicts college accomplishment in the same area, but nonclassroom high school achievement is only minimally related to academic potential and achievement.

To our knowledge, there has been no previous work using SJTs in student groups prior to that reported by Oswald et al. (2004). However, previous research predicting employment success has been very encouraging (McDaniel, Hartman, Whetzel, & Grubb, 2007; McDaniel,

Morgeson, Finnegan, Campion, & Braverman, 2001). Validities from these meta-analyses against a variety of job performance measures averaged .20 (.26 when corrected for measurement error). There has been debate on the nature of the constructs measured by the SJT. They do correlate with the Big Five in the .20s and .30s and with cognitive ability approximately .30 (McDaniel et al., 2007), but there is considerable unique and reliable variance in these measures relative to other major individual difference constructs (Schmitt & Chan, 2006). Perhaps the unique variance captured by the SJT is something called practical intelligence (Wagner & Sternberg, 1991). The original study introducing SJT (Motwidlo et al., 1990) to the employment arena was oriented to the measurement of practical intelligence. Our SJT included items addressing practical everyday problems that a student would address both in and outside of the classroom. Perhaps most similar to our use of the SJT was a study by Lievens, Buyse, and Sackett (2005) in which the authors describe the actual use of a SJT for the selection of medical students in Belgium. They found it predicted grades incrementally over cognitive ability in an interpersonally-oriented curriculum.

Alternative Outcomes

Earlier in this paper, we mentioned the importance of alternative outcomes and the fact that our measures were specifically developed to predict these outcomes. We constructed two such composites for use in our research. First, we constructed a 12-item measure that reflected each of the 12 dimensions defined in Table 1. Using senior level undergraduate students, we developed behavioral definitions of each of the 12 dimensions and identified very positive and very negative behaviors on each of these dimensions. The latter were used as anchors for each of the scales. Students were asked to rate their performance on each of these scales. Earlier research indicated that there was little evidence of discriminant validity for these scales and that they were

best represented as unidimensional (Oswald et al., 2004), so the 12 ratings were combined in a single measure. The second alternative outcome was a measure of organizational citizenship behavior (OCB). In the employment arena, OCBs are activities that contribute in a positive way to organizational climate and sometimes organizational performance, but that are not directly related to the task activities defined by a job or a student's academic performance in this instance (Borman & Motowidlo, 1993; Organ, 1988). In the case of the student participants in our research, these included activities such as serving as leaders, participating in extracurricular activities and community outreach projects, representing one's school to prospective students, and tutoring other students. More detail on each of these measures is contained in the Methods section below. Conceptually, we believe these outcomes are those that were implied by the objectives of universities whose student goals we reviewed at the beginning of the project. In a more traditional vein, we also examined the prediction of cumulative college GPA, graduation at the end of the students' fourth year of college, and class attendance.

Longitudinal Multi-Institution Validation Effort (2004 to 2008)

In the summer and early fall of 2004, we began a new and larger validation effort examining a subset of items from the total available item pool for both the biodata and the SJT. This longitudinal and multi-institution validation project produced validity data for the two noncognitive measures of college student performance. It is important to note that the biodata and SJT measures were administered after students had been admitted to their respective universities in the first few weeks of their freshman year; hence, these measures were not used in making admissions decisions. The results of the original data collection with 2771 students at 10 colleges and universities in the United States and follow-up collection of outcome data at the end

of their first, second, and third semesters were described in an earlier report by Drzakowski et al. (2005) and some were the subject of the profiling attempt reported by Schmitt et al. (2007).

Purpose of the Study

The purpose of the present paper is to examine the criterion-related validity results for the biodata and SJT based on outcome data collected during and after the fourth year of college. These data are of interest because, traditionally, this is when we expect the majority of students to complete their college careers, though we recognize that many institutions are now evaluating five and even six year graduation status (ACT, http//www.act.org/research/policymakers/pdf/retain2008.pdf). We not only examine the validity coefficients of the two noncognitive measures but also focus on the capability of these measures to predict student outcomes above and beyond more traditional measures like SAT and ACT test scores and high school cumulative grade-point average (HSGPA). We also expanded the nature of the outcomes considered to include not only cumulative college GPA, but also graduation status, class absenteeism, self-reported performance on the 12 dimensions listed in Table 1, and organizational citizenship behavior. Finally, we examined how inclusion of the biodata and situational judgment measures in a composite that includes HSGPA and SAT/ACT to make admissions decisions will change the ethnic diversity of the student population and their expected grades.

The nature of the relationships we believed would be significant and substantial is summarized in Table 2 based on the nature of the constructs we assumed underlay the predictors and criteria. Given the extensive body of research indicating the validity of HSGPA and SAT/ACT in the prediction of college grades, we assumed we would find substantial relationships between these two variables and cumulative college grade point averages. We also

believed that our biodata measures of Knowledge and Continuous Learning which request information about academic activities in high school would relate to college academic performance. Similarly we believed that background activity related to Responsibility and Perseverance would relate to college grades. Finally, as a measure of judgment in practical situations related to college life, we felt that the SJT should relate to college GPA. We also expected that Knowledge, Responsibility, and Perseverance as well as the SJT, SAT/ACT, and HSGPA would be positively related to graduation in four years. In addition, we felt that those who were high in Career Orientation would be more likely to finish college on time as would those who showed Leadership skills. We did not anticipate a relationship between Continuous Learning and graduation status as we felt individuals high on this set of behaviors would pursue interests that would not be so directly related to timely graduation.

Insert Table 2 about here

A third outcome we considered was class attendance. We felt that class attendance was primarily motivational hence hypothesized that the largest bivariate relationships would be observed for the biodata and SJT measures. Specifically, we felt that the biodata dimensions of Leadership, Responsibility, Health, Career Orientation, Perseverance, and Ethics would be related to class attendance as would the SJT measure. Because the self rating of performance on the behaviorally anchored rating scale (BARS) was directly tied to the dimensions we were trying to measure with both biodata and SJT, we hypothesized a relationship between these measures. We felt the organizational citizenship measure should be most highly related to the biodata measures of Leadership, Responsibility, and Ethics as well as the SJT.

Methods

Sample

A total of 2,771 incoming freshman college students across 10 U.S. colleges and universities participated in our Time 1 data collection in early Fall, 2004. These institutions included five Big Ten universities, two historically Black institutions, a large school in the west, a southern state university, and a large private institution in the Midwest. Biodata measures, the SJT, and demographic variables were collected at this initial data collection. Following this data collection, we also requested and received archival data (HSGPA and SAT/ACT scores) from the admissions offices of the universities involved. While the HSGPA and SAT/ACT scores were collected during the admissions process and presumably used to make admissions decisions, the biodata and SJT measures were collected shortly after the students began their first term in college.

Approximately three and a half years after the original data collection at Time 1 (start of their undergraduate career), we conducted the follow-up of data collection that provided several of the outcome variables in this study. In a web-based survey, we asked students to respond to a range of self-report outcome measures including academic satisfaction, organizational citizenship behaviors, class absenteeism, and college behaviors relating to the 12 dimensions described in Table 1. Five hundred ninety-three of the original participants responded to this survey. In addition, we obtained college GPA data and graduation status from seven of the ten original institutions. Unfortunately one of the historically black schools and the institution in the west that initially included a large number of Hispanic students refused to participate. College graduation data were available on 2086 students; we had cumulative four-year college GPA data

on 1,165 students. Because of missing data, the analyses reported below were conducted on less than the total sample.

Basic demographic statistics (gender, age, and race) for subsamples at the first and final time points are presented in Table 3. As mentioned above, the total original sample size across all 10 schools was 2,771. At the final data collection, 593 responded to our web-based survey and archival data from the seven institutions were available for 2,086 students. The original sample was primarily between the ages of 18 and 20, and 64.2% were female. Although the majority of participants were Caucasian, there was a fair representation of large minority groups in the U.S. (i.e., African, Hispanic, and Asian American). The final sample at follow-up was different from the original sample in that we lost a large portion of the African American respondents and gained proportionately among Asian and Caucasian American students.

Insert Table 3 about here

Measures

Predictors

During the informed consent process, participants signed optional data release forms. For the participants who signed these forms, HSGPA data and SAT and/or ACT scores were obtained from college or university registrars. All participants had taken one of these tests (SAT & ACT) and many had taken both as part of their application to different universities, so these variables were standardized on national norms within test, combined, and used as a single index of the participants' ability and/or preparation to do college work. This was accomplished by converting any raw ACT (composite) test scores to equivalent SAT scores, using a conversion

table from www.collegeboard.com (or see Dorans, Lyu, Pommerich, & Houston, 1997). SAT and ACT composite scores were correlated .85. When a participant took both the ACT and SAT, an average of the converted ACT score(s) and the raw SAT score(s) was computed and served as an index of cognitive ability. No data were obtained for participants who did not sign the release forms.

Biodata measures. As described above, biodata items requested information about an individual's background and life history. Similar information is contained within college applications, but is often provided by students in an open-ended way and is used by admissions officers in an intuitive or implicit manner (e.g., the use of applicants' extracurricular activity lists and resumes). By contrast, biodata provide a systematic and quantitative assessment of the same information. This could provide admissions officers a more efficient and consistent method to incorporate this information in their admissions decisions.

Each of the biodata scales consisted of approximately 10 multiple-choice items that were objectively scored. Items inquired about one's previous experiences, similar to tests used in job selection processes. The revised biodata instrument used in this data collection effort consisted of 112 standard multiple-choice questions covering 11 dimensions. These 112 items were selected based on content and item properties from a previous version of the biodata, which contained 206 items. A twelfth dimension, a measure of interpersonal skills (see Table 1), was excluded because of our inability to construct a psychometrically adequate measure of the dimension.

Reliabilities of these biodata scales (as well as the other measures when they were available) are displayed on the diagonal of Table 4 along with the intercorrelations, both observed and corrected for unreliability, between all study variables. These results replicate earlier research (Oswald et al., 2004) that produced similar levels of reliability and

intercorrelation for the same dimensions. Exploratory and confirmatory factor analyses of these scales (Gillespie, Kim, Manheim, Yoo, Oswald, & Schmitt, 2002) also provided empirical support for the meaningfulness of these dimensions.

Situational Judgment Test (SJT). The SJT content also reflected the dimensions of college student success (see Table 1). Each SJT item presented a scenario that a typical college student might face related to one of these 12 dimensions. Response options represented possible behavioral responses to the scenario presented. For each scenario, the participant selected the response option that represented his or her "most likely" response and the option that represented his or her "least likely" response. Each SJT item was scored from -2 to +2, with higher scores indicating situational judgment that is in line with scoring keys developed with the help of a set of students deemed to be experts (i.e., junior and senior college students who have successfully persevered through at least two years of college). A more detailed description of item scoring can be found in Friede, Gillespie, Kim, Oswald, Ramsay, & Schmitt (2003).

In this data collection, we administered a 36-item version of the SJT (see Drzakowski et al., 2004 for a description of the selection of the 36 items from the 153-item bank). Three items reflected each of the 12 dimensions of student performance. However, earlier work on this measure did not provide evidence for the discriminant or convergent validity of the individual sets of items designed to measure each of the 12 dimensions (Oswald et al., 2004). Thus, only a single composite score was calculated.

Range Restriction Concerns. For the predictors used in our study, validity may be underestimated if lower scoring individuals were eliminated during the admissions process. Because the biodata and SJT were not collected until after the admissions process was complete and these measures were minimally correlated with HSGPA and SAT/ACT scores that were

available for use (see Table 4), we assumed there was minimal, if any, restriction of range in the scores of entry-level freshmen. For SAT/ACT scores, we examined standard deviations for the first-term college students in our sample. The SAT Verbal and Math score standard deviations were 112 and 120 respectively. Corresponding standard deviations published on the SAT website were 114 and 118. ACT score standard deviations in our sample ranged from 4.47 to 5.25 across four subtests whereas national data published on the ACT website showed standard deviations ranging from 4.6 to 6.0. These data may indicate little, if any, restriction of range in entry-level ability indices at the total sample level. We do not know the range of scores of the applicants within the various universities. Sizable differences in means and standard deviations in SAT/ACT scores across universities suggest possible variance in range restriction. College GPA data were corrected for institutional differences in entry-level ability using the procedure described in our "data analyses" section below.

A second potential range restriction problem is possible since the final sample for whom four-year outcome data are available was much smaller than the original sample (see Table 3). If there was a differential tendency on the part of students in one part of the score distribution to drop out, this could produce range restriction. Therefore, we compared the standard deviations of all variables for our total sample with the standard deviations of our final 2008 sample. Standard deviations differed by less than .04 across all variables with the exception of the standardized ability composite (i.e., SAT/ACT). The standard deviation of this composite was .92 for the first term college freshmen, but only .64 for the 2008 sample indicating that some of the less able students had not continued in college. We provide range restriction corrections to the validity data in the results section below and discuss the possible implications for other analyses.

Outcome Measures

A cumulative GPA for each student was collected from seven of the participating institutions. Because admissions policies at our different schools meant that students with widely different SAT/ACT scores were admitted, we corrected college GPA using a procedure that College Board employs in assessing the validity of the SAT in similar instances. Specifically, we first standardized the GPA variable within university. We then regressed the standardized grades across universities on the ability measure (i.e., the summed composite of SAT and ACT scores) along with a set of dummy variables representing each college and university. The coefficients for the dummy variables indicate the differences in grades that would be expected for students with comparable SAT scores at the various universities. Grades for students at each school were then adjusted by that school's regression coefficient such that students at universities with higher average SAT scores received a relatively higher adjusted college GPA, and conversely, students at universities with lower average SAT scores received a relatively lower adjusted college GPA.

The Behaviorally-Anchored Rating Scales (BARS) assessed students' self-reported performance on the 12 dimensions of college student success. The BARS provide descriptions of each dimension of success and example behaviors that reflect different levels of performance on that dimension. Respondents rate their performance on a five-point scale ranging from 1 (very low) to 5 (very high). Coefficient alpha for this 12-item scale was .77.

In the organizational behavior literature, organizational citizenship behaviors are defined as discretionary, extra-role behaviors that are not officially part of one's job, but that contribute to the overall functioning of the organization (Organ, 1988). To assess student citizenship behavior, we adapted the items created by Moorman and Blakely (1995). We adapted this measure by selecting content from three of the five subscales included in the original instrument

and altering the items such that they reflect an academic, rather than organizational, setting. The three sub-categories of citizenship behaviors that we considered relevant to university settings were interpersonal helping, loyal boosterism, and individual initiative. The two remaining subscales, compliance and sportsmanship, were excluded because we did not believe that they reflected OCBs in a college context. Interpersonal helping refers to voluntarily helping other university members. An example of an interpersonal helping item that we used is, "Helped students who have been absent from class." The loyal boosterism category refers to promoting the university to outsiders and defending it against criticism. An example of a loyal boosterism item that we used is "Actively promoted your school to people who might want to attend." The category of individual initiative refers to going above and beyond the minimum required to make the university a better place. An example of an individual initiative item that we used is, "Participated in student government or other clubs that try to make your school a better place." Two additional items created by the research team were added to the Individual Initiative subscale. These two items were, "Did things to improve your school," and "Participated in student government or other clubs that try to make your school a better place." Response options were on a five-point scale ranging from Very Infrequently/Never to Very Frequently/Always. Coefficient alpha for the resultant 10 item scale was .80.

For the class attendance outcome, we asked students to indicate "the extent to which you have missed regularly scheduled classes in the past six months." There were five response options ranging from "Missed less than 5 times" to "Missed more than 30 times." Participants were asked to self report absenteeism on two items. One item asked them to provide information on controllable absences (e.g., missed class to socialize with friends or because they found the class boring). The second question asked them to report uncontrolled absences (e.g., being sick,

an emergency). In the analyses that follow we used the controllable absence measure. This measure was positively skewed (1.41) and efforts to reduce the skew prior to analysis were not successful. Correlations with this measure were likely lower than they would have been had this measure been more nearly normally distributed.

Finally we obtained graduation data from seven schools and coded it 1 for graduation as of spring 2008 and 0 for not graduated. It is important to note that this index represented graduation in four years; many students will graduate later. It is also the case that the graduation index does not include information on students who left the original institution and graduated elsewhere (these people were coded as "not graduated"). The proportion of the sample that graduated in four years (52.1%) was almost exactly half the people for whom institutions provided archival data.

It is important to point out that some of the predictors and outcomes in this study were collected via self-report measures. Self-reported variables have a tendency to be inflated in high-stakes admissions situations; in the present study data, the sample was not high-stakes (students were already admitted into their universities) and there was considerable reliable variance in the predictor and outcome scores (i.e., they are not all high) leaving the potential for predictive relationships. The criteria were also measured several years after the predictors. The time lag in measurement not only attenuates the effects of common method bias but also that of normative implicit theories students may have about predictor-criterion relationships.

Data Analyses

For four of the five outcomes (cumulative college GPA, class attendance, OCB, satisfaction, and the BARS), we regressed the outcome on HSGPA and SAT/ACT scores in Step 1 followed by the 11 biodata measures and the SJT on Step 2. HSGPA and SAT/ACT measures

were entered first because our interest was in assessing both validity and incremental validity of the new noncognitive measures of student potential. For graduation status which was a dichotomous variable, we used hierarchical logistic regression in which the predictor variables were entered in the same manner.

To assess the degree to which potential use of noncognitive predictors might affect the demographic distribution of those students admitted we computed a simple sum of the biodata and SJT scores. This sum as well as the HSGPA and SAT/ACT variables were standardized and all three were then simply added together to form an "admissions composite." Scores on this composite were then used to rank order the students. Then we examined the ethnic composition of the top 15%, 50%, and 85% of the students simulating a very selective, moderately selective, and minimally selective admissions policy. The same procedure was used with an "admissions composite" consisting of only the SAT/ACT score and HSGPA as might be a more traditional selection strategy. A comparison of the demographic status of students "selected" under these two strategies gave us an index of the degree to which the use of a noncognitive component in admissions decisions might have an impact on the demographic composition of the student body.

A more traditional question is whether use of these predictor composites produced any differential prediction of cumulative college GPA. To answer this question, we used the procedure described originally by Cleary (1968). We regressed the cumulative college GPA on the outcomes of the three components of the admissions decision plus race plus the products of the predictors and race for each minority group separately using an analysis outlined by Bartlett, Bobko, Mosier, and Hannan (1978). Thus three regression analyses were conducted in each of which the prediction of GPA for one minority group (i.e., African, Hispanic, or Asian Americans) was compared with that for Caucasian students. In this analysis, an intercept

difference in subgroup regressions is indicated by the significance of the regression weight associated with race. A difference in subgroup regression slopes is indicated by the significance of the regression weight associated with the product of the predictor and race. This regression analysis was done only for cumulative college GPA since the other outcomes with the exception of graduation status were all collected on the follow-up survey for which subgroup sample sizes were inadequate.

Results

Means, standard deviations, and intercorrelations for the variables we assessed are presented in Table 4. The last five rows of this table contain information relevant to the set of hypothesized relationships contained in Table 2. In examining the degree to which these correlations matched our a priori expectations with respect to predictor-outcome relationships, we used a correlation above .10 as our criterion as a practically significant level of validity since most correlations in the table are statistically significant given the large sample sizes. For cumulative college GPA, the expected correlations with HSGPA and ACT/SAT scores were quite high ($r = .53$ in both instances) and consistent with past literature. In addition, correlations with the Knowledge, Continuous Learning, Responsibility, Perseverance and SJT scores were all above .10 as hypothesized. In addition, however, several other biodata scores (Artistic and Multicultural Appreciation, Health, and Ethics) exhibited sizable correlations with cumulative GPA. Unexpectedly, the Career Orientation score was negatively related (-.14) to cumulative college GPA. One possible explanation, though entirely post hoc, is that many of our African-American students were first generation college students for whom career mobility and a career orientation was a major reason for college attendance; these students also received lower GPAs. For graduation status, expected large correlations with Knowledge, Leadership, Responsibility,

and SJT as well as SAT/ACT and HSGPA were observed. Expected correlations with Career Orientation and Perseverance were not large, but two other correlations (with Artistic Appreciation and Ethics) were larger than expected. For class attendance, we expected negative correlations as this outcome was the number of self-reported absences. Consistent with our expectations, Health, Perseverance, and Ethics were related to class attendance, but we did not observe large correlations with Leadership, Responsibility, and Career Orientation as expected. Correlations with Adaptability, Knowledge, and Artistic Appreciation were higher than expected and SAT/ACT scores were actually positively related to class absenteeism though the correlation was relatively small (r = .11).

Insert Table 4 about here

Correlations with nontraditional student outcomes were different. Most of the noncognitive predictors were relatively highly correlated (.20 to .40) with BARS while the SAT/ACT and HSGPA measures were not. For the OCB measure, the four expected correlations (Leadership, Responsibility, Ethics, and SJT) were relatively high, but so also were correlations with Continuous Learning, Multicultural Appreciation, Adaptability, Perseverance, and Ethics. SAT/ACT scores were negatively related (r = -.11) to OCB.

Overall, using the .10 correlation as confirmation of our hypotheses, we found lack of support in five of the 37 instances (see Table 2) in which we expected to find relatively large correlations. In addition, we observed 15 sizable correlations we did not expect and three that were in a direction opposite than expected (i.e., Career Orientation with cumulative college GPA, SAT/ACT scores with class absenteeism, and OCB with SAT/ACT scores). These

bivariate correlations are encouraging in a predictive validity sense, but they also suggest that more work on the construct validity of the biodata measures should be conducted. The multivariate analyses described in the following section are more appropriate indices of overall predictability and the relative predictability of various sets of predictors. Hierarchical regressions of four outcomes on both cognitive and noncognitive predictors are presented in the next section.

As indicated above, there was a significant degree of range restriction in the SAT/ACT variable in the final sample relative to the sample of first term college students. We used the standard deviations of these two groups to correct the validity of this variable in predicting college GPA and graduation status (Guion, 1998, p. 315) only because correlations with the other outcomes were near zero. These corrected correlations were .67 versus .53 for college GPA and .41 versus .30 for graduation status. There was no practical difference in the standard deviations between first and fourth year students in any other variable and we did not have applicant scores on any of the measures, so no other corrections were made. We included observed correlations in the following regression analyses which should serve to underestimate the role of ability in predicting final college GPA and graduation status given the range restriction in the ability measure.

Hierarchical Regression Analyses

The results of hierarchical regression analyses of cumulative college GPA, class absenteeism, BARS, Academic Satisfaction, and OCB are presented in Table 5. In each case, SAT/ACT and HSGPA were entered at Step 1 in the analysis followed by the biodata and SJT predictors. Cumulative GPA showed the usual large R^2 for HSGPA and SAT/ACT scores (.398), but the noncognitive predictors as a set displayed incremental validity (.029) and four of the regression weights for individual scales (SJT, Continuous Learning, Health, and Career

Orientation) were statistically significant. As in the case of the bivariate correlation, the regression weight for Career Orientation was negative.

Insert Table 5 about here

For the BARS outcome, change in R^2 was significant only for the noncognitive predictors. Individual regression weights were significant for Multicultural Appreciation, Health, Perseverance, and Ethics. Bivariate correlations for the remaining noncognitive measures were all in the .20s and .30s so nonsignificant regression weights for the remaining predictors is certainly a function of the relatively high intercorrelations of the predictor set.

The cognitive predictors were statistically significantly related to OCB and Absenteeism but at much lower levels than was the case for GPA. In the case of OCB, the SAT/ACT score was a significant predictor, but negatively so. Noncognitive predictors produced an R^2 change of .201; individual regression weights for Artistic and Multicultural Appreciation, Leadership, Responsibility, and Adaptability were statistically significant ($p < .05$). SAT/ACT scores were significantly related to class absenteeism, but the noncognitive measures were the most important predictors of absenteeism. The overall R^2 for Absenteeism (.149) was lower than that for the other outcomes, but regression weights for four of the noncognitive predictors (Leadership, Health, Ethics, and SJT) were statistically significant.

Because the scores on some of the biodata scales are highly correlated (see Table 3), the beta coefficients in the hierarchical and logistic regressions should be interpreted with caution. Examples can be found for each of the regression models in which the sign of the betas is not in the expected direction. The zero-order correlations, however, are generally in the

expected direction. The fact that the sign of the beta coefficients is opposite that of the correlations is due to collinearity effects. Despite the multicollinearity among the scores on the biodata scales, the models still indicate how well the entire set of biodata measures predicts the outcome variables of interest. We believe the most interpretable index of an individual predictor-outcome relationship is the correlation coefficient.

We also conducted hierarchical analyses in which we reversed the order of the biodata and SJT measures in the second and third steps of a hierarchical regression analysis. In all cases the SJT added incrementally ($p < .05$) to the prediction of outcome variables above the prediction afforded by SAT/ACT, HSGPA, and the biodata variables. Likewise, the biodata added incrementally to the prediction of outcome variables after SAT/ACT, HSGPA, and SJT had been entered.

Graduation Status

Because Graduation Status was a dichotomous variable, we used logistic regression to examine its predictability. The results of the logistic regression analysis are presented in Table 6. In addition to high school GPA and ACT/SAT, graduation in four years was predicted by Continuous Learning, but negatively so. The correlation, though, between Continuous Learning and graduation was positive (see the last row of Table 4), implying the negative regression weight is a function of colinearity. The odds ratios reported in Table 6 indicate the increased likelihood (or decreased likelihood in the case of values under 1.00) of graduation. For example, for each standard deviation change in HSGPA, the students in our sample were nearly four times likely to graduate from college. The odds ratio for SAT/ACT was only 1.304, but this lower ratio was certainly partly a function of the correlation between SAT/ACT and HSGPA. Of the noncognitive predictors, the odds ratio for the SJT was the largest (1.355), but it was not

statistically significant. The last row of Table 4 contains the correlations of all variables with graduation status. As can be seen, several of the correlations between graduation status and noncognitive predictors are between .10 and .15. Finally, the graduation status of these students using the results of the logistic regression was successfully predicted by HSGPA and SAT/ACT in 62.9% of the cases (recall that 52.1% graduated which is the correct prediction rate if there were no relationship between graduation and the two predictors). Adding the noncognitive predictors to the logistic regression did produce a statistically significant change in the correct prediction rate to 65.0%. Graduation rates across subgroups did vary; 81, 79, 61 and 66 per cent of the Hispanic, Asian, African American and Caucasian groups graduated in the four years covered by our data collection.

Insert Table 6 about here

Impact of the Use of Noncognitive Predictors on Admissions of Members of Different Groups

To assess what implication the use of noncognitive measures might have on the proportion of different groups admitted to these universities (assuming these students were actually applicants), we computed two composites. The first composite consisted of a sum of the standardized SAT/ACT scores and standardized HSGPA. The second composite consisted of the standardized sum of the 12 noncognitive measures (11 biodata scores and the SJT) as well as the standardized measures of HSGPA and SAT/ACT. The latter composite, then, represents an approximately equally-weighted sum of noncognitive measures, HSGPA, and SAT/ACT scores. These two composites were then used to rank order the participants in our research. We then computed the proportion of each of four groups (students of Hispanic, Asian, African and

Caucasian descent) who would have been admitted to these universities if the universities admitted the top 15%, top half, or top 85% of these students. We also computed the proportion of each of these four groups who graduated in four years using these two admissions strategies and the average cumulative GPA of those who did graduate. Obviously, these analyses are hypothetical since all these students were admitted to their respective universities and in the case of the noncognitive measures, responses were not made in the usual high-stakes admissions context.

The results of these analyses are presented in three sections in Table 7. In Table 7A, we see that the use of a battery that includes a noncognitive composite will have little impact on the proportion of students in different groups that are admitted when the university is not very selective (i.e., the top 85% are admitted). However, the proportion of Hispanic and African American students who are admitted differs increasingly as a university becomes more selective. Proportions of Caucasian and Asian students are correspondingly smaller if one incorporates noncognitive measures into the admissions procedure as the admissions strategy becomes more selective.

Insert Table 7 about here

Tables 7B and 7C address the question of how different students admitted under these various strategies will perform once they are admitted. In Table 7B, we present the average cumulative college GPA of students admitted using the two different batteries of admissions procedures under various levels of selectivity. Comparing the average GPA of students admitted under the two sets of admissions procedures in a relatively unselective situation reveals very

little difference. There are small differences in average college GPA in more selective situations (the largest is the .10 difference for Hispanics in the highly selective situation, d = .25). A computation of overall GPA across all demographic groups revealed that the total group of students selected at each of the three levels of selectivity by the cognitive only or cognitive plus strategies differed only by very small amounts. Those students selected under high and moderate levels of selectivity using only cognitive tests would have GPAs that are larger by .018 and .006 relative to a strategy that included both cognitive and noncognitive tests. At very low levels of selectivity, use of both types of tests actually resulted in slightly superior GPA (difference equaled .03) over use of cognitive tests only. Table 7C shows the proportion of members of the various ethnic groups that graduated in four years. One obvious difference in this table is that the proportion of Caucasian students who have graduated at each level of selectivity is notably smaller than that of the other groups. This may be due to differences in the universities that members of different groups attended as a large number of Asians and African American students went to private universities whereas Caucasian students were more likely to attend large public universities. The important comparison, though, for purposes of this paper is the proportion who graduated based on different admissions criteria. There are almost no differences in graduation rates when students are admitted using either of the two batteries of tests. The largest difference between the graduation rates for the two batteries is for African Americans at the most selective level. At this level, more African American students would graduate if they were selected using the traditional cognitive measures than with the new composite measure. This difference, however, does not exist at the less selective levels or for the other subgroups. The generally increasing proportion that graduate under increasingly selective admissions strategies reflects the level of validity of both batteries of tests in the prediction of four-year

graduation. At the very least, graduation rates would be unaffected by this use of noncognitive and cognitive predictors while increasing the diversity of the student body.

In these analyses, we have focused on the traditional student outcomes of graduation and cumulative college GPA. This was done because most universities likely consider these the most important student outcomes. Similar analyses of the other outcomes would likely have demonstrated much larger differences between purely cognitive measures and a mixed set of instruments since the noncognitive tests were more predictive of these outcomes (see Tables 4 and 5).

It is also the case that we computed unit-weighted composites of the predictor variables. Other weighting systems taking account of the intercorrelations among the predictors or using other weighting schemes that reflect different institutional values would certainly produce a different set of outcomes (DeCorte, Lievens, & Sackett, 2007).

Differential Prediction

Our final attempt to evaluate the use of noncognitive tests in college admissions involved an analysis of the differential prediction (Cleary, 1968) of the performance of students in the four ethnic groups. In this analysis, we used the standardized noncognitive composite, SAT/ACT, and HSGPA as predictors. Each of these was multiplied by a dichotomous Caucasian versus minority group (African, Asian or Hispanic descent) variable. Then college GPA was regressed on the set of predictors (Step 1) and race (Step 2), and the set of three product terms (Step 3) following Bartlett et al. (1978). This was done separately for each of the three minority groups. The results of these analyses are presented in Table 8. In these analyses, a significant interaction term indicates a difference in subgroup slopes while a significant race effect indicates a difference in subgroup intercepts.

Insert Table 8 about here

These analyses reveal some evidence of differential prediction, particularly for African American students. The race by HSGPA interaction was statistically significant and negative indicating that the slope of the relationship between college GPA and HSGPA was less positive for African American students than for Caucasian students. The reverse was true for the race by SAT/ACT interaction; that is, the slope of the relationship between college GPA and SAT/ACT scores was more positive for African American students than for Caucasian students. The negative regression weight for race in the African American-Caucasian analysis would indicate some overprediction of African American students' grades though this is, of course, modified by the two significant interactions. There was no evidence of differential prediction for Hispanic and Asian American students though the race by HSGPA interaction was negative and relatively large in both instances. These results were consistent with the validity coefficients for SAT/ACT and HSGPA for the various subgroups. Validities of these variables were much higher for African American students than for students in the other three subgroups while the validity of HSGPA was somewhat higher for Caucasian students than for students in the other groups.

In all three analyses, the role of SAT/ACT and HSGPA in predicting cumulative college GPA was more important than that of the noncognitive component as was expected from previous analyses. Since SAT/ACT and HSGPA were relatively highly correlated we see evidence of a suppressor effect for these two predictors in the Hispanic and African American analyses. It should also be noted that these analyses were limited in that members of the three minority groups were relatively small (73, 62, and 25 for African, Asian, and Hispanic American

students respectively) especially when compared with the sample size for Caucasian students (N = 934). We were unable to obtain cumulative GPA for students from two universities at which large numbers of African and Hispanic American students had originally enrolled. Similar analyses against first year college GPA available from the senior author with much larger numbers (566, 130, and 167 for African, Hispanic and Asians respectively) yielded very similar results. Analyses of the differential prediction of the alternative outcomes were very different as would be expected given the results presented in Tables 4 and 5. Differential prediction analyses of the BARS outcome, for example, indicated that only the noncognitive composite was a valid predictor and that only for Asians was there some evidence of overprediction of self-rated performance on the BARS dimensions relative to Caucasian students.

Finally, we provide standardized subgroup mean differences for all predictor and outcome variables in Table 9. The referent group in all cases was the Caucasian student group. These means were computed for those students for whom data were available on all variables so subgroup sample sizes are relatively small (N=22, 63, 36, and 364 for Hispanics, Asians, African Americans, and Caucasians respectively). However, standardized subgroup differences for the much larger original sample contained in parentheses in Table 9 are nearly the same. For the SAT/ACT measures and HSGPA measures, the usual large subgroup differences are observed for Caucasian-African American and Caucasian-Hispanic comparisons with the latter difference being somewhat smaller. For both of these measures, Asian students' scores are slightly higher than Caucasian students. Consistent with prior estimates of subgroup differences on noncognitive measures (e.g., Roth, & Potosky, 1999; Hedlund et al., 2006), subgroup differences observed on noncognitive measures are generally much smaller than the differences on the cognitively loaded predictors, though some may be practically important. Asians and Hispanic

students achieve higher scores on the Multicultural Appreciation dimension than do Caucasians and African Americans and African Americans are more career-oriented than other subgroups. On the outcome measures, African Americans receive college grades that are one more than one standard deviation lower than those of Caucasians while Asians receive higher grades. African Americans also self report lower performance (BARS) than do the other groups.

Depending on how predictors are used in an admissions context, the non-cognitive measures could produce lower selection rates among African Americans, but the impact would be substantially less than that produced by use of the SAT/ACT or HSGPA measures alone. It should be noted that the impact of supplementing cognitive measures with non-cognitive measures depends on several factors including the number of predictors, the magnitude and direction of subgroup differences, the intercorrelation of the noncognitive predictors, and the correlation between the cognitive and noncognitive measures. There is a potential scenario in which subgroup differences could be exacerbated by the addition of non-cognitive measures. For example, when a non-cognitive measure produces subgroup differences that are smaller but in the same direction as a cognitive measure, and the measures are not highly correlated, a composite may exhibit larger subgroup differences than either measure alone (Sackett & Ellingson, 1997). While this was not the case in our study (i.e., the present set of predictors showed the potential to reduce subgroup differences), practitioners should generally remain cognizant of these factors when attempting to recruit a more diverse set of individuals through the use of non-cognitive selection tools.

Discussion

The results reported in this paper replicated the finding that SAT/ACT scores and HSGPA have relatively high levels of validity for the prediction of college grades, even

cumulative grades over four years in college. While previous meta-analyses (Hezlett et al. 2001; Kunzel et al., 2004; Sackett et al. 2009) have included a great many studies reporting the validity of cognitive predictors (such as test scores and HSGPA) in the prediction of first year college GPA, studies of the prediction of cumulative grades over four years are not as numerous. Validity for both HSGPA and SAT/ACT was .53 for our complete sample. This study demonstrated that objectively-scored noncognitive measures in the form of biodata and SJT added incrementally to the prediction of undergraduate college GPA though responses to these measures was not made in high-stakes situation in which motivation to fake would exist. These results are consistent with employee selection research on the SJT that indicates incremental validity of the SJT over cognitive predictors (McDaniel et al., 2007). Also, research in a medical school context (Lievens, Buyse, & Sackett, 2005; Lievens & Sackett, 2007) indicated that the SJT had incremental validity over cognitive predictors in predicting performance in an interpersonally-oriented curriculum. The level of incremental prediction of college GPA observed in the study reported in this paper was not large, but could certainly be practically important. Also, a number of correlations with college GPA were relatively large including those for the biodata measures of Knowledge, Artistic Appreciation, and Ethics as well as the SJT. Correlations with graduation status at the end of four years were similar in pattern, but smaller, partly because the outcome variable was dichotomous. Odds ratios from a logistic regression showed that the most important predictor of graduation status was HSGPA.

One major purpose of our study was to demonstrate that if the set of criteria used to assess college student performance is expanded, these criteria would be better predicted by the noncognitive measures we constructed. The BARS instrument was constructed as a direct measure of the same dimensions derived from our examination of the student performance goals

that universities indicate are important to them. Our noncognitive predictors were all significantly and relatively highly related (> .20) to the BARS and cognitive variables were not. While this variable was a self report of performance, it was collected nearly four years after the predictor variables and in a different format. The student version of OCB was predicted nearly as well; the individual correlations with the SJT, Leadership, Responsibility, Perseverance, and Adaptability were particularly large. Correlations with Controllable Absenteeism were not as large, but Perseverance, Ethics and the SJT exhibited relatively high levels of validity in the prediction of Absenteeism. If these outcomes are important to universities as their literature implies, the admission of students with high scores on these noncognitive measures would certainly result in superior student performance on these dimensions.

In addition to the examination of validity, we examined the degree to which use of an admissions battery that included both cognitive and noncognitive components would affect the demographic composition of admitted students assuming various levels of selectivity and that our sample was representative of applicants. This is not completely the case since our participants were already admitted students, but the use of students as opposed to applicants is most likely to underestimate observed relationships because of range of restriction in the predictors. These analyses indicated that a greater proportion of Hispanic and African American students would be admitted using both cognitive and noncognitive measures as opposed to a simple sum of HSGPA and SAT/ACT scores. Correspondingly lower proportions of Asian and Caucasian American students would have been admitted when using the set of three predictors. This trend is more evident under conditions of greater selectivity. Because cognitive variables were more highly related to college GPA than were the noncognitive variables, one would expect that there would be some decrement in average college GPA when noncognitive variables are

used to make admissions decisions particularly when only very able students are admitted. This is generally true, but only very small differences in average GPA were observed. There was also a very small difference in the proportion of students who graduated in four years using the two different batteries of admissions procedures. The conclusion of this combination of analyses is that a university could use noncognitive indices to make admissions decisions and increase the diversity of its student body while having little or no negative impact on their academic performance. In addition, there could be relatively large positive changes in other spheres of student activity and performance.

As mentioned in the introduction, the problems of academic admissions officers are similar to those encountered in personnel selection though research efforts in these two arenas are relatively isolated (Sackett et al., 2001). We find cognitive measures quite predictive of performance, but use of these measures as the sole gatekeepers will result in lower admissions rates for some minority groups (see Table 7). Use of noncognitive measures as part of a selection or admissions procedure will alleviate, but not remove, subgroup differences in composite scores and hence differential selection rates (Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). There would have been a minor decrement in college GPA in this instance; the magnitude of this decrement will vary as a function of the selectivity of the university as well as the weighting of the various admissions components. Considering other outcomes will certainly make this "tradeoff" less negative, and perhaps positive, if the university values a broader set of outcomes. The situational judgment measure which has been primarily used in employment selection contexts proved to be a significant contributor to the prediction of several important student outcomes. Rationally derived biodata measures proved to be related to outcome measures in ways that were specified a priori supporting the notion that these measures can be used to

measure identifiable and interpretable constructs which has long been a debatable issue in employment selection arenas (Mitchell & Klimoski, 1982; Mumford & Stokes, 1992; Stokes & Cooper, 2001). However, inconsistencies between the expected and observed correlations between biodata and outcome measures suggests that more work on the contruct validity of the biodata measures should be conducted. As is the case in personnel selection, there does seem to be more acceptance of the notion that GPA is not the only standard against which student success should be judged and, if so, then different admissions instruments may provide valuable information about student potential.

One predictor, Career Orientation, was unexpectedly negatively related to GPA. One post hoc speculation regarding this relationship is that it is due to the fact that many of the first-generation, mostly minority students in our sample went to college primarily because they felt this was the means to better employment possibilities. These students also were less well prepared academically than their peers. While this may account for the negative correlation, we did not have the data to confirm or disconfirm the possibility. There may be other explanations as well.

Limitations and Future Research

Continued research on the issues addressed in this paper should address both limitations and additional questions. In several places in this article, we alluded to school differences in student characteristics and policies. These issues should be addressed with a larger set of universities and in multilevel analyses. Differences in school policies, goals and climate as well as school differences in the type of students enrolled may all influence the nature of some of the predictor-outcome relationships observed in the current study.

Even though our initial sample included relatively large numbers of members of different racial subgroups, the unfortunate loss of participants from two schools as well as normal attrition severely limited the number of students in minority groups for whom we had final college GPA and graduation status. Issues of representativeness of the sample at the national level, much less an international one, and generalizability are concerns. Future work with larger groups of minority students and more and different samples of participants would be valuable.

The data collected in this study are longitudinal in the sense that all predictors were collected before or immediately after students arrived at their universities. However, the noncognitive predictors were not used to make admissions decisions and participants knew that was the case. Hence, additional work in which these data are collected during the application process should be collected. This is important for purely psychometric reasons (e.g., range restriction concerns) as well as motivational issues. Participant motivation to "fake" or distort their responses to gain admission would likely be greater in that situation. More broadly, the possibility that students may fake these measures remains an important concern about their implementation and may suggest their use in a counseling/guidance as opposed to admissions context.

Several of our alternative outcomes (BARS, OCBs, and class absenteeism) were self reports of performance. This may have served to inflate relationships between noncognitive predictors and BARS and OCBs, but the data were collected nearly four years apart and in a different format (web vs. paper-and-pencil and on different scales). For class absenteeism, there was little motivation on the part of students to inflate or deflate their reports. Although we used actual GPA provided by institutions, it correlated over .9 with self reports of GPA that we also collected suggesting that social desirability did not heavily influence scores in this study.

Consistent with our findings, the Kuncel, Crede, and Thomas (2005) meta-analysis reported a correlation of .90 between college GPA obtained from school records and self-reported college GPA across twelve studies though there were significant differences in means. Student self reports of subjective information such as personality have generally also been found to correlate highly with parent reports (e.g., Barbaranelli et al., 2008; Parker & Stumpf, 1998). However, attempts to gather information about these alternative outcomes from peers and/or faculty members may be a very informative adjunct to the self reports reported in this article.

Practical Implications

The results reported in this paper would suggest that admissions personnel could use the biodata and SJT to predict who will perform well in different areas of relevance to most universities including students' academic performance. In addition to seeking admissions measures with predictive validity, it has also been suggested that colleges evaluate assessment instruments on their ease of administration, costs, and efficiency (Camara, 2005). One aim in developing these measures was to make it easier for universities with large applicant pools to evaluate noncognitive aspects of student potential which are usually assessed with essays, interviews, letters of reference, etc. The biodata instrument appears to represent conceptually some of the same constructs (e.g., interests, hobbies). The SJT represents judgments students make in similar domains to those that are likely the target in university attempts to measure noncognitive domains. Future research might consider the incremental validity of the biodata and SJT over more traditional noncognitive measures. An additional consideration for future research is the extent to which our noncognitive measures reduce subgroup differences in the presence of other noncognitive admissions components. Simulation by Sackett and Ellingson (1997) and Schmitt et al. (1997) demonstrate that the magnitude of d associated with a composite

of predictors is a complex function of the number of predictors, their intercorrelation , and the level of d demonstrated by the individual predictors.

One significant issue concerning the biodata and SJT instruments remains to be addressed and that is the degree to which scores might be inflated if they are actually used to make high-stakes decisions and test preparation courses become available. Certainly students can and do obtain help in preparing essays or interview statements and may be able to influence who writes letters of reference (Willingham & Breland, 1982). It is not clear to what degree a multidimensional biodata form would be fakable if a university desires a given profile of scores as opposed to high scores on a set of dimensions. While difficult to conduct, an evaluation of these procedures in a high-stakes situation is clearly required. At this point, faking and the impact of coaching programs may be a major reason why organizations such as the College Board have not proceeded to promote the use of measures such as these in actual admissions decisions.

Another practical implication of our results is that alternative student outcomes that most universities espouse are not highly related to academic performance and are not highly predicted by ability measures. If one is to maximize student performance on these outcomes, noncognitive measures are likely to be more valid than ability. Similarly, use of noncognitive measures in combination with traditional indices of ability will result in some increases in the diversity of the student population.

Conclusions

The results of the research reported in this article indicate that biodata and SJT relate in predictable and important ways to several student outcomes even over a four year period. They add incrementally to the prediction of college GPA and graduation status over and above

HSGPA and SAT/ACT and are the major predictors of other student performance outcomes. Our data indicate that use of these noncognitive predictors would diversify the student body with minor changes in student performance though data on the latter point are limited by sample size.

References

Ackerman, P. L. (1989). Within task intercorrelations of skilled performance: Implications of predicting individual differences. *Journal of Applied Psychology, 74*, 360-364.

Anastasi, A., Meade, M.J., & Schneiders, A.A. (1960). *The validation of a biographical inventory as a predictor of college success: Development and validation of the scoring key*. New York: College Entrance Examination Board.

Barbaranelli, C., Fida, R., Paciello, M., Di Giunta, L., & Caprara, G. V. (2008). Assessing personality in early adolescence through self-report and other-ratings a multitrait-multimethod analysis of the BFQ-C. *Personality and Individual Differences, 44*, 876-886.

Bartlett, C. J., Bobko, P., Mosier, S. B., & Hannan, R. (1978). Testing for fairness with a moderated multiple regression strategy: An alternative for differential analysis. *Personnel Psychology, 31*, 233-241.

Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology, 74,* 478-494.

Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology, 52*, 561-589.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71-98). San Francisco: Jossey-Bass.

Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: The meaning for personnel selection research. *Human Performance, 10*(2), 99-109.

Boyer, S. P., & Sedlacek, W. E. (1988). Non-cognitive predictors of academic success for international students: A longitudinal study. *Journal of College Student Development, 29*, 218-223.

Breland, H. M. (1998). *National trends in the use of test scores in college admissions*. Princeton, NJ: Educational Testing Service.

Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). *Predictions of freshman grade point average from the revised and recentered SAT I: Reasoning Test* (College Board Rep. No. 2000-1). New York: College Entrance Examination Board.

Camara, W. J. (2005). Broadening predictors of college success. In W. J. Camara & E. W. Kimmel (Eds.), *Choosing Students: Higher education admissions tools for the 21st century*. Mahwah, NJ: Lawrence Erlbaum Associates.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W.C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement, 5*, 115-124.

Clifton, T. C., Mumford, M. D., & Baughman, W. A. (1999). Background data and autobiographical memory: Effects of item types and task characteristics. *International Journal of Selection and Assessment, 7*, 57-71.

Crede, M., & Kuncel, N. R.(2008). Study Habits, Skills, and Attitudes: The Third Pillar Supporting Collegiate Academic Performance. *Perspectives on Psychological Science, 3*, 425-453.

Dalal, R. S. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology, 90,* 1241-1255.

DeCorte, W., Lievens, F., & Sackett, P. R. (2007). Combining predictors to achieve optimal trade-offs between selection quality and adverse impact. *Journal of Applied Psychology, 92,* 1380-1393.

Dorans, N. J., Lyu, C. F., Pommerich, M., & Houston, W. M. (1997). Concordance between ACT assessment and recentered SAT I sum scores. *College and University, 73*(2), 24-35.

Drzakowski, S., Friede, A., Imus, A., Kim, B., Oswald, F. L., Schmitt, N., & Shivpuri, S. (2005). *Assessment of Life Experiences Questionnaire (ALEQ) Biodata Measure and Life Events Assessment and Development (LEAD) Situational Judgment Inventory: Results of an expanded data collection across ten U.S. colleges and universities*. East Lansing: Michigan State University. Report submitted to College Board.

Friede, A. J., Gillespie, M. A., Kim, B. H., Oswald, F. L., Ramsay, L. J., & Schmitt, N. (2003). *Summary of the Interim and Final Reports: Biodata and Situational Judgment Inventories as Measures of College Success.* E. Lansing, MI, Michigan State University. Report submitted to College Board.

Gillespie, M. A., Kim, B. H., Manheim, L. J., Yoo, T.Y., Oswald, F. L., & Schmitt, N. (2002). The development and validation of biographical data and situational judgment inventories in the prediction of college student success. Paper presented at the annual meeting of the American Psychological Society. New Orleans, LA.

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: Lawrence Erlbaum.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94,* 562-575.

Hedlund, J., Wilt, J. M., Nebel, K. L., Ashford, S. J., & Sternberg, R. J. (2006). Assessing practical intelligence in business school admissions: A supplement to the graduate management admissions test. *Learning and Individual Differences, 16,* 101-127.

Henry, R. A., & Hulin, C. L. (1989). Changing validities: Ability-performance relations and utilities. *Journal of Applied Psychology, 74,* 365-367.

Hezlett, S. A., Kuncel, N. R., Vey, M. A., Ahart, A. M., Ones, D. S., Campbell, J. P., & Camara, W. (2001, April). The predictive validity of the SAT: A meta-analysis. Paper presented in D. Ones & S. Hezlett (Chairs), *Predicting performance: The interface of I-O psychology and educational research.* Symposium presented at the 16th Annual Convention of the Society for Industrial and Organizational Psychology, San Diego, CA.

Hough, L. M. (1998). Personality at work: Issues and evidence. In M. D. Hakel (Ed.), *Beyond multiple choice: Evaluating alternatives to traditional testing for selection.* Mahwah, NJ: Erlbaum.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment, 9,* 152-194.

Kuncel, N.R., Crede, M., & Thomas, L.L. (2007). A comprehensive meta-analysis of the predictive validity of the Graduate Management Admission Test (GMAT) and undergraduate grade point average (UGPA). *Academy of Management Learning and*

Education, 6, 51-68.

Kuncel, N. R., Crede, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75,* 63-82.

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

Kuncel, N.R., Hezlett, S.A., & Ones, D.S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127,* 162–181.

Kuncel, N.R., Hezlett, S.A., & Ones, D.S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all?. *Journal of Personality and Social Psychology, 86,* 148–161.

Landau, E. (2008). More colleges move toward optional SATs, May 30, www.cnn.com.

Lievens, F., Buyse, T., & Sackett, P. R. (2005). The operational validity of a video-based SJT for medical college admissions: Illustrating the importance of matching predictor and criterion domains. *Journal of Applied Psychology, 90,* 442-452.

Lievens, F., & Sackett, P. R. (2007). Situational judgment tests in high-stakes settings: Issues and strategies with generating alternate forms. *Journal of Applied Psychology, 92,* 1043-1055.

Mael, F. A. (1991). A conceptual rationale for the domain and attributes of biodata items. *Personnel Psychology, 44,* 763-792.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L., III. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology, 60*, 63-91.

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Predicting job performance using situational judgment tests: A clarification of the literature. *Journal of Applied Psychology, 86*, 230-24.

Melancon, J.G. (2002). Reliability, structure, and correlates of Learning and Study Strategies Inventory scores. *Educational and Psychological Measurement, 62*, 1020–1027.

Mitchell, T. W., & Klimoski, R. J. (1982). Is it rational to be empirical? A test of methods for scoring biographical data. *Journal of Applied Psychology, 67*, 311-317.

Moorman, R. H. & Blakely, G. L. (1995). Individualism-collectivism as an individual difference predictor of organizational citizenship behavior. *Journal of Organizational Behavior, 16*(2), 127-142.

Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75*, 640-647.

Mumford, M. D., Connelly, M. S., & Clifton, T. C. (1990). *The meaning of life history measures: Implications for scaling strategies*. Unpublished report. Fairfax, VA: George Mason University, Center for Behavioral and Cognitive Studies.

Mumford, M. D., & Owens, W. A. (1987). Methodology review: Principles, procedures, and findings in the application of background data measures. *Applied Psychological Measurement, 11*, 1-31.

Mumford, M. D., & Stokes, G. S. (1992). Developmental determinants of individual action: Theory and practice in applying background measures. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, pp. 61-138). Palo Alto, CA: Consulting Psychologists Press.

National Association for College Admissions Counseling. (2008). *Report of the Commission on the Use of Standardized Tests in Undergraduate Admissions*. Arlington, VA: Author.

Nickels, B. J. (1994). The nature of biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook* (pp.1-16). Palo Alto, CA: CPP Books.

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college success. *Journal of Applied Psychology, 89,* 187-208.

Owens, W. A. (1976). Background data. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 609-644). Chicago: Rand McNally.

Owens, W. A., & Schoenfeldt, L. F. (1979). Towards a classification of persons. *Journal of Applied Psychology, 64,* 569-607.

Pantages, P. J., & Creedon, C. F. (1978). Studies of college attrition: 1950-1975. *Review of Educational Research, 48,* 49-101.

Parker, W. D., & Stumpf, H. (1998). A validation of the five-factor model of personality in academically talented youth across observers and instruments. *Personality and Individual Differences, 25,* 1005-1025.

Payne, D. A., Rapley, F. E., & Wells, R. A. (1973). Application of a biographical data inventory to estimate college academic achievement. *Measurement and Evaluation in Guidance, 6*(3), 152-156.

Ployhart, R. E., Schneider, B., & Schmitt, N. (Eds.). (2006). *Staffing organizations: Contemporary practice and theory*. Mahwah, NJ: Erlbaum.

Richards, J.M., Jr., Holland, J.L., & Lutz, S.W. (1967). Prediction of student accomplishment in college. *Journal of Educational Psychology, 58*, 343-355.

Ridgell, S., & Lounsbury, J.W. (2004). Predicting collegiate academic success: General intelligence, "Big Five" personality traits, and work drive. *College Student Journal, 38*, 607–618.

Rigol, G. W. (2003). *Admissions decision-making models: How U. S. institutions of higher education select undergraduate students*. NY: College Board.

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.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology, 50*, 707-721.

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

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M.B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative action world. *American Psychologist, 56*, 302-318.

Schmidt, F. L, & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262-274.

Schmitt, N., & Chan, D. (2006). Situational judgment tests: Method or construct? In J. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests* (pp.135-156). Mahwah, NJ: Erlbaum.

Schmitt, N., Gillespie, M., Kim, B., Manheim, L., Oswald, F., & Yoo, T. Y. (2003) Impact of elaboration on social desirability and the validity of biodata measures. *Journal of Applied Psychology, 88*, 979-989.

Schmitt, N., Oswald, F. L., Kim, B. H., Imus, A., Drzakowski, S., Friede, A., & Shivpuri, S. (2007). The use of background and ability profiles to predict college student outcomes. *Journal of Applied Psychology, 92*, 165-18.

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency using various predictor combinations. *Journal of Applied Psychology, 82*, 719-73.

Stokes, G. S., & Cooper, L. A. (2001). Content/construct approaches in developing life history measures for selection. *International Journal of Selection and Assessment, 9*, 1-14.

Wagner, R. K., & Sternberg, R. J. (1991). *The common sense manager*. San Antonio: Psychological Corporation.

Zeegers, P. (2001). Approaches to learning in science: A longitudinal study. *British Journal of Educational Psychology, 71*, 115–132.

Zimmerman, W.S., Parks, H., Gray, K., & Michael, W.B. (1977). The validity of traditional cognitive measures and of scales of the Study Attitudes and Methods Survey in the prediction of the academic success of Educational Opportunity

Program students. *Educational and Psychological Measurement*, *37*, 465–470.

Table 1

Conceptual Definitions of Student Performance Dimensions Represented in the Biodata scales, the SJT, and the Self-rating BARS Instrument

---

Knowledge and mastery of general principles (Knowledge)
Gaining knowledge and mastering facts, ideas and theories and how they interrelate, and the relevant contexts in which knowledge is developed and applied. Grades or GPA can indicate, but not guarantee, success on this dimension.

Continuous learning, and intellectual interest and curiosity (Learning)
Being intellectually curious and interested in continuous learning. Actively seeking new ideas and new skills, both in core areas of study as well as in peripheral or novel areas.

Artistic and cultural appreciation (Artistic)
Appreciating art and culture, either at an expert level or simply at the level of one who is interested.

Appreciation for diversity (Diversity)
Showing openness, tolerance, and interest in a diversity of individuals and groups (e.g., by culture, ethnicity, religion, or gender). Actively participating in, contributing to, and influencing a heterogeneous environment.

Leadership (Leadership)
Demonstrating skills in a group, such as motivating others, coordinating groups and tasks, serving as a representative for the group, or otherwise performing a managing role in a group.

Interpersonal skills (Interpersonal)*
Communicating and dealing well with others, whether in informal social situations or more formal school-related situations. Being aware of the social dynamics of a situation and responding appropriately.

Social responsibility and citizenship (Responsibility)
Being responsible to society and the community, and demonstrating good citizenship. Being actively involved in the events in one's surrounding community, which can be at the neighborhood, town/city, state, national, or college/university level. Activities may include volunteer work for the community, attending city council meetings, and voting.

Physical and psychological health (Health)
Possessing the physical and psychological health required to engage actively in a scholastic environment. This would include participating in healthy behaviors, such as eating properly, exercising regularly, and maintaining healthy personal and academic relations with others, as well as avoiding unhealthy behaviors, such as alcohol/drug abuse, unprotected sex, and ineffective or counterproductive coping behaviors.

Table 1 (cont'd)

Conceptual Definitions of Student Performance Dimensions Represented in the Biodata scales, the SJT, and the Self-rating BARS Instrument

Career orientation (Career)	Having a clear sense of career one aspires to enter into, which may happen before entry into college, or at any time while in college. Establishing, prioritizing, and following a set of general and specific career-related goals.
Adaptability and life skills (Adapt)	Adapting to a changing environment (at school or home), dealing well with gradual or sudden and expected or unexpected changes. Being effective in planning one's everyday activities and dealing with novel problems and challenges in life.
Perseverance (Persevere)	Committing oneself to goals and priorities set, regardless of the difficulties that stand in the way. Goals range from long-term goals (e.g., graduating from college) to short-term goals (e.g., showing up for class every day even when the class isn't interesting).
Ethics and integrity (Ethics)	Having a well developed set of values, and behaving in ways consistent with those values. In everyday life, this probably means being honest, not cheating (on exams or in committed relationships), and having respect for others.

Note. Abbreviations for each dimension appear in parentheses; these abbreviations are used in subsequent tables. *The Interpersonal Skills scale was not incorporated into our profiling or regression analyses due to a lack of internal consistency and high intercorrelations with the other biodata scales.

Table 2

Hypothesized Relationships between Predictor and Outcome Variables

			Class		
Predictor	GPA	Grad Status	Absenteeism	BARS	OCB
Knowledge	X	X		X	
Cont. Lrng.	X			X	
Art. Apprec.				X	
Multic.				X	
Leadership		X	X	X	X
Responsibility	X	X	X	X	X
Health			X		
Car. Ornt.		X	X	X	
Adaptability				X	
Perseverance	X	X	X	X	
Ethics			X	X	X
SJT	X	X	X	X	X
SAT/ACT	X	X			
HSGPA	X	X			

Note. SJT = situational judgment test, OCB = organizational citizenship behavior, BARS = behaviorally anchored rating scale, Cont. Lrng = Continuous Learning, Multic. = Multicultural Appreciation, Car. Ornt. = Career Orientation.

Table 3

Demographics for 2004 and 2008 Samples

	2004 Sample		2008 Self-report Sample		2008 GPA Sample	
	N	%	N	%	N	%
Gender						
Male	990	34.8	221	34.6	632	33.3
Female	1772	62.3	418	65.4	1265	66.7
Race/Ethnicity						
White	1530	53.8	437	69.5	1211	64.0
Black	683	24.0	50	7.9	346	18.3
Asian	209	7.4	84	13.4	161	8.5
Hispanic	162	5.7	19	3.0	67	3.5
Native American/ Pacific Islander	22	.7	2	.3	10	.6
Other	143	5.0	37	5.8	97	5.1
Age						
18	2445	86.0	0	0		
19	259	9.1	0	0		
20	28	1.0	2	.3		
21	7	.2	253	39.5	1670	88.2
22	3	.1	365	56.9	196	1.3
23 and above	11	.5	21	3.2	28	1.6

Table 4

Means, Standard Deviations, and Intercorrelations between Study Variables[a]

	N	Mean	SD	High School GPA	Ability	Knowledge	Learning	Artistic	Multicultural	Leadership	Responsibility	Health	Career
High School GPA	2532	3.50	.43	*1.00*	.58	.39	.07	.15	.06	.17	.24	.17	-.09
Ability Score (ACT/SAT)	2589	.61	.92	.58	*1.00*	.32	.19	.28	.14	.15	.20	.15	-.25
Biodata - Knowledge	2765	3.15	.47	.32	.26	*.67*	.65	.36	.37	.36	.35	.42	.32
Biodata – Learning	2765	3.09	.61	.06	.17	.47	*.78*	.54	.65	.43	.39	.19	.40
Biodata - Artistic	2765	2.91	.82	.14	.26	.27	.44	*.86*	.72	.40	.46	.02	.05
Biodata - Multicultural	2765	2.98	.66	.06	.12	.27	.51	.60	*.80*	.46	.52	.05	.20
Biodata - Leadership	2768	3.07	.81	.16	.14	.28	.35	.34	.38	*.86*	.64	.26	.33
Biodata - Responsibility	2768	3.32	.76	.21	.18	.25	.31	.38	.41	.53	*.79*	.14	.31
Biodata - Health	2768	3.25	.51	.14	.13	.28	.14	.01	.04	.20	.11	*.68*	.13
Biodata - Career	2768	3.32	.65	-.08	-.22	.23	.31	.04	.16	.27	.24	.09	*.77*
Biodata - Adaptability	2768	3.38	.45	.10	.08	.33	.22	.08	.15	.30	.17	.46	.20
Biodata - Perseverance	2768	3.73	.49	.12	-.03	.50	.38	.17	.26	.42	.31	.31	.41
Biodata - Ethics	2768	3.86	.54	.20	.13	.43	.25	.20	.16	.18	.26	.19	.17
SJT	2730	.66	.33	.26	.20	.33	.21	.21	.22	.21	.31	.12	.13
Cumulative GPA	1256	3.45	.56	.53	.53	.26	.13	.18	.11	.09	.13	.11	-.14
BARS Total	612	3.78	.51	.08	.01	.21	.24	.20	.29	.29	.26	.26	.22
Controllable Absent.	556	1.74	1.08	.00	.17	-.17	-.08	-.08	-.12	-.01	-.10	-.13	-.08
OCB	593	3.28	.64	-.01	-.10	.06	.13	.00	.18	.35	.29	.09	.16
Graduation Status	1934	.55	.50	.30	.25	.11	.03	.13	.10	.13	.13	.08	-.03

	Adaptability	Perseverance	Ethics	SJT	Cumulative GPA	BARS	Controlled Absences	OCB	Graduation
High School GPA	.13	.14	.24	.29	.53	.09	-.01	-.01	.30
Ability Score (ACT/SAT)	.09	-.04	.16	.23	.53	.01	.14	-.10	.25
Biodata – Knowledge	.50	.70	.64	.46	.31	.29	-.18	.08	.14
Biodata - Learning	.31	.49	.34	.27	.14	.31	-.07	.20	.03
Biodata – Artistic	.10	.21	.27	.26	.19	.24	-.09	.00	.14
Biodata – Multicultural	.20	.33	.22	.28	.13	.37	-.11	.22	.11
Biodata – Leadership	.40	.52	.23	.25	.10	.35	-.03	.43	.14
Biodata – Responsibility	.24	.41	.36	.40	.15	.34	-.09	.36	.15
Biodata – Health	.69	.43	.27	.16	.13	.36	-.15	.12	.09
Biodata – Career	.28	.54	.23	.17	-.16	.28	-.09	.21	-.04
Biodata – Adaptability	*.65*	.68	.28	.23	.06	.40	-.12	,28	.07
Biodata --Perseverance	.47	*.75*	.44	.36	.08	.47	-.18	.29	.09
Biodata – Ethics	.18	.31	*.67*	.61	.20	.32	-.29	.15	.14
SJT	.16	.27	.44	*.76*	.25	.29	-.20	.15	.14
Cumulative GPA	.05	.07	.17	.22	*1.00*	.25	-.05	.03	.27
BARS Total Score	.28	.36	.23	.23	.22	*.77*	-.30	.50	.16
Controllable Absenteeism	-.10	-.16	-.24	-.17	-.05	-.27	*1.00*	-,11	-.14
OCB	.20	.22	.11	.12	.03	*.26*	-.10	*.80*	.06
Graduation Status	.06	.08	.11	.12	.27	.14	-.14	*.06*	*1.00*

[a]SJT = situational judgment test, BARS = behaviorally anchored rating scale, OCB = organizational citizenship behavior, Learning = Continuous Learning, Multicultural = Multicultural Appreciation, Career = Career Orientation. Correlations above .08 for variables with 600 cases are significant ($p < .05$) at r = .05. Reliabilities of the measures are in italics on the diagonal of the matrix. Intercorrelations above the diagonal have been corrected for unreliability in both measures in those instances in which a reliability measure was available.

Table 5

Validity of Biodata and SJT Measures: Hierarchical Regression of Cumulative College GPA, Behaviorally Anchored Rating (BARS), Absenteeism, and Organizational Citizenship Behavior (OCB) on the Predictors

	Cumulative GPA		BARS		OCB		Absenteeism	
Step 1	r	b	r	b	r	b	r	b
HSGPA	.531	.351*	.079	.070	-.007	.008	-.011	-.050
ACT/SAT	.539	.327*	.011	.002	-.098	-.107*	.149	.216*
ΔR^2		.398*		.008		.012*		.033*
Step 2								
Knowledge	.263	.046	.192	-.096	.056	-.045	-.161	-.063
Learning	.129	.064*	.226	-.011	.116	-.003	-.072	.041
Artistic	.174	.020	.200	.029	-.001	-.195*	-.082	-.037
Multicultural	.115	.022	.302	.199*	.184	.159*	-.115	-.084
Leadership	.092	-.014	.300	.043	.372	.286*	-.024	.112*
Responsibility	.136	.003	.274	.038	.298	.160*	-.107	-.004
Health	.101	.061*	.236	.157*	.093	-.036	-.121	-.131*
Career	-.144	-.132*	.220	.081	.174	.014	-.098	-.002
Adaptability	.054	-.043	.284	.087	.213	.110*	-.097	.004
Perseverance	.063	-.004	.363	.148*	.231	-.003	-.181	-.040

Ethics	.164	.004	.239	.098*	.102	.031	-.276	-.203*
SJT	.211	.070*	.223	.088	.087	-.021	-.194	-.102*
ΔR^2		.029*		.240*		.201*		.116*
Overall R^2		.426*		.247*		.214*		.149*
Adj. R^2		.419		.227		.193		.127
N		1155		547		558		556

---

*Indicates a significant beta, $p \leq .05$. Correlations above .06 are all statistically significant for the GPA relationships. Correlations above .09 are statistically significant for relationships with BARS, Absenteeism, and OCB. b refers to standardized regression weights.

Table 6

Logistic Regression of Graduation Status on Biodata and SJT

	Graduation Status	
	B	Odds Ratio
HSGPA	1.327	3.771*
ACT/SAT	.266*	1.304*
X^2 (df=2)		178.25*
Nagelkerke R^2		.128
Knowledge	-.214	.807
Learning	-.282*	.754*
Artistic	.133	1.143
Multicultural	.159	1.173
Leadership	.107	1.113
Responsibility	.067	1.069
Health	.102	1.212
Career	-.092	.912
Adaptability	.011	1.011
Perseverance	.136	1.146
Ethics	.158	1.171
SJT	.304	1.356
X^2 (df=12)		38.891*
Nagelkerke R^2		.154

Table 7

A. Demographic Composition: Percent of Various Groups Admitted under Various Levels of Selectivity using Cognitive Measures and Both Cognitive and Noncognitive Measures[a]

	Hispanic		Asian		African		Caucasian	
	Cog	Cog+	Cog	Cog+	Cog	Cog+	Cog	Cog+
Hi Selec.(Top 15%)	4.3	6.4	17.8	14.9	.9	4.1	77.0	74.6
Mod.Selec.(Top Half)	4.0	4.6	1.5	1.1	8.3	1.0	77.1	75.3
Min. Selec.(Top 85%)	4.5	4.7	7.6	7.7	18.4	18.7	69.5	69.0
All	3.7		9.0		19.4		67.8	

B. GPA: Mean and Standard Deviation of Members of Various Groups Admitted under Various Levels of Selectivity using Cognitive Measures and Both Cognitive and Noncognitive Measures

	Hispanic				Asian				African				Caucasian			
	N	Cog	N	Cog+	N	Cog	N	Cog+	N	Cog	N	Cog+	N	Cog	N	Cog+
Hi S	4	3.67(.37)	6	3.57(.40)	23	3.88(.36)	18	3.84(.33)	0		1	3.66(--)	208	3.88(.36)	197	3.87(.37)
Mod.S	17	3.51(.56)	17	3.42(.58)	55	3.65(.44)	54	3.63(.45)	15	3.46(.46)	15	3.38(.46)	651	3.66(.43)	638	3.66(.43)
Min. S	23	3.44(.56)	22	3.47(.55)	64	3.62(.44)	64	3.62(.44)	49	3.05(.51)	47	3.08(.48)	929	3.52(.49)	905	3.54(.48)

C. Persistence: Proportion of Various Groups Admitted under Various Levels of Selectivity using Cognitive Measures and Both Cognitive and Noncognitive Measures Who Graduated in Four Years

	Hispanic		Asian		African		Caucasian	
	Cog	Cog+	Cog	Cog+	Cog	Cog+	Cog	Cog+
Hi Selec.(Top 15%)	93	90	89	91	100	92	72	70
Mod.Selec.(Top Half)	79	78	72	75	87	87	59	60
Min. Selec.(Top 85%)	77	76	68	68	79	78	51	52

---

[a]Cog represents the use of an admissions procedure in which a sum of the standardized SAT/ACT composite and HSGPA are used to make admissions decisions. Cog+ represents an admissions procedure in which a sum of the standardized SAT/ACT composite, HSGPA, and the noncognitive composite are used to make admissions decisions. Hi S refers to a strategy in which the top 15% of the applicants are admitted; Mod S is a strategy in which the top half is selected; and Min S is a strategy in which the top 85% are admitted.

Table 8

Moderated Regression of Cumulative College GPA on Noncognitive Composite, HSGPA and SAT/ACT with Race as a Moderator

Variable[a]	African b	Hispanic b	Asian b
Constant	3.437	3.340	3.227
Noncog. Composite	.053	-.092	-.013
SAT/ACT	-.011	.497*	.260*
HSGPA	.372*	-.309	.354*
R^2_{Δ}	.427*	.338*	.339*
Race	-.111	.007	.099
R^2_{Δ}	.008*	.001	.000
Race X Noncog. Composite	-.002	-.072	.064
Race X SAT/ACT	.215*	.147	-.056
Race X HSGPA	-.122*	-.279	-.104
R^2_{Δ}	.006*	.004	.002
Overall R^2	.441	.342	.341
N	1007	959	996

[a]b refers to the nonstandardized regression weights. Noncog = noncognitive. The regression weights are those from the full regression equation produced at the step of the hierarchical regressions.

*$p < .05$.

Table 9

Standardized subgroup mean differences for Predictor and Outcome Measures

Variable	Hispanic	Asian	African American
Knowledge	-.04 (-.17) [a]	.00 (.11)	-.29 (-.30)
Cont. Learning	-.02 (-.05)	.10 (.16)	-.10 (-.11)
Art. Appreciation	.09 (.04)	.18 (.17)	-.23 (-.26)
Multic. Appreciation	.49 (.41)	.69 (.63)	.09 (.06)
Leadership	-.02 (-.10)	-.06 (-.06)	-.11 (-.14)
Responsibility	.07 (.01)	.16 (.09)	-.12 (-.15)
Health	-.16 (-.31)	-.41 (-.39)	-.43 (-.43)
Career Orientation	.05 (.06)	-.11 (-.02)	.52 (.53)
Adaptability	-.05 (-.14)	-.23 (-.23)	-.08 (-.10)
Perseverance	.08 (.00)	-.08 (-.09)	.20 (.19)
Ethics	-.26 (-.28)	-.06 (-.07)	-.20 (-.21)
Situational Judgment	-.13 (-.15)	.01 (.01)	-.20 (-.22)
SAT/ACT	-.80 (-.83)	.22 (.35)	-1.18 (-1.18)
HSGPA[b]	-.63 (-.62)	.07 (.07)	-.88 (-.88)
BARS	.08	.00	-.58
OCB	.04	-.14	-.30
Attendance	-.29	.15	.04
College GPA	-.05	.37	-1.06

Table 9 continued

[a] Numbers in parentheses are the standardized subgroup mean differences from the Caucasian group for the total sample when first measured as freshman college students. Ns for these groups were 161, 187, 678, and 1506 for Hispanic, Asian, African American, and Caucasian students respectively. The numbers outside parentheses are the mean differences for the sample for whom outcome data were available. Ns for these subgroups were 22, 63, 36, and 364 for Hispanic, Asian, African American groups respectively.

[b] HSGPA-high school grade point average, BARS= self rated performance on the behaviorally anchored rating scales, OCB=organizational citizenship behavior.