RESEARCH REPORT 2011-1

CollegeBoard
inspiring minds™

Development and Validation of Measures of Noncognitive College Student Potential

By Neal Schmitt, Abigail Billington, Jessica Keeney, Matthew Reeder, Timothy J. Pleskac, Ruchi Sinha and Mark Zorzie

Neal Schmitt and **Tim Pleskac** are professors in the Department of Psychology at Michigan State University.

Abigail Billington, **Jessica Keeney**, **Matthew Reeder**, **Ruchi Sinha**, and **Mark Zorzie** are graduate students in the Organizational Psychology Program at Michigan State University.

The College Board

The College Board is a mission-driven not-for-profit organization that connects students to college success and opportunity. Founded in 1900, the College Board was created to expand access to higher education. Today, the membership association is made up of more than 5,900 of the world's leading educational institutions and is dedicated to promoting excellence and equity in education. Each year, the College Board helps more than seven million students prepare for a successful transition to college through programs and services in college readiness and college success — including the SAT® and the Advanced Placement Program®. The organization also serves the education community through research and advocacy on behalf of students, educators and schools.

For further information, visit www.collegeboard.org.

Contents

Abstract

In considering and evaluating approaches to the admission of college students, the usual approach is to try to measure past academic achievement and primarily verbal and math ability on the assumption that these abilities will predict subsequent college academic grades and achievement. Measures such as the SAT®/ACT and high school GPA do predict classroom achievement (Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008). However, it is also the case that most universities claim to develop students in areas not well represented by classroom grades such as leadership, social responsibility, integrity, multicultural appreciation, and others. In an effort to develop noncognitive measures that would predict subsequent student performance, we began with a "job analysis" of the "job" of undergraduate students. We developed a list of expectations universities claim to have of students and derived a list of constructs that were hypothesized to be essential to success. This set of constructs has been central to the development of a set of biodata measures and a situational judgment inventory. We present evidence for the validity of these measures across two different studies. Also presented are standardized differences across various demographic subgroups. In general, our results show that these tools, along with traditional classroom achievement measures, provide a better representation of the totality of relevant college student outcomes.

A central part of the college admission process includes the use of standardized test scores and some record of high school achievement. These measures do predict college success quite well. (Hezlett et al., 2001; Kobrin et al., 2008; Kuncel, Hezlett, & Ones, 2004; Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). Standardized cognitive ability tests or achievement tests (e.g., SAT/ACT) can be administered to large numbers of students simultaneously and are objectively and efficiently scored, which makes them practical especially to institutions that must process large numbers of applicants. Throughout this report, we often refer to SAT/ACT tests as cognitive measures; this does not reflect our opinion that they are measures of g as that construct is understood in the literature on intelligence testing. In fact, these tests are more likely considered measures of the accumulated knowledge gained during the test-takers' previous educational experiences. On the downside, standardized tests display large differences in subgroup performance, which typically disadvantage minority students when these tests are used to make admission decisions (Wilds & Wilson, 1998; Sackett, Schmitt, Ellingson, & Kabin, 2001). In addition, many argue that other noncognitive variables are needed to predict adequately which students succeed or fail (Camara & Kimmel, 2005; Sedlacek, 2004; Sternberg & The Rainbow Project Collaborators, 2006) and to provide a more holistic view of student potential.

To evaluate noncognitive student attributes, many institutions evaluate student portfolios that include letters of recommendation, applicant essays, evidence of involvement in extracurricular and community activities, interviews, and other procedures. Schools vary widely in their assessment of the information in these supplemental materials. Some use them to make ratings of the students' potential (Schmitt, Billington, Golubovich, et al., 2009); others simply rate essays or extracurricular activities and then combine these ratings with standardized test scores to arrive at admission decisions. In making the ratings, it is unclear to what material different raters refer. For example, one rater might form ratings based on personal qualities identified in a letter of recommendation, a second may be most impressed by information about involvement in volunteer community activities, and a third may actually weight knowledge of standardized test scores heavily in making an overall rating. This makes it difficult to discern what constructs might be evaluated across people or to what degree different raters will agree on student potential. Hence we do not know the nature of what is being evaluated or whether such evaluation is systematic across student applicants.

In spite of these potential problems or inconsistencies, there is growing interest in the inclusion of noncognitive factors in the admission process. Past studies have included the examination of meta-cognitive skills (e.g., Zeegers, 2001), study attitudes (e.g., Zimmerman, Parks, Gray, & Michael, 1977), creativity and problem solving (Sternberg & The Rainbow Collaborators, 2006), study motivation (Melancon, 2002), and personality traits (Ridgell & Lounsbury, 2004). A recent meta-analysis (Crede & Kuncel, 2008) documented that factors like study habits, study motivation and study skills among other attitudinal constructs have incremental validity above that provided by standardized test scores and high school grades.

Rationale for and Literature Regarding Dimensions of Student Success

Academic institutions have long claimed that the college experience is multifaceted, developing not only knowledge or "book smarts" but also leadership, personal values and character (Taber & Hackman, 1976; Willingham, 1985). Despite these claims, and as stated above, today's institutions continue to evaluate the outcomes of the college experience on relatively narrow criteria such as grade point average (GPA) and graduation. In addition, efforts to validate college admission procedures often rely solely on GPA and very often only freshman GPA. Aside from providing a better sense of what is expected of student outcomes, the move to a broader conceptualization of performance could result in a student body that is more diverse. A more complete consideration of these alternate dimensions of student performance should lead to the consideration of noncognitive measures of student potential. As opposed to cognitive measures, noncognitive measures usually display minimal subgroup differences; their use to make admission decisions should increase the probability of admission of members of subgroups whose scores on the SAT/ACT tend to be lower. The implementation and use of assessment methods and performance measures that evaluate the real breadth of student development is long overdue.

The implementation and use of assessment methods and performance measures that evaluate the real breadth of student development is long overdue.

If we accept that accurate reflections of the broad development we expect to see in college students are contained in the marketing materials of U.S. colleges and in the statements by various stakeholders in the academic enterprise (e.g., college presidents, state legislators, parents, the students themselves), then college student success must be a broader domain than is reflected by traditional criteria of college success. Our view is that the measurement of college success should be multidimensional. Certainly, given the breadth and complexity of the criterion domain of college student success as articulated in most universities' mission statements, we should develop and implement evaluation procedures that are commensurate in scope.

Conceptualizing and evaluating successful development as a college student depends on the multiple outcomes desired by students and the school administration (Willingham, 1985). In an early attempt to discern dimensions of college performance, Taber and Hackman (1976)

found 17 dimensions, academic and nonacademic, to be important in classifying successful and unsuccessful college students. Examples of these dimensions are intellectual perspective and curiosity, communication proficiency, and ethical behavior. Furthermore, other researchers report that college students actively engaged across numerous domains tend to achieve greater success in their overall college experience as reflected in their scholastic involvement, accumulated achievement record or their graduation (Astin, 1984; Willingham, 1985). High school honors, school references, personal statements and a measure of persistence or follow-through have been shown, above and beyond high school rank and SAT score, to predict scholarship, leadership, nonacademic accomplishments and overall success ratings (Willingham, 1985). In related work, Ra (1989) found measures of high school leadership, athletic success, persistence in extracurricular activities, honors, personal statements and references predicted college GPA, but not perceived success and leadership experience. More recently, Cress, Astin, Zimmerman-Oster, and Burkhardt (2001) found that those who participated in leadership education and training showed development in personal and societal values, civic responsibility, multicultural awareness, leadership skills, and understanding of leadership theories. These empirical findings reinforce the usefulness of developing measures of multidimensional performance measures for the college context.

Determination of Targeted Dimensions

Determining the nature and number of dimensions of college success in the research described in this report was necessarily an exploratory information gathering process that followed two primary guidelines. First, the number of dimensions should not be so many that the information is unwieldy, yet not be so few that the domain of college success is not appropriately represented (as we have pointed out, the latter point was the prime motivation for our research). Second, we wanted to understand how a variety of stakeholders who are invested in the process of college education defined college success because relying on one source alone could lead to inadequate definitions and representations of college success. Also, it would not provide information about how different stakeholders converge and diverge in the dimensions of college success they identify as important.

In identifying the dimensions, we first examined the Web pages of colleges and universities, including both public and private institutions. We examined their mission statements and the stated educational objectives of their programs. We engaged in this effort for two reasons. First, we used use this information to achieve a clearer sense of those student characteristics and constructs that may be central to student success. Once dimensions were established from this information, we began the process of developing the situational judgment inventory (SJI) and biodata instruments. Second, we used this information as the basis of the development of outcome measures against which to validate our predictor instruments. Outcome measures developed as a result of this process included self-report activity checklists and self- and peer-ratings on behaviorally anchored rating scales.

Specifically, 35 colleges and universities' websites were searched, either by manually seeking stated educational objectives or mission statements from the home page or by entering search terms such as "university mission statement" in the universities' search engines. The 35 colleges and universities varied on characteristics such as public/private, student body size and the geographic region in which they were located. Twenty-three institutions provided usable information. Institutions not providing usable information did not explicitly state their educational objectives, nor did they provide a university mission statement. There was no discernable relationship between those institutions that provided usable information and those that did not. After identifying relevant information from each institution, this information

was distilled into the discrete constructs represented in each statement in order to separately identify all criteria contained within the statement, retaining the original wording to the extent possible, and without introducing extraneous wording. For example, the sentence fragment "...promote a commitment to learning, freedom, and truth." was decomposed into "promote a commitment to learning," "promote a commitment to freedom," and "promote a commitment to truth." This decomposition resulted in 174 separate fragments (including redundancies from multiple institutions). Three researchers independently sorted the resulting fragments into dimensions rationally derived from the criteria, then agreed on 12 dimensions through joint discussion of their independent sorts.

Next, we interviewed a lead administrator in the department of Residence Life at a large midwestern university, and we analyzed University Residence Life materials. Dimensions identified at this level included academics, social development (e.g., roommate and romantic relationships), personal development (e.g., substance abuse, leisure time, time management), ethical and character development (e.g., challenge to values, maintaining community standards), and multicultural competence (e.g., seeking out other cultures and lifestyles). The information gathered from this interview was consistent with hypotheses about the dimensionality of student performance already gathered from our examination and coding of university mission statements.

Finally, the criteria identified through our Web search and from university resources were compared against those identified in other related research (Beatty, Greenwood, & Linn, 1999; Patelis & Camara, 1999; Sackett, et al., 2001; University of Pennsylvania, 2000; and Wightman & Jaeger, 1998). This literature review did not lead us to new dimensions. Table 1 contains the list of 12 dimensions derived from these various sources of information.

Using the 12 dimensions listed in Table 1, the same three raters independently re-sorted the 134 (of the original 174) fragments that were judged to be nonredundant into one of the 12 dimensions. Of those 134 fragments, 85 (62 percent) were agreed upon by all three raters as to the dimension they best reflected, and 129 statements (96 percent) were agreed upon by at least two raters. The five on which there was no agreement were deleted from further consideration in our work. After this resorting task, each of the identified dimensions was compared to similar dimensions in the industrial and organizational (I/O), educational and vocational psychology literature involving a college population. In some instances, the labels and definition of each dimension were modified to be more consistent with usage in the current literature in I/O, educational and/or vocational psychology. As indicated above, we used these 12 dimensions as a guide in writing items for our situational judgment inventory and biodata instruments and also in developing student performance criteria against which to validate these measures. The dimensions also served as the basis for subsequent confirmatory analyses of the dimensionality of both the predictors and the outcome measures. Our final list as indicated in Table 1 includes knowledge (the usual academic objective assessed in studies of college student success), continuous learning, artistic/cultural appreciation, multicultural appreciation, leadership, interpersonal skills, social responsibility, physical and psychological health, career orientation, adaptability, perseverance, and ethics.

Development of Potential Admission Tools

These 12 dimensions or constructs served as the blueprint for the development of our measures. The first measure we developed was an assessment of biographical data, or biodata. Biodata contain information about one's background and life history. Despite the informal use of similar information in college applications (e.g., extracurricular activity lists

and résumés), we undertook the development of a biodata inventory with standard multiple-choice responses to questions about one's previous experiences, similar to tests used in job selection processes. The reasons for developing this inventory included the desire for efficient data collection and scoring and improved standardization of the data collected and used to document students' potential in noncognitive domains.

Multiple sources were searched for preexisting biodata items that would fit as many of the 12 criterion dimensions as possible. Our initial search produced 197 that fit one of our dimensions and had relevance to the college context. The majority of the items were adapted from Pulakos, Schmitt, and Keenan (1994) and Mumford (2001). However, other sources that were searched include the following: the University of Georgia Biographical Questionnaire (Owens Albright, & Glennon, 1966), the Assessment of Background and Life Experiences (Hough, Eaton, Dunnette, Kamp, & McCloy, 1990), the Personnel Reaction Blank (Gough & Arvey, 1998), a biographical questionnaire by Russell, Green and Griggs (undated manuscript) and Schmitt and Kunce (2002). When necessary, item stems were modified for appropriateness to the college context. After this procedure, some dimensions lacked a sufficient number of items. We then rationally generated additional items for those dimensions.

The reasons for developing this inventory included the desire for efficient data collection and scoring and improved standardization of the data collected and used to document students' potential in noncognitive domains.

The dimensional structure of the inventory was further established by assessing inter-rater agreement on a rational sort of the items. Six researchers resorted all items back into the 12 dimensions. Items on which five of six raters agreed with the originally assigned dimension were retained; those on which four of six agreed were discussed and rewritten or dropped, and those with less agreement were discarded. When all six raters assigned an item to one dimension other than the one originally specified, it was reassigned to that new dimension. Using these criteria, five additional items were discarded, several were reassigned to a new dimension, and 11 items that were not assigned to any of our 12 dimensions but appeared to be good predictors were retained in a miscellaneous category for exploratory purposes. The final biodata inventory, at that point, consisted of approximately 200 items, each scored on a 4- or 5-point scale. Subsequently, additional items were developed so that we now have a pool of 296 items.

We also developed a situational judgment measure that included items representing situations related to the 12 dimensions, along with alternative courses of action from which a student applicant was asked to identify the option that they would most and least likely pursue. Situational judgment inventories (SJIs) are measures in which respondents choose or rate possible actions in response to hypothetical situations or problems, usually in a paper-and-pencil format. Though SJIs have been in and out of favor in employment contexts for more than 80 years, there has been a renewed interest in SJIs due to their validity as

employment tests designed to predict job performance. A recent meta-analysis (McDaniel, Hartman, Whetzel, & Grubb, 2007; McDaniel, Morgeson, Bruhn-Finnegan, Campion, & Braverman, 2001) found SJIs to have an overall criterion-related validity of ρ = .34, though there appears to be substantial variability associated with that value ($\sigma\rho$ = .14 with a 90 percent credibility interval of .09 to .69), with job complexity as a potential moderator (Huffcutt, Weekley, Wiesner, DeGroot, & Jones, 2001). SJIs are also less costly to construct and administer than more complex simulations like work samples and assessment centers (Motowidlo, Dunnette, & Carter, 1990). In personnel selection, the use of SJIs usually reduces adverse impact for minorities relative to that of cognitive tests (Sackett et al., 2001; Pulakos & Schmitt, 1996). The SJI produces favorable test-taker reactions (Hedlund et al., 2001) as well as high perceptions of face validity (Clevenger, Pereira, Wiechmann, Schmitt, & Harvey, 2001; McDaniel et al., 2001). Such support for SJIs in employment settings then suggests that they may be a viable supplement or alternative to traditional cognitive ability testing in our particular domain of interest, college admissions. One research endeavor in college admissions (Hedlund et al., 2001) found a SJI to have rather small incremental validity above GMAT scores for MBA students (ΔR^2 = .03). Our SJI development is based on a different set of considerations and methods, including a broader definition of student success.

A search of existing situational judgment inventory measures used in previous studies with managers, FBI special agents and school principals produced item stems that fit many of the aforementioned 12 dimensions. We next recruited and paid undergraduate students at a large midwestern university to participate in developing our SJI further. This development process consisted of three separate data-collection phases. First, students generated critical incidents for use as additional item stems for dimensions underrepresented by existing SJI items. This critical incident generation task was administered in a forward- and reverse-ordered format to ensure that time limitations didn't preclude at least some individuals paying attention to each performance dimension in the set. Second, an independent set of students created multiple response options for each item stem.

Third, rating data were collected from an advanced undergraduate psychology course. Students were all junior and senior college undergraduates, who completed the instrument as part of a project for a psychological measurement class. Graduate students not affiliated with the research project administered the instrument during laboratory sessions. The two-hour task consisted of undergraduates (a) reading a series of SJI items reflecting situations that college students tend to face, with item content representing situations reflecting the aforementioned 12 dimensions of college success; and (b) for each item/situation, rating the effectiveness of anywhere from four to seven different responses to the situation, also indicating the very best and very worst response. One half of the 112-item instrument was administered to each of six lab sessions, with three sessions getting the first half and three sessions getting the second half. The two test forms, Forms A1 and A2, comprised 576 and 561 items, respectively, which included eight demographic questions. These ratings of the effectiveness of the various options of dealing with the situations were used to develop a scoring key for the situational judgment items. Use of college juniors and seniors as our experts in developing scoring keys assumes that they are successful students. Subsequent use of resident hall advisers and a group of senior African American students as experts resulted in very similar scoring keys.

The scoring key was empirically developed to reflect the fact that individuals taking the SJI should get a higher score on items where they tend to agree with student experts on the best and worst responses to a situation. Scores for each item range from -2 to +2 and follow the SJI scoring procedure developed and reported in Motowidlo, Dunnette and Carter (1990). Subsequent item development work has increased the SJI item pool to 150 items.

Reliability and Dimensionality of Measures

To assess the reliability and dimensionality of both biodata and the SJI, we computed coefficient alpha for each scale and did exploratory and confirmatory factor analyses to confirm the hypothesized dimensionality of both the biodata and SJI items using responses of 647 college freshmen who participated in the initial development work. Reliabilities and interscale correlations (both observed and those corrected for attenuation due to unreliability in the scales) are presented in Table 2 for 11 biodata scales and a composite SJT scale. Efforts to develop an internally consistent biodata measure of interpersonal skills were not successful, so a measure of that dimension is not included. These data are from our last study and reflect considerable item selection and refinement over the last several years based on three separate studies. Four scales (Knowledge, Adaptability, Health and Ethics) have reliabilities less than .70, which is sometimes considered marginal though none are lower than .65. All of the corrected intercorrelations are substantially less than 1.00, though moderate in most cases (i.e., .30 to .70), which is evidence for their discriminant validity.

A priori subscales for the SJI were not internally consistent, and correlations between subscales indicated a lack of discriminant validity. Exploratory factor analysis of the SJI items (Gillespie, Oswald, Schmitt, Manheim, & Kim, 2002) revealed the presence of a relatively large general factor accounting for three times the variance of the second factor. Although there were a number of additional factors with eigenvalues above 1.00, these factors accounted for small portions of variance and also were difficult to interpret substantively. In addition, the reliabilities of the intended scales were all relatively low; hence we have combined all SJI items in a single 36-item scale in our subsequent research. We have also demonstrated our ability to construct multiple additional forms of the SJI from our item pool that have highly similar psychometric characteristics (Oswald, Friede, Schmitt, Kim, & Ramsay, 2005). In the data presented in Table 2, the SJI has a reliability of .76 and relatively low correlations with all biodata scales.

Applicant Versus Research Participant Responses and Score Inflation

One of the primary concerns associated with possible use of noncognitive measures is the possibility that, in high-stakes situations, respondents would inflate or fake their scores to reflect how they think a desirable applicant should respond. To get some sense of the amount of inflation that might occur, we have compared responses of participants who were responding as part of a research project with

One of the primary concerns associated with possible use of noncognitive measures is the possibility that, in high-stakes situations, respondents would inflate or fake their scores to reflect how they think a desirable applicant should respond.

responses of applicants. The latter group was told that the measures were not being used to make admission decisions, but the data were collected as part of the application process prior to an admission decision. Means and standard deviations of these two groups of participants are presented in Table 3. As can be seen in this table, applicants' scores are higher than those of research participants (incumbent students). The standardized mean differences range from a low of .06 for Health to a high of .74 for Knowledge. Most are different by .3 to .5 in standard deviation units. In spite of the obvious inflation, there are very small or nonexistent differences in the standard deviations of the responses and, as we will report below, the validity of the scales against various outcomes is nearly the same for these different groups of respondents.

We did conduct several studies in which we evaluated various methods of assessing and reducing response inflation (Ramsay, Kim, Gillespie, & Friede, 2003; Ramsay, Schmitt, Oswald, Kim, & Gillespie, 2006) suggested by the previous literature. Score inflation is particularly worrisome in high-stakes testing programs since various firms such as Kaplan will almost certainly develop and provide coaching courses for examinees. There have been few studies assessing the effect of coaching on biodata and SJI, but Miller (2001) found that a 15-minute coaching program with sample items and a description of the construct being measured resulted in increases in scores on a Conscientiousness scale equal to nearly a half standard deviation. Cunningham, Wong and Barbee (1994) found that a brief written description of an integrity test raised scores by about 10 percent. Similarly, Ramsay et al. (2006) conducted a 10-minute coaching session in which the constructs measured by the biodata and SJI in this study were described. This brief coaching resulted in increases in scores equal to about .5 standard deviation units.

Perhaps the most effective deterrent to score inflation is the use of a warning that responses will be checked as to their accuracy. Various authors (Becker & Colquitt, 1992; Dwight & Donovan, 2003; Vasilopoulos, 1999) have found that such a warning diminished score inflation on biodata, personality and situational judgment measures. These warnings were most effective when they included the notion that faking would be identified, and the consequences of such faking were described. In our study, we included such warnings, and it did have some impact on the resultant scores. However, the effect was much smaller than the impact of coaching.

We also examined the nature of the biodata items to see if some were more or less susceptible to faking. Experts rated the content of the items using the characteristics discussed by Mael (1991). These judgments of item objectivity and verifiability and the degree to which a respondent would have control over the behavior being reported in an item were correlated -.22, -.18 and -.15, with mean biodata responses while item relevance to an academic situation was correlated .10 with mean biodata responses. This is indirect evidence, at least, that items should be written so that they are objective and appear to be verifiable, if checked, and that they represent behavior that is under the control of the respondent. The correlation with relevance suggests that the more "face valid" items are, the more likely applicants will enhance their responses. A similar effect has been noted in the employment arena (Dwight & Alliger, 1997).

An additional effort was made to evaluate the use of requirements that examinees elaborate their responses (Schmitt & Kunce, 2002) to biodata items (e.g., asking a respondent to indicate what leadership positions they held in high school if they say they occupied four such positions). This requirement did lower scores on items for which elaboration was required (Ramsay et al., 2006); however, these lowered scores occurred primarily on items for which elaboration was required (Schmitt et al., 2003). We concluded that requiring elaboration of all

items would be counterproductive as the instrument would take respondents much longer to complete and admission personnel much longer to score and verify.

Efforts to assess the degree of "faking" and correct for it took two forms. First, we used bogus items to detect those who were not reading the instrument carefully or who were responding randomly. A bogus item is one for which a particular response is clearly untrue or impossible (e.g., responding affirmatively to the statement that "all grass is larger than trees" clearly indicates a lack of attention to the item content). We used four or five items like these in our instrumentation to detect and remove careless responders. Usually, only a small minority (less than two percent) were identified as careless responders in any of our studies.

Attempts to identify socially desirable responding have proven more difficult. Socially desirable responding may occur when the characteristics being captured by a test or test item are transparent to the respondent, and those characteristics are regarded as attractive by the respondent or more generally attractive in the respondent's culture. Personality tests are criticized for being susceptible to socially desirable responding; for instance, conscientiousness and emotional stability factors of the Big Five are regarded as socially desirable and adaptive. This transparency potentially makes personality measures, such as those capturing the Big Five, susceptible to inflation. On average, the mean scores on Big Five measures have been shown to be inflated by about half a standard deviation under instructions to fake good (Viswesvaran & Ones, 1999). This degree of inflation is similar to the difference between applicant and incumbent students observed in Table 3. To be able to control for this tendency, social desirability scales have been created to gather information about the tendency of individuals to respond in a socially desirable manner. However, these scales themselves are fakable. For example, Viswesvaran and Ones (1999) showed in their meta-analysis a faking effect size of more than one standard deviation on social desirability scales.

Paulhus (1984) presents evidence that there are two components of socially desirable responding: self-deception and impression management. *Self-deception* is the unconscious inclination that an individual has toward claiming that desirable characteristics apply to him or her. *Impression management* is the conscious dissembling that an individual engages in to present a favorable impression. Paulhus' Balanced Inventory of Desirable Responding (BIDR) is an example of a social desirability measure that captures both dimensions. It is this impression management component that is regarded as being most closely linked to inflation. We used Paulhus' measures of these two components in several of our data collection efforts. The results indicate that scores on the biodata are correlated with both self-deception (.37 and .46) and impression management (.25 and .39) in two different studies (Ramsay et al., 2003; Schmitt et al., 2003). There is evidence, then, that scores on the biodata instrument, at least, are related to social desirability. However, this inflation does not seem to have an impact on the validity of the scales or the variance in applicant responses (see Table 3 and evidence presented below on validity).

In terms of the practical advice that might be derived from these studies and the research literature on faking, we suggest the following. First, biodata items that are objective and verifiable should be preferred over those that represent opinions, attitudes or preferences. Second, examinees should be warned that it is possible to verify their responses and that evidence of the inflation of their response will have consequences. Third, bogus items should be used to detect those who are careless responders. Finally, there remains the likelihood that responses will be inflated as a function of coaching and the motivation accompanying high-stakes testing. While our research indicates this inflation does not impact the variability or validity of responses, researchers and users of biodata and SJI information should be

aware of the problem and continue to monitor its impact on the response quality of student applicants in high-stakes situations.

While our research indicates this inflation does not impact the variability or validity of responses, researchers and users of biodata and SJI information should be aware of the problem and continue to monitor its impact on the response quality of student applicants in high-stakes situations.

Validity

As was mentioned, the traditionally used measure of success has been college GPA. Universities, however, value a number of other outcomes for students. A number of noncognitive measures of concern to administrators were administered to students to assess these alternate conceptualizations of success. Behaviorally anchored ratings, academic and social satisfaction, organizational citizenship behaviors, deviance, problematic drinking behaviors, drug use, controllable absenteeism, and intentions to drop out were assessed using scales constructed by the research team. The following section describes each of the outcome measures.

Behaviorally Anchored Rating Scales (BARS) were used to measure students' self-reported performance on the previously identified dimensions of college student success (Drzakowski et al., 2005; Oswald, Schmitt, Kim, Gillespie, & Ramsay, 2004). The BARS provides descriptions of each dimension of success and example behaviors that reflect different levels of performance on that dimension. Respondents rate their performance on a 5-point scale ranging from 1 (very low) to 5 (very high).

Organizational Citizenship Behaviors (OCBs) refer to nonrequired behaviors that promote the welfare of the university (Organ, 1988). The measure of this construct consisted of 10 five-point frequency-based scales with response options ranging from "Very Infrequently/Never" to "Very Frequently/Always." Example items included "Gone out of your way to make new students feel welcome at school," "Defended your school when others tried to criticize it," and "Participated in student government or other clubs that try to make your school a better place."

Deviance refers to a measure of behaviors that are detrimental to the university or to society in general. This measure consisted of 13 items, all with five-point frequency-based response options ranging from "Very Infrequently/Never" to "Very Frequently/Always." Example items included "Made a derogatory ethnic, religious or racial remark at school," "Let someone copy from your homework or cheat off of you in class," and "Illegally copied or downloaded computer software."

Participants were asked to self-report *absenteeism* on two items. They were asked to indicate "the extent to which you have missed regularly scheduled classes in the past 6 months." 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). There were five response options ranging from "Missed less than 5 times" to "Missed more than 30 times." All analyses reported in this paper are on the controlled absence measure only.

Students' intentions to drop out or transfer were assessed using three self-report items on a 5-point Likert-type scale ranging from 1 (strongly disagree) to 5 (strongly agree). The *intent to turnover* items were adapted from the intent to turnover scales described by Eaton and Bean (1995) and Griffeth and Hom (1988).

Student *academic satisfaction* was measured with five items with a five-option response scale ranging from "strongly disagree" to "strongly agree." Representative items included "I'm happy with what I learn in my classes" and "All in all, I'm satisfied with the education I get at this school."

Four items measured *social satisfaction* in college with the same response scale. Sample items include "I'm satisfied with the number of friends I have here" and "Overall, I'm satisfied with my social life at this school."

In addition to the outcomes above, *first-year college GPA* was collected from the participating universities. College GPA was corrected by university based on SAT scores to account for differences in the admitted sample at each school (Oswald et al., 2004).

Problem Drinking and Drug Use Scales were developed to assess the extent to which students participated in problematic behaviors related to the consumption of alcohol and other drugs. Response scales for items on each of these scales differ. The drug use scale consists of three items assessing cigarette, marijuana and other drug use.

The majority of these outcome measures were developed for use in the 2004 sample. All were administered at the end of these students' first, second, third, and seventh semesters except for the drinking and drug use scales. These were developed later in the course of this longitudinal data collection and were administered only in the data collection during the students' seventh semester. All of the aforementioned outcome measures were again administered in the sample of students who began college 2007, providing the research team with two samples of students who had completed the outcome measures.

The means, standard deviations and reliability estimates as well as their relationships with various predictors are presented in Table 4 for the 2004 and 2007 samples. Correlations among the biodata scales and the SJI are presented in Table 2 above based on responses of the 2004 sample. Correlations between biodata and the SJT represent a reasonable level of discriminant validity. In addition, the generally low correlations with high school GPA and SAT/ACT scores indicate that the noncognitive measures are potentially capable of providing incrementally predictive information.

Table 4 and Tables 5 and 6 contain two additional biodata scales (Awards and Jobs) that were developed based on questions in the Common Application Blank used by many universities in the admission process. These two scales reflect the various awards received while in high school as well as the jobs respondents held while in high school. The zero-order correlations between predictor scales and a number of outcomes are statistically significant ($p < .05$), though given the relatively large sample size, correlations above .08 were significant. To identify the combined predictability of these biodata scales and the SJT on these outcomes, regression analyses were conducted with both the 2004 and 2007 samples as well as meta-analytically combined data. High school GPA and SAT/ACT scores were entered in the first

step of these regressions, and the biodata and SJT scores were entered in the second step. In order to maximize power and obtain the most accurate results, the outcome variables measured most closely to the predictor data collections were used. In most cases, this meant that the outcome data collected at the end of students' first semesters were used in the analysis. However, in the 2004 sample, the drug and alcohol use measures were not administered until students' seventh semesters, and these were the data used in analyses. The results of predictions across all four years of college for some of these outcomes are contained in Schmitt, Billington, Keeney, et al. (2009).

The results of the regression analyses, including standardized beta coefficients and percentage of variance explained (R^2), are provided in Table 5. Table 5 also contains the zero-order correlations (validity coefficients) between the various outcomes and the predictors.

Organizational citizenship behaviors were significantly predicted in the combined regression (R^2 = .21). There were some differences in the correlations between these two samples. The Knowledge, Artistic, Multicultural, Leadership, Citizenship, Health, Careers, Adaptability, Perseverance and Jobs scales were all significant. These associations were generally in the positive direction, though the Knowledge, Artistic and Health scales showed evidence of suppression, with standardized regression coefficients that were negative despite positive zero-order correlations. Deviance was also predicted by these instruments (R^2 = .17). Knowledge, Artistic Appreciation, Ethics and Situational Judgment negatively predicted deviance, while Learning and Leadership positively predicted it. These scales contributed to an incremental improvement over HSGPA and to an overall significant R^2 for predicting this outcome.

The BARS scale was the best-predicted noncognitive outcome, with Artistic Appreciation, Multicultural Appreciation, Leadership, Health, Career Orientation, Adaptability, Perseverance and Ethics all positively contributing to a significant overall R^2 of .35. Voluntary absenteeism was negatively predicted by Knowledge, Health, Perseverance, Ethics and Situational Judgment but was positively predicted by Continuous Learning, Adaptability and the Jobs scale.

Turnover intent was weakly, but significantly, related to the predictors. Only the SJT emerged as a predictor in the combined regression, but it was still significant overall (R^2 = .03). The traditional cognitive outcome of first-year college GPA was predicted largely by HSGPA and SAT scores, although the Careers and Jobs scales were negative predictors. Correlations for several noncognitive predictors were statistically significant and of reasonable magnitude (e.g., Knowledge, Awards, Ethics and the SJT). The overall regression was significant and quite large (R^2 = .36), but again, this was largely due to the variables in the initial step of this regression.

Academic satisfaction was positively predicted by Adaptability and Perseverance but was negatively predicted by Leadership, though this appeared to be a suppressor effect. The overall variance explained was relatively low (R^2 = .05), but it was significant. Social satisfaction was positively predicted Multicultural Appreciation, Citizenship, Health and Adaptability, while it was negatively predicted by Knowledge and Artistic Appreciation. Again, these negative regression weights appeared to be a result of multicollinearity among the predictors. The overall regression was significant (R^2 = .08).

Problem drinking was positively predicted by Leadership, Health, Adaptability and the Jobs scale, while it was negatively predicted by the Learning, Careers, Ethics and Awards scales, as well as the SJT. The overall regression was significant (R^2 = .16). Drug use was positively predicted by Continuous Learning, Artistic Appreciation, Multicultural Appreciation and the Jobs scale, while it was negatively predicted by Ethics, yielding a significant overall regression (R^2 = .11).

Overall, these scales do predict a variety of outcomes. When both samples are taken into account, they significantly predict all of the outcomes (though weakly so for turnover intent and first-year college GPA). The predictive ability of these scales was notably greater than the traditional cognitive predictors for the noncognitive outcomes of OCBs, deviance, the BARS, problem drinking behaviors and drug use. These results indicate that, in general, the noncognitive measures achieve the goal of predicting these alternate measures of students' success.

These results indicate that, in general, the noncognitive measures achieve the goal of predicting these alternate measures of students' success.

Subgroup Differences

At the beginning of this report, we stated the belief that noncognitive measures of student potential would likely reveal small or nonexistent subgroup mean differences which might be very important since their use in the admission process may result in a more diverse student body. Weighted mean averages across the several samples from which we collected data were calculated to compare the performance of gender and racial subgroups on the predictor measures. The calculations were performed for three samples: a sample of Michigan State University and University of Illinois students collected in 2003 for which black students were deliberately oversampled, the sample of new freshmen collected from 10 universities in 2004, and the sample of college applicants collected in 2007–2008. The findings were meta-analyzed across the three samples using the conventions of Hedges and Olkin (1985). Following Cohen's (1988) convention, *d*-values above .2 were considered meaningful. Table 6 is a summary of the results of the meta-analysis. All subgroup differences in that table and the discussion below are presented in terms of standardized mean differences.

The first set of comparisons was between males and females. Males moderately outperformed females on the Health scale (.48) and the SAT/ACT composite (.41). Females moderately outperformed males on the SJI (-.44) and less so on the Artistic (-.25), Multicultural (-.24), Responsibility (-.29), Career (-.21) and Perseverance (-.23) scales. As was true in these samples, males typically outperform females on standardized tests of ability, such as the SAT and ACT. A higher score for men on the Health dimension may be a function of the inclusion of a number of items on the participation in exercise and sports activities on this scale.

The second set of comparisons was between white and black students. White students scored much higher than black students on HSGPA (.98) and the SAT/ACT composite (1.46). White students slightly outperformed black students on the Knowledge (.29), Artistic (.24), Health (.39), Ethics (.24) and Jobs (.34) scales. Black students moderately outperformed white students on the Career scale (-.49).

The third set of comparisons was between white and Asian students. White students scored slightly higher on the Health (.36) and Adaptability (.26) scales, and much higher on the Jobs scale (.86). Asian students scored moderately higher than white students on the Multicultural scale (-.52) and the SAT/ACT composite (-.41).

There were no systematic differences that indicate that the noncognitive measures may discriminate against a particular subgroup.

The fourth set of comparisons was between white and Hispanic students. White students outperformed Hispanic students on the SAT/ACT composite (.91) and HSGPA (.68), moderately outperformed them on the Jobs scale (.51), and slightly outperformed them on the Health scale (.22). Hispanic students slightly to moderately outperformed white students on the Multicultural scale (-.36).

In summary, although there were meaningful subgroup differences on many of the predictor variables, there were no systematic differences that indicate that the noncognitive measures may discriminate against a particular subgroup. In all cases, there were measures on which minority students outperformed majority students. For all comparisons involving ethnic minority groups, it was the case that the SAT/ACT score differences were larger than differences on the biodata or SJI.

Implications of Use as Admission Tools for Demographic Composition of the Student Body

Given the low levels of performance differences among ethnic subgroups on the biodata and SJI, we attempted to estimate how the use of the noncognitive measures in admission might impact the proportion of different subgroups admitted to the universities. We calculated two composites of predictors and rank ordered the sample of new freshman students from 10 universities whose responses to the measures were originally collected in 2004 based on their scores on the composite measures. The first composite included only the cognitive predictors. It was a sum of the standardized SAT/ACT scores and standardized high school grade point average. The second composite was an approximately equally weighted sum of the noncognitive measures (11 biodata scales and the SJI; the biodata; Awards and Jobs scales were excluded from the composite) and the cognitive measures (SAT/ACT and HSGPA). We computed the proportion of each of four groups (students who identified as Hispanic, Asian, black and white) who would have been admitted to the universities if the universities admitted a highly selective (top 15 percent), moderately selective (top 50 percent) or minimally selective (top 85 percent) cohort. The results of this analysis are presented in Table 7.

The use of the battery that included the noncognitive measures would have little impact on the proportion of students from different groups admitted under conditions of minimal selectivity. As selectivity increases, however, the proportion of Hispanic and black students who would be admitted if the noncognitive measures were included increased. Correspondingly, the proportion of white and Asian students would be smaller under those conditions.

An additional set of analyses sought to address the question of whether using a noncognitive measure in admission might affect the average academic performance observed of the hypothetically selected students. We found that a selection composite that included the noncognitive measure resulted in differences of no more than .10 in college GPA from a cognitive only composite. Differences were even smaller under conditions of minimal

selectivity. Seven institutions also provided information on students' four-year graduation status. Graduation rates were not very different under the two different selection scenarios, with the inclusion of the noncognitive composite producing the largest difference for African Americans under conditions of high selectivity (92 percent versus 100 percent graduation).

To examine whether any of the measures differentially predicted college student performance by race, we used the moderated multiple regression procedure recommended by Bartlett, Bobko, Mosier and Hannan (1978). The predictive validity for each minority group (i.e., African Americans, Hispanics, and Asians) was compared against that observed for white students. There was some evidence of predictive bias for the ability and school achievement for African American students; HSGPA had lower predictive validity, and SAT/ACT had higher predictive validity. However, there was no evidence that the noncognitive composite differentially predicted college GPA.

The above findings demonstrate that supplementing cognitive measures with noncognitive measures can reduce subgroup differences, while having only minimal impact on academic performance. It should be noted, however, that the noncognitive measures were not responded to in the usual high-stakes admission context. Also, because analyses were performed for a group of students who had already been admitted, range restriction in the cognitive composite may have underestimated some of the reported relationships. Finally, whereas our hypothetical scenarios for selecting students involved the use of unit-weighted composites, other weighting schemes (e.g., based on predictor intercorrelations or institutional values) could produce different results with respect to the demographic composition of a student body.

The findings demonstrate that supplementing cognitive measures with noncognitive measures can reduce subgroup differences.

Student Profiles and Counseling Interventions

Another attempt we have made is to identify groups of students who have similar profiles on the biodata, SJI and ability measures. The characteristics and performance of these groups of students were then compared to determine if we could identify groups of students with whom differing interventions might be helpful or effective in making them better students or improving their adjustment to the college situation.

Cluster analyses were conducted to identify groups of students who had empirically distinct motivational and ability profiles. A set of cognitive and noncognitive predictors were used to form the profiles (Schmitt et al, 2007). The student profiles were developed on the basis of a combination of biodata scales, the SJI measure of students' responses to hypothetical academic and nonacademic situations (SJI), and traditional indicators of ability (SAT/ACT and High School GPA). The cluster analysis revealed the following five clusters of students (cluster names are based on Schmitt et al., 2007):

1. ***Cluster 1: Low academic, career oriented —*** This group was characterized by the lowest average scores on the SAT/ACT dimension and the lowest HSGPA scores. On the SJI and most other biodata dimensions, these participants scored about average. On the career orientation dimension, they received the highest average score (closely followed by the fifth cluster, described below).

2. ***Cluster 2: High ability, culturally limited —*** This group was characterized by average SAT/ACT scores but relatively high HSGPA. These individuals' biodata scores were relatively low, especially on the artistic and diversity dimensions. However, their scores on the health, adaptability and ethics biodata dimensions were above average, as was their SJI score.

3. ***Cluster 3: Marginal —*** This group was characterized by SAT/ACT and HSGPA scores that averaged 0.5 standard deviations below the mean. On most of the biodata dimensions and the SJI, their scores were between 0.5 and 1.0 standard deviation units below the average of the total group. This group's average scores on the knowledge and perseverance dimensions were also particularly low.

4. ***Cluster 4: Able artistic —*** This group was characterized by the highest average SAT/ACT scores, the highest artistic biodata scores, and the second highest diversity biodata scores. Scores on most other biodata dimensions were average to below average, with the lowest scores being on the biodata dimensions of career orientation, adaptability and perseverance.

5. ***Cluster 5: Academically able, well rounded —*** This group's average standing on the SAT/ACT composite and HSGPA was nearly 0.5 standard deviation units above the average. Although they were not the brightest group as represented by these two cognitive indexes, their scores were the highest on the knowledge, continuous learning, diversity, leadership, social responsibility, adaptability, perseverance and ethics biodata dimensions as well as on the SJI.

The profile of scores for each of these clusters is also depicted in the profiles in Figure 1. The profiles of these five groups differ in both level and shape. The most notable demographic differences across clusters include the following:

1. Men were slightly more likely to be members of the marginal group and the high ability, culturally limited group than were women. It is perhaps surprising that women were nearly twice as likely to fit the profile of the low academic, career-oriented category as were men.

2. Hispanic and African American students were four to five times as likely to be members of the low academic, career-oriented cluster as were Asian and Caucasian students. They were also somewhat more likely to be members of the marginal group and less likely to be identified as members of the able and artistic cluster than were Caucasians and, in particular, Asian American students.

3. Engineering majors were disproportionately represented in the high ability, culturally limited cluster. Those in the fine arts and humanities were heavily represented in the able and artistic group. Students in the natural sciences were most likely to be in the high-ability, culturally limited group and the academically able, well-rounded group. Students who had not declared a major or whose major was in the other category were more likely to be in the marginal cluster. Finally, students with an undeclared major were also unlikely to be in the low academic, career-oriented cluster and likely to be in the able and artistic cluster, which suggests that many students in this category are quite unsure of what they want to do and perhaps why they are in college, despite their potential to do well.

There were also differences among the five subgroups, as characterized by these profiles, on various performance and motivational outcomes external to the variables on which the profiles were based. On all outcome dimensions considered, there were statistically significant and interpretable differences among the students in the five clusters; that is, students among the five clusters earned different grades in college; contributed differentially in nonacademic ways; and differed in satisfaction levels, class attendance, and stated intention to leave college. Table 8 includes a summary of the differences in outcomes associated with students in the five clusters.

Interventions Implications of Cluster Membership

These cluster results may have practical implications for designing interventions for students that could be used in high school counseling settings, in college admission, and in identification of college students who are at high risk for college failure. These profiles provide a holistic understanding of student's motivation, cognitive ability and behavior and may be effective tools for designing interventions and career guidance programs tailored to students' specific deficiencies. Below are some suggestions for different interventions that might be introduced to aid students to adapt to college life and optimize their college experience (Schmitt et al., 2007). Further research is required to test for the influence of such interventions on student performance outcomes.

These cluster results may have practical implications for designing interventions for students that could be used in high school counseling settings, in college admission, and in identification of college students who are at high risk for college failure.

Schmitt et al. (2007) also speculated as to interventions that might be most appropriate with students in each of the five clusters. Students in the Low Academic, Career-Oriented group usually come from minority groups who have had less-than-ideal educational opportunities. These students are highly motivated and career-oriented but lack essential academic skills. The best intervention for such students would be one that provides additional academic skills and one that helps the students to understand the link between college work and future job and career possibilities. Alternatively, those who are dissatisfied with university life should be best advised to attend technical programs or schools of interest to them in which they receive preparation that is more directly linked to a specific job or occupation.

Students in the High Ability, Culturally Limited group do well academically, engage in citizenship behaviors and are usually satisfied with their academic experience. However, such students are low on continuous learning and artistic and multicultural appreciation. Interventions that provide exposure to a wider range of academic and cultural experiences and motivate students to engage in them would help to make these students more well rounded.

Students in the Marginal group have a low standing on average across all cognitive and noncognitive profile variables. Such students are at the highest immediate risk for failure in college and thus are most in need of immediate and wide-ranging interventions. Interventions for such students should be aimed at providing academic skills and vocational counseling, thereby providing students with direction (e.g., help in selection of major) and appropriate career guidance counseling.

Able and Artistic group consists of students that do well academically, are particularly high on artistic and multicultural appreciation and are on the whole satisfied with their college experience. Such students do not necessarily need targeted interventions to be successful in college. If these individuals have any problems in college, they are likely to be the result of the students' own high standards for performance.

Finally, the Academically Able, Well-Rounded group of students is the complete opposite of the marginal group. These students have clear educational goals, are academically able and are well rounded. These students do not require developmental interventions. In fact, they would make ideal peer mentors who can provide support to students in the marginal and low academic groups.

Conclusions and Future Work

The work described in this report was actually several projects conducted over the past eight years with the cooperation of multiple researchers, at least two dozen universities and their support personnel, thousands of college students, and the support of the College Board. Any effort to summarize the findings of this research will necessarily be limited, but we feel the following are the most noteworthy conclusions. First, the ACT/SAT and HSGPA are very good predictors of college GPA, particularly early in students' careers; this is an affirmation of a large body of literature. Second, and more relevant to our purpose, is the fact that noncognitive attributes as we have measured them do correlate with college GPA, but the incremental validity associated with these measures is relatively small. Third, the noncognitive measures are correlated with other valued dimensions of student performance beyond the achievement reflected in college grades. They were especially strongly related to students' self-rated performance on our targeted dimensions (BARS) and various organizational citizenship behaviors and, to a lesser extent, with negative behavior such as class absenteeism, deviance or destructive behavior, drug and alcohol abuse and intent to leave the university. All of these latter behaviors were minimally related to traditional predictors of college student success (i.e., SAT/ACT or HSGPA). Fourth, there were much smaller differences between males and females and ethnic subgroups compared to the differences displayed by cognitive predictors; use of a battery of admission tools that includes both sets or measures will usually result in a more demographically diverse student body at little or no decrement in average student performance.

There is need for more research, particularly in situations in which the noncognitive measures are actually used to make admissions decisions. The 2007 sample described above did respond to the measures during the admission process, but they were told that their responses would not be used to make admission decisions. The need to evaluate responses in an actual decision making situation is critical, given the concern with student motivation to do well and the evidence that students can and do inflate their responses in these situations. Studies on the role of faking summarized in this paper provide some suggestions on the control of response inflation, but these efforts will not likely resolve this issue. We also think that more research could be usefully directed to the use of these measures as counseling

tools that may remove some of the concern about faking. The profiling effort described the possibility of some interventions for different groups of people, but our primary focus has been on prediction rather than remediation or interventions with students who might develop problems in college. Finally, because our measures are designed to measure noncognitive attributes, they might be particularly useful in identifying able students who would otherwise not be given opportunities based solely on cognitive indices. The degree to which this is possible with populations who typically score low on traditional measures should be further examined.

References

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel, 25*, 297–308.

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.

Beatty, A., & Greenwood, R., & Linn, R.L. (Eds.) (1999). *Myths and tradeoffs: The role of tests in undergraduate admissions*. Washington D.C.: National Academy Press.

Becker, T. E., & Colquitt, A. L. (1992). Potential versus actual faking of a biodata form: An analysis along several dimensions of item type. *Personnel Psychology*, *45*, 389–406.

Camara, W. J., & Kimmel, E. W. (2005). (Eds.). *Choosing students: Higher education admissions tools for the 21st century*. Mahwah, NJ: Lawrence Erlbaum Associates.

Clevenger, J., Pereira, G. M., Wiechmann, D., Schmitt, N., & Harvey, V. S. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology*, *86*, 410–417.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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.

Cress, C. M., Astin, H. S., Zimmerman-Oster, K., & Burkhardt, J. (2001). Developmental outcomes of college students' involvement in leadership activities. *Journal of College Student Development*, *42*, 15–27.

Cunningham, M. R., Wong, D. T., & Barbee, A. P. (1994). Self-presentation dynamics on overt integrity tests: Experimental studies of the Reid Report. *Journal of Applied Psychology*, *79*, 643–658.

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. Report submitted to the College Board. E. Lansing, MI: Michigan State University.

Dwight, S. A., & Alliger, G. M. (1997). Reactions to overt integrity test items. *Educational and Psychological Measurement*, *57*(6), 937–948.

Dwight, S. A., & Donovan, J. J. (2003). Do warnings not to fake reduce faking? *Human Performance*, *16*, 1–23.

Eaton, S., & Bean, J. P. (1995). An approach/avoidance behavioral model of college student retention. *Research in Higher Education*, *36*, 617–645.

Gillespie, M., Oswald, F. L., Schmitt, N., Manheim, L., & Kim, B. (2002). Construct validation of a situational judgment test of college student success. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology. Toronto, Canada.

Gough, H. G., & Arvey, R. D. (1998). *Personnel Reaction Blank*. Palo Alto, CA: Consulting Psychologists Press.

Griffeth, R. W., Hom, P. W. (1988). A comparison of different conceptualizations of alternatives in turnover research. *Journal of Organizational Behavior*, *13*, 103–111.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hedlund, J., Plamondon, K., Wilt, J., Nebel, K., Ashford, S., & Sternberg, R. J. (2001, April). Practical intelligence for business: going beyond the GMAT. Paper presented in J. Cortina (Chair), *Out with the old, in with the new: Looking above and beyond what we know about cognitive predictors.* Symposium presented at the 16th Annual Convention of the Society for Industrial and Organizational Psychology, San Diego, CA.

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., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, *75*, 581–595.

Huffcutt, A. I., Weekley, J. A., Wiesner, W. H., DeGroot, T. G., & Jones, C. (2001). Comparison of situational and behavior description interview questions for higher-level positions. *Personnel Psychology*, *54*, 619–644.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average*. New York: College Board Research Report No. 2008-5.

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.

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., Bruhn-Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Predicting job performance using situational judgment tests. *Journal of Applied Psychology*, *86*, 730–740.

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

Miller, C. E. (2001). The susceptibility of personality selection tests to coaching and faking. (Doctoral dissertation, University of Akron, 2000). *Dissertation Abstracts International*, *61*, 3888.

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. (2001). *Rationally-keyed biographical data items for constructs related to sales performance*. Fairfax, VA: George Mason University.

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

Oswald, F. L., Friede, A. J., Schmitt, N., Kim, B. K., & Ramsay, L. J. (2005). Extending a practical method for developing alternate test forms using independent sets of items. *Organizational Research Methods*, *8*, 149–164.

Oswald, F. L., Schmitt, N., Kim, B. H., Gillespie, M. A., & Ramsay, L. J. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology*, *89*, 187–207.

Owens, W. A., Albright, L. E., & Glennon, J. R. (1966). *A catalog of life history items*. Chicago, IL: The Creativity Research Institute and the Richardson Foundation.

Patelis, T., & Camara, W. (1999). *Preliminary Report of Statewide Assessments in the United States*. The College Board: New York, NY

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*, 598–609.

Pulakos, E. D., & Schmitt, N. (1996). An evaluation of two strategies for reducing adverse impact and their effects on criterion-related validity. *Human Performance*, *9*, 241–258.

Pulakos, E. D., Schmitt, N., & Keenan, P. A. (1994). *Validation and implementation of the FBI Special Agent entry level selection system*. Alexandria, VA: HumRRO, Report No. FR-PRO-94 20.

Ra, J. B. (1989). Validity of a new evaluative scale to aid admissions decisions. *Evaluation and Program Planning*, *12*, 195–204.

Ramsay, L., Kim, B., Gillespie, M., & Friede, A. (2003). Score Inflation on Biodata and Situational Judgment Inventory Items Used in College Admissions. Report submitted to the College Board. E. Lansing, MI: Michigan State University.

Ramsay, L., Schmitt, N., Oswald, F. L., Kim, B. H., & Gillespie, M. (2006). The impact of situational context variables on responses to biodata and situational judgment inventory items. *Psychology Science*, *48*, 268–287.

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.

Russell, C. J., Green, S. G., Griggs, D. (undated). Biographical questionnaire items designed to discriminate between transitional and transformational leaders. W. Lafayette, IN: Purdue University.

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.

Schmitt, N., Billington, A. Q., Golubovich, J., Keeney, J., Pleskac, T., Reeder, M., et al. (2009). Report of the First-Year Follow-up of College Applicants at Twelve Universities. Michigan State University: Submitted to the College Board.

Schmitt, N., Billington, A. Q., Keeney, J., Oswald, F. L., Pleskac, T., Sinha, R., et al. (2009). Prediction of four-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students. *Journal of Applied Psychology*, *94*, 1479–1498.

Schmitt, N., & Kunce, C. (2002). The effect of required elaboration of answers to noncognitive measures. *Personnel Psychology*, *55*, 569–588.

Schmitt, N., Oswald, F. L., & Gillespie, M. A. (2005). Implications of broadening the performance domain for the prediction of academic success. In W. F. Camara and E. Kimmel (Eds.). *Choosing students: Higher Education Admission Tools for the 21st Century* (pp. 195–214). Mahwah, NJ: Lawrence Erlbaum Associates.

Schmitt, N., Oswald, F. L., Kim, B. H., Gillespie, M. A., & Ramsay, L. J. (2004). The impact of justice and self-serving bias explanations for the perceived fairness of different types of selection tests in college admissions. *International Journal of Selection and Assessment*, *12*, 160–171.

Schmitt, N., Oswald, F. L., Kim, B. H., Gillespie, M. A., Ramsay, L. J., & Yoo, T. (2003). Impact of elaboration on social desirability and the validity of biodata measures. *Journal of Applied Psychology*, *88*, 979–988.

Schmitt, N., Oswald, F. L., Kim, B. H., Imus, A., Merritt, S., Friede, A.,et al. (2007). The use of background and ability profiles to predict college student outcomes. *Journal of Applied Psychology*, *92*, 165–179.

Sedlacek, W. E. (2004). *Beyond the big test: Noncognitive assessments in higher education*. San Francisco: Jossey-Bass.

Sternberg, R. J., & The Rainbow Project Collaborators. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence*, *34*, 321–350.

Taber, T. D., & Hackman, J. D. (1976). Dimensions of undergraduate college performance. *Journal of Applied Psychology*, *61*, 546–558.

University of Pennsylvania, Graduate School of Education. (2000). *Reporting on issues in Education reform* (CPRE Publication RB-31-June). Consortium for Policy Research in Education.

Vasilopoulos, N. L. (1999). The impact of job familiarity and warning of response verification on the relationship between response latency and impression management: A field investigation. *Dissertation Abstracts International*, *59*, 4521.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, *59*(2), 197–210.

Wightman, L. F., & Jaeger, R. M. (1998). High stakes and ubiquitous presence: An overview and comparison of *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.

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

Wilds, D. J., & Wilson, R. (1998). *Minorities in higher education, 1997–1998: Sixteenth annual status report*. Washington, DC: American Council on Education. Office of Minority Concerns.

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
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
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
Appreciating art and culture, either at an expert level or simply at the level of one who is interested.

Multicultural appreciation
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
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
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
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
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.

Career orientation
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
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
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
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.

Table 2

Means, Standard Deviations and Intercorrelations Between Measures of the Biodata Dimensions and SJI*

	N	*Mean*	SD	Knowl-edge	Learn-ing	Artistic	Multi-cultural	Leader-ship	Respon-sibility	Health	Career	Adapt-ability	Perse-verance	Ethics	SJT
Knowledge	2765	3.15	.47	*.67*	.65	.36	.37	.36	.35	.42	.32	.50	.70	.64	.46
Learning	2765	3.09	.61	.47	*.78*	.54	.65	.43	.39	.19	.40	.31	.49	.34	.27
Artistic	2765	2.91	.82	.27	.44	*.86*	.72	.40	.46	.02	.05	.10	.21	.27	.26
Multicultural	2765	2.98	.66	.27	.51	.60	*.80*	.46	.52	.05	.20	.20	.33	.22	.28
Leadership	2768	3.07	.81	.28	.35	.34	.38	*.86*	.64	.26	.33	.40	.52	.23	.25
Responsibility	2768	3.32	.76	.25	.31	.38	.41	.53	*.79*	.14	.31	.24	.41	.36	.40
Health	2768	3.25	.51	.28	.14	.01	.04	.20	.11	*.68*	.13	.69	.43	.27	.16
Career	2768	3.32	.65	.23	.31	.04	.16	.27	.24	.09	*.77*	.28	.54	.23	.17
Adaptability	2768	3.38	.45	.33	.22	.08	.15	.30	.17	.46	.20	*.65*	.68	.28	.23
Perseverance	2768	3.73	.49	.50	.38	.17	.26	.42	.31	.31	.41	.47	*.75*	.44	.36
Ethics	2768	3.86	.54	.43	.25	.20	.16	.18	.26	.19	.17	.18	.31	*.67*	.61
SJI	2730	.66	.33	.33	.21	.21	.22	.21	.31	.12	.13	.16	.27	.44	*.76*

*SJI = situational judgment inventory, Learning = Continuous Learning, Multicultural = Multicultural Appreciation, Career = Career Orientation. Correlations above .04 significant ($p < .05$). 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. These data were derived from the responses of the set of respondents who first responded to our survey in 2004.

Table 3

Comparison of Applicants to Incumbent Students in Previous Data Collection

	Average Score of Applicants	*N*	Average Score of Incumbents	*N*	Difference (*d*-value)
Knowledge	3.50 (.48)	725	3.15 (.47)	2711	.74
Learning	3.43 (.61)	665	3.09 (.61)	2711	.56
Artistic	3.26 (.76)	663	2.91 (.82)	2711	.44
Multicultural	3.30 (.70)	821	2.98 (.66)	2711	.47
Leadership	3.31 (.77)	666	3.07 (.81)	2714	.30
Responsibility	3.70 (.74)	722	3.32 (.76)	2714	.51
Health	3.28 (.51)	725	3.25 (.51)	2714	.06
Career	3.40 (.63)	729	3.32 (.65)	2714	.12
Adaptability	3.49 (.43)	663	3.38 (.45)	2714	.25
Perseverance	3.90 (.47)	665	3.73 (.49)	2714	.35
Ethics	4.09 (.42)	723	3.86 (.54)	2714	.48
SJT	.40 (.14)	507	.33 (.17)	2676	.45

Note: Standard deviations are in parentheses next to the means. SJT = Situational Judgment Test.

Table 4

Correlations Between Predictor and Outcome Variables (Combined Meta-analytically) in the 2004 and 2007 Samples and the Intercorrelations Between the Outcomes

	2004 Mean	2004 SD	2007 Mean	2007 SD	Knowl-edge	Learning	Artistic	Multicul-tural	Leader-ship	Citizen-ship	Health	Career	Adapt-ability	Perse-verance	Ethics
BARS	3.63	.48	3.74	0.46	**.35**	.36	.31	.35	.39	.30	.29	.28	.35	.45	.31
College GPA	3.31	.73	3.49	0.62	**.24**	.10	.16	.10	.12	.14	.12	-.12	.05	.09	.18
Absenteeism	1.44	.29	2.74	1.09	**-.17**	-.02	-.04	-.01	-.01	-.07	-.10	-.02	-.01	-.14	-.23
OCB	3.28	.63	4.38	0.94	**.11**	.15	.09	.23	.32	.31	.06	.24	.20	.31	.11
Deviance	1.39	.39	1.56	0.51	**-.22**	-.06	-.11	-.06	.01	-.07	-.04	-.03	-.07	-.13	-.38
Academic Sat.	4.15	.52	4.20	0.56	**.14**	.09	.05	.05	.04	.08	.12	.06	.14	.15	.11
Social Sat.	3.81	.87	3.89	0.87	.02	.06	.01	.11	.13	.14	.16	.10	.21	.15	.03
Drug Use	1.39	.83	1.31	0.70	**-.04**	.07	.08	.09	-.04	-.06	-.01	-.07	.03	-.06	-.21
Problem Drinking	2.49	.97	7.65	3.45	**-.11**	-.10	-.06	-.05	.05	-.01	.09	-.11	.07	-.07	-.24
Turnover Intentions	1.41	.63	1.24	0.58	**-.08**	-.03	-.05	-.04	-.03	-.08	-.07	-.02	-.08	-.08	-.08

Note: SJI = Situational Judgment Inventory. OCB = Organizational Citizenship Behavior. Sat.= Satisfaction. H.S.=High School. Cronbach's alpha values are displayed along the diagonal; the first value displayed is the 2007 alpha and the second is the 2004 alpha. Significant correlations are in bold and are at the $p < .05$ level.

Table 4 (*continued*)

Correlations Between Predictor and Outcome Variables (Combined Meta-analytically) in the 2004 and 2007 Samples and the Intercorrelations Between the Outcomes

	Awards	Jobs	SJT	H.S. GPA	SAT/ ACT	BARS	College GPA	Absen-teeism	OCB	Devi-ance	Aca-demic Sat.	Social Sat.	Drug Use	Problem Drink-ing	Turn-over Inten-tions
BARS	**.22**	**.06**	**.24**	**.08**	**.08**	*.72, .74*	**.18**	**-.17**	**.37**	**-.22**	**.25**	**.30**	**-.13**	**-.16**	**-.23**
College GPA	**.20**	**-.08**	.19	**.51**	**.49**	**.24**	*1.00*	**-.17**	-.05	-.05	**.13**	-.04	-.02	.08	**-.15**
Absenteeism	-.05	.07	**-.16**	-.06	.06	**-.21**	**-.23**	*1.00*	.00	**.26**	-.03	**.08**	**.12**	**.17**	.00
OCB	**.13**	**.15**	**.14**	.00	**-.09**	**.46**	.04	-.02	*.83, .88*	.04	**.12**	**.38**	-.04	.03	-.15
Deviance	-.06	.07	**-.24**	**-.13**	-.06	**-.26**	**-.18**	**.27**	.04	.77, .81	**-.15**	.03	**.20**	**.25**	**.17**
Academic Sat.	.06	-.01	.10	.09	.08	**.30**	**.16**	**-.17**	**.26**	**-.13**	*.83, .81*	**.32**	.02	-.01	**-.36**
Social Sat.	.02	**.08**	**.06**	-.02	-.02	**.34**	-.03	-.03	**.51**	.00	**.30**	*.88, .91*	.00	**.09**	**-.34**
Drug Use	-.07	**.12**	-.07	**-.08**	.03	**-.16**	-.07	**.24**	-.06	**.31**	-.05	-.03	*.61, .64*	**.39**	**.07**
Problem Drinking	**-.09**	**.22**	-.14	-.06	.05	**-.21**	**-.11**	**.23**	.05	**.38**	-.02	**.09**	**.44**	*.82, .76*	-.06
Turnover Intentions	-.02	-.02	**-.11**	-.06	-.07	**-.17**	-.07	**.12**	**-.17**	**.11**	**-.29**	**-.22**	.05	-.01	*.82, .76*

Note: SJI = Situational Judgment Inventory. OCB = Organizational Citizenship Behavior. Sat.= Satisfaction. H.S.=High School. Cronbach's alpha values are displayed along the diagonal; the first value displayed is the 2007 alpha and the second is the 2004 alpha. Significant correlations are in bold and are at the $p < .05$ level.

Table 5

Incremental Validity of Biodata and SJT: Hierarchical Regression Results

	OCB						Deviance					
	2004		2007		Com.		2004		2007		Com.	
Step 1	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***
HSGPA	-.05	.02	.06	.09	.00	.06	-.13	-.16*	-.14	-.06	-.13	-.13*
ACT/SAT	-.10	.11*	-.08	-.14*	-.09	-.12*	-.06	.00	-.07	-.06	-.06	.00
ΔR^2		.02*		.02*		.01*		.03*		.01		.02*
Step 2												
Knowledge	.08	.07	.16	-.12	.11	-.06*	-.21	-.01	-.23	-.19*	-.22	-.08*
Learning	.10	.02	.21	.07	.15	-.04	-.06	.01	-.06	.07	-.06	.07*
Artistic	.06	.10*	.14	-.10	.09	-.09*	-.09	-.08*	-.13	-.08	-.11	-.07*
Multicultural	.20	-.08*	.28	.18*	.23	.15*	-.04	.06	-.09	-.09	-.06	.01
Leadership	.27	-.18*	.40	.12	.32	.14*	-.01	.09*	.04	.21*	.01	.10*
Citizenship	.28	-.18*	.34	.22*	.31	.16*	-.07	-.01	-.07	.02	-.07	.01
Health	.01	.02	.12	.01	.06	-.06*	-.08	.00	.02	.04	-.04	.05
Career	.22	-.05	.26	.07	.24	.07*	-.05	-.01	-.01	.08	-.03	.01
Adaptability	.18	-.05	.23	.25	.20	.07*	-.09	-.06	-.04	-.12	-.07	-.03
Perseverance	.27	-.12*	.36	.20*	.31	.16*	-.15	-.04	-.11	.07	-.13	-.02
Ethics	.11	-.03	.10	.00	.11	.02	-.41	-.35*	-.34	-.17*	-.38	-.32*
Awards	.10	.01	.15	.01	.13	-.02	-.03	.07*	-.08	-.03	-.06	.00
Jobs	.13	-.04	.18	.07	.15	.08*	.04	-.01	.11	.08	.07	.05
SJT	.13	.01	.16	-.03	.14	.02	-.27	-.14*	-.19	-.13*	-.24	-.09*
ΔR^2		.16*		.29*		.20*		.19*		.15*		.15*
Overall R^2		.18*		.30*		.21*		.22*		.16*		.17*
Adj. R^2		.17		.27		.20		.21		.12		.17
N		920		323		1246		920		324		1244

* Indicates a coefficient is significant at $p < .05$.
Note: *b* refers to standardized regression coefficients.

Table 5 (continued)

Incremental Validity of Biodata and SJT: Hierarchical Regression Results

	BARS						Absenteeism					
	2004		2007		Com.		2004		2007		Com.	
Step 1	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***
HSGPA	.08	.08*	.09	.03	.08	.05	-.06	-.13	-.07	-.08	-.06	-.12*
ACT/SAT	.08	-.02	.08	.09	.08	.05	.06	.11	.06	.05	.06	.12*
ΔR^2		.01*		.01		.01*		.02*		.01		.01*
Step 2												
Knowledge	.34	.04	.36	.05	.35	.02	-.24	-.13*	-.08	-.18*	-.17	-.11*
Learning	.35	.05	.37	.09	.36	.04	-.10	-.05	.07	-.12	-.02	.06*
Artistic	.28	.11*	.34	.12	.31	.11*	-.04	.03	-.04	-.02	-.04	-.04
Multicultural	.36	.13*	.33	.08	.35	.11*	-.05	.02	.04	.09	-.01	.05
Leadership	.38	.12*	.40	.09	.39	.13*	-.01	.08*	-.02	.09	-.01	.05
Citizenship	.31	.00	.29	.07	.30	.01	-.06	-.01	-.08	-.12	-.07	-.04
Health	.31	.15*	.27	.11	.29	.13*	-.12	-.09*	-.08	-.11	-.10	-.08*
Career	.27	.09*	.30	.14	.28	.09*	-.05	.04	.01	-.04	-.02	.04
Adaptability	.38	.15*	.32	.00	.35	.09*	-.05	.03	.04	.28*	-.01	.10*
Perseverance	.44	.12*	.46	.22*	.45	.17*	-.18	-.05	-.08	-.06	-.14	-.09*
Ethics	.34	.18*	.26	.07	.31	.12*	-.30	-.21*	-.15	-.14*	-.23	-.16*
Awards	.20	-.03	.24	.01	.22	-.03	-.06	-.03	-.04	-.09	-.05	-.01
Jobs	.03	-.01	.10	.02	.06	.01	.07	.05	.08	.09	.07	.06*
SJT	.24	.01	.24	-.03	.24	.01	-.19	-.09*	-.13	-.07	-.16	-.08*
ΔR^2		.37*		.40*		.34*		.12*		.15*		.09*
Overall R^2		.38*		.41*		.35*		.14*		.15*		.10*
Adj. R^2		.37		.38		.34		.13		.11		.09
N		1013		322		1335		844		323		1167

* Indicates a coefficient is significant at $p < .05$.
Note: *b* refers to standardized regression coefficients.

Table 5 (*continued*)

Incremental Validity of Biodata and SJT: Hierarchical Regression Results

	Turnover Intent						College GPA					
	2004		2007		Com.		2004		2007		Com.	
Step 1	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***
HSGPA	-.06	-.03	-.07	-.07	-.06	-.03	.56	.41*	.39	.23*	.51	.36*
ACT/SAT	-.09	-.08*	-.05	-.06	-.07	-.05	.47	.30*	.52	.43*	.49	.32*
ΔR^2		.01*		.01		.01*		.35*		.32*		.34*
Step 2												
Knowledge	-.05	.07	-.11	.02	-.08	-.02	.25	.05	.22	.13	.24	.04
Learning	-.03	-.01	-.04	-.06	-.03	.03	.09	.01	.11	-.07	.10	.00
Artistic	-.05	-.03	-.05	-.02	-.05	-.03	.16	.01	.15	-.03	.16	.02
Multicultural	-.04	.06	-.05	.02	-.04	.00	.09	.01	.13	.14	.10	.03
Leadership	-.03	.06	-.02	.15*	-.03	.05	.12	.00	.13	-.02	.12	.01
Citizenship	-.10	-.10*	-.06	-.12	-.08	-.06	.16	.01	.08	.04	.14	.01
Health	-.10	-.05	-.03	.03	-.07	-.03	.17	.08*	.02	-.04	.12	.04
Career	-.02	-.03	-.01	.08	-.02	.00	-.12	-.09*	-.13	-.12*	-.12	-.09*
Adaptability	-.11	-.09*	-.04	.03	-.08	-.04	.07	-.02	.00	-.07	.05	-.04
Perseverance	-.06	-.05	-.11	-.14	-.08	-.04	.07	-.04	.15	.08	.09	.04
Ethics	-.05	-.01	-.13	-.14*	-.08	-.01	.18	.01	.17	.04	.18	.04
Awards	-.03	.07*	-.02	.04	-.02	.03	.21	-.03	.20	.05	.20	.01
Jobs	.00	.02	-.05	-.08	-.02	-.01	-.05	.02	-.14	-.03	-.08	-.07*
SJT	-.08	-.03	-.15	-.12	-.11	-.07*	.25	.14*	.05	-.09	.19	.03
ΔR^2		.03*		.10*		.02*		.04*		.05		.02*
Overall R^2		.04*		.11*		.03*		.39*		.36*		.36*
Adj. R^2		.02		.06		.02		.38		.33		.36
N		910		323		1233		1308		325		1633

* Indicates a coefficient is significant at $p < .05$.
Note: *b* refers to standardized regression coefficients.

Table 5 (*continued*)

Incremental Validity of Biodata and SJT: Hierarchical Regression Results

	Academic Satisfaction						Social Satisfaction					
	2004		2007		Com.		2004		2007		Com.	
Step 1	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***
HSGPA	.10	.06	.07	.02	.09	.07*	-.07	-.09*	.04	.04	-.02	-.01
ACT/SAT	.08	.06	.07	.08	.08	.05	-.03	.01	.00	-.01	-.02	-.01
ΔR^2		.01*		.01		.01*		.01*		.00		.00
Step 2												
Knowledge	.14	.05	.15	.10	.14	.02	.00	-.12*	.04	-.11	.02	-.09*
Learning	.06	-.04	.12	-.01	.09	.02	.06	-.02	.06	.01	.06	-.02
Artistic	.05	.01	.04	-.07	.05	.01	.00	-.06	.02	-.14	.01	-.08*
Multicultural	.04	-.05	.07	.06	.05	-.01	.10	.08	.12	.16	.11	.11*
Leadership	.05	-.03	.03	-.15*	.04	-.08*	.11	.02	.15	-.09	.13	.02
Citizenship	.08	.04	.09	.12	.08	.05	.10	.08	.20	.19*	.14	.10*
Health	.12	.01	.13	.10	.12	.05	.15	.07	.17	.19*	.16	.09*
Career	.05	-.02	.07	.08	.06	.02	.11	.08*	.08	.03	.10	.03
Adaptability	.15	.10*	.13	.01	.14	.07*	.23	.17	.18	.05	.21	.14*
Perseverance	.15	.13*	.14	.00	.15	.09*	.14	.07	.16	.08	.15	.06
Ethics	.06	-.04	.18	.22*	.11	.02	.00	-.02	.06	.09	.03	-.02
Awards	.07	-.02	.05	.04	.06	-.01	-.01	-.08	.05	.01	.02	-.04
Jobs	-.01	-.04	-.01	.01	-.01	-.03	.07	.01	.10	.05	.08	.04
SJT	.13	.08*	.07	-.04	.10	.03	.03	-.01	.09	-.06	.06	.01
ΔR^2		.05*		.13*		.04*		.09*		.12*		.08*
Overall R^2		.06*		.13*		.05*		.10*		.12*		.08*
Adj. R^2		.04		.09		.04		.08		.08		.07
N		934		325		1259		933		325		1258

* Indicates a coefficient is significant at $p < .05$.
Note: *b* refers to standardized regression coefficients.

Table 5 (*continued*)

Incremental Validity of Biodata and SJT: Hierarchical Regression Results

	Problem Drinking						Drug Use					
	2004		2007		Com.		2004		2007		Com.	
Step 1	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***	***r***	***b***
HSGPA	.02	-.06	-.12	-.14*	-.06	-.11*	-.05	-.09	-.10	-.09	-.08	-.12*
ACT/SAT	.14	.16*	-.02	.03	.05	.10*	.06	.10	-.01	.01	.03	.09*
ΔR^2		.02*		.02		.01*		.01		.01		.01*
Step 2												
Knowledge	-.09	.09	-.13	.00	-.11	.02	.01	.08	-.08	-.02	-.04	.04
Learning	-.08	.01	-.11	-.10	-.10	-.07*	.13	.15*	.03	.00	.07	.09*
Artistic	-.08	.00	-.05	.03	-.06	-.02	.09	.10	.07	.14*	.08	.08*
Multicultural	-.12	-.07	.00	.05	-.05	.02	.07	.02	.11	.06	.09	.10*
Leadership	.05	.22*	.05	.12	.05	.10*	-.05	.03	-.03	-.04	-.04	-.06
Citizenship	-.05	.06	.01	.07	-.01	.05	-.11	-.10	-.03	.03	-.06	-.05
Health	.10	.05	.09	.07	.09	.11*	.01	.01	-.02	-.04	-.01	.02
Career	-.13	-.06	-.10	-.19*	-.11	-.10*	-.07	-.05	-.07	-.14*	-.07	-.06
Adaptability	.03	-.01	.09	.06	.07	.07*	.04	.06	.03	.10	.03	.06
Perseverance	-.11	-.10	-.05	.06	-.07	-.04	-.05	.00	-.06	-.04	-.06	-.03
Ethics	-.26	-.25*	-.22	-.18*	-.24	-.22*	-.23	-.26*	-.19	-.17*	-.21	-.24*
Awards	-.07	-.18*	-.13	-.10	-.09	-.09*	-.08	-.15*	-.06	-.03	-.07	-.06
Jobs	.19	.14*	.24	.18*	.22	.20*	.06	.05	.16	.20*	.12	.13*
SJT	-.23	-.20*	-.08	-.08	-.14	-.07*	-.14	-.06	-.02	.06	-.07	.01
ΔR^2		.21*		.13*		.15*		.13*		.11*		.10*
Overall R^2		.23*		.15*		.16*		.14*		.11*		.11*
Adj. R^2		.20		.10		.15		.11		.07		.10
N		457		319		776		461		322		783

* Indicates a coefficient is significant at $p < .05$.
Note: *b* refers to standardized regression coefficients.

Table 6

Meta-analytic Findings for Subgroup Differences from Three Samples

	Male–Female			White–Black			White–Asian			White–Hispanic		
	g	*Lower bound*	*Upper bound*	*g*	*Lower bound*	*Upper bound*	*g*	*Lower bound*	*Upper bound*	*g*	*Lower bound*	*Upper bound*
Knowledge	-.06	-.09	-.03	*.29*	.25	.33	.02	-.03	.08	.09	.03	.16
Learning	.18	.14	.22	.09	.04	.14	-.16	-.23	-.09	.00	-.09	.09
Artistic	*-.25*	-.30	-.20	*.24*	.17	.30	-.15	-.24	-.06	-.03	-.14	.09
Multicultural	*-.24*	-.29	-.20	-.07	-.12	-.02	*-.52*	-.59	-.45	*-.36*	-.45	-.27
Leadership	-.15	-.21	-.10	.12	.06	.19	.14	.06	.23	.07	-.04	.19
Responsibility	*-.29*	-.34	-.24	.15	.09	.21	-.08	-.17	.00	.02	-.08	.13
Health	*.48*	.45	.51	*.39*	.35	.43	*.36*	.31	.42	*.22*	.15	.29
Career	*-.21*	-.25	-.17	*-.49*	-.54	-.44	.07	.00	.14	-.09	-.18	.00
Adaptability	.04	.01	.07	.10	.07	.14	*.26*	.21	.31	.11	.05	.17
Perseverance	*-.23*	-.26	-.20	-.16	-.20	-.12	.19	.14	.25	-.07	-.14	.00
Ethics	-.19	-.22	-.16	*.24*	.20	.29	.12	.06	.18	.18	.11	.25
Awards	-.15	-.20	-.10	.07	.01	.13	-.02	-.11	.06	.14	.04	.24
Jobs	-.03	-.09	.03	*.34*	.27	.41	*.86*	.76	.97	*.51*	.39	.64
SJI	*-.44*	-.45	-.42	.08	.05	.11	.17	.14	.20	.05	.01	.09
HSGPA	-.08	-.11	-.05	*.98*	.95	1.01	-.09	-.13	-.04	*.68*	.63	.73
SAT/ACT	*.41*	.35	.47	*1.46*	1.40	1.52	*-.41*	-.50	-.31	*.91*	.79	1.02

Note: Effect size values over .2 are italicized to indicate significance according to Cohen's convention (1988).

Table 7

Demographic Composition: Percent of Ethnic Subgroups Admitted under Various Levels of Selectivity Using Composites of Cognitive or Both Cognitive and Noncognitive Measures

	Hispanic		Asian		Black		White	
Selectivity	**Cog**	**Cog+**	**Cog**	**Cog+**	**Cog**	**Cog+**	**Cog**	**Cog+**
High (Top 15%)	4.0	6.4	17.8	14.9	.9	4.1	77.0	74.6
Moderate (Top 50%)	4.3	4.6	15.0	11.0	8.3	10.0	77.1	75.3
Minimal (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	

Note. Cog represents the first composite, which included only the cognitive predictors. Cog+ represents the second composite, which included both the cognitive and noncognitive predictors.

Table 8

Means and Standard Deviations of the Clusters on Attitudinal, Behavioral and Performance Outcomes

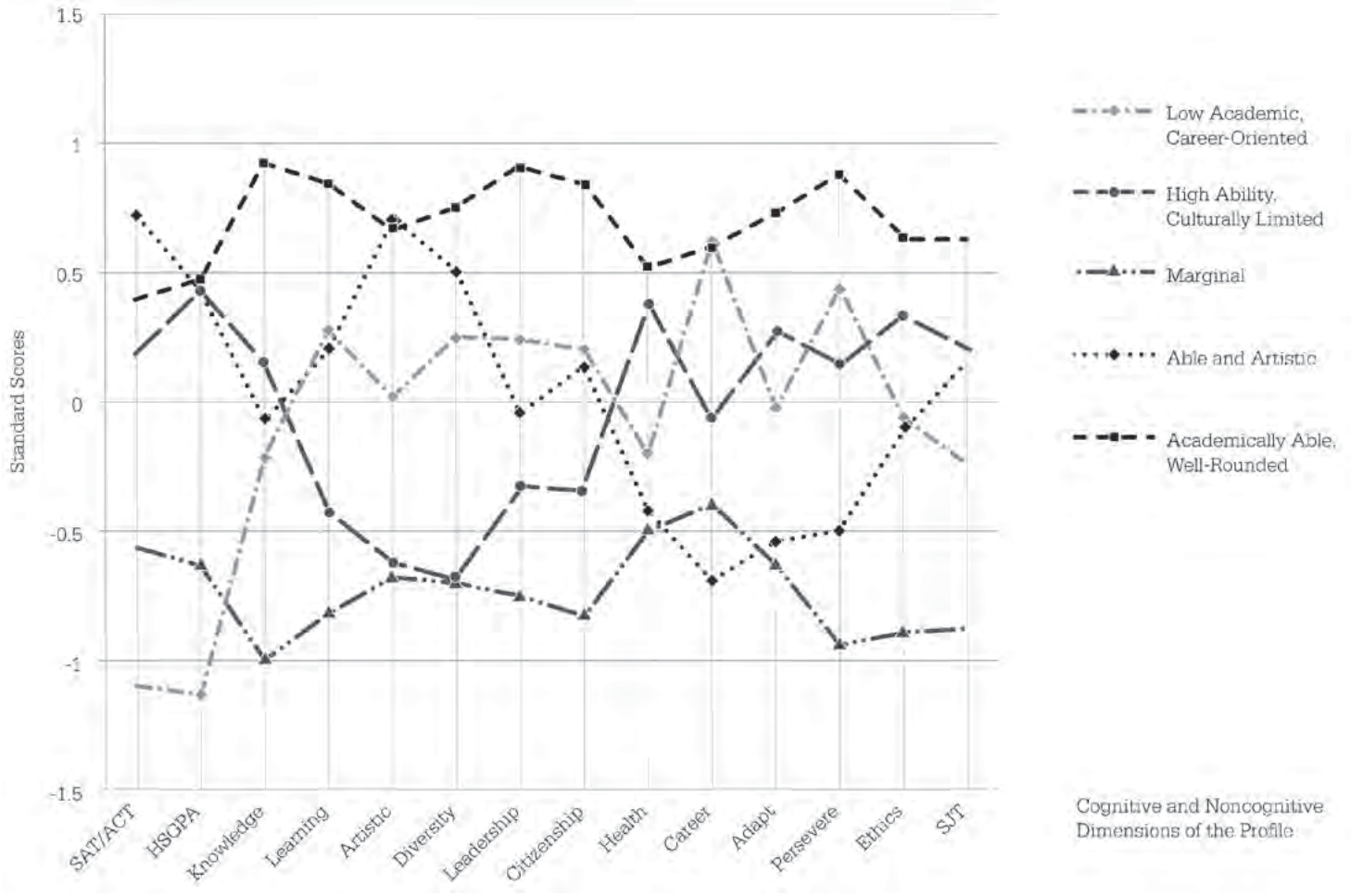
	First-Year College GPA			BARS Self-Rated Performance			Class Absenteeism		
Cluster	*Mean*	*SD*	*d*	*Mean*	*SD*	*d*	*Mean*	*SD*	*d*
Low Academic, Career-Oriented	2.22	.83	1.62	3.67	.57	.55	1.48	.75	.67
High Ability, Culturally Limited	3.24	.70	.28	3.51	.40	.91	1.69	1.03	.46
Marginal	2.50	.91	1.25	3.29	.45	1.40	2.16	1.13	---
Able and Artistic	3.45	.72	---	3.58	.41	.76	1.92	1.07	.24
Academically Able, Well Rounded	3.39	.70	.08	3.92	.49	---	1.60	.95	.55

	Intent to Quit			*OCBs*			*Satisfaction*		
Cluster	*Mean*	*SD*	*d*	*Mean*	*SD*	*d*	*Mean*	*SD*	*d*
Low Academic, Career-Oriented	1.58	.85	.01	2.64	.53	.69	3.89	1.08	.40
High Ability, Culturally Limited	1.39	.66	.30	3.04	.60	---	4.09	.73	.15
Marginal	1.59	.82	---	3.02	.65	.03	4.06	.79	.19
Able and Artistic	1.36	.58	.34	2.96	.58	.14	4.05	.75	.20
Academically Able,									
Well Rounded	1.35	.66	.36	2.67	.56	.64	4.21	.84	---

Note: *d* equals the standardized mean difference between the group with the highest average on an outcome and the average of the group to which it is compared. All overall mean differences were statistically significant at $p < .05$. Satisfaction refers to overall students' satisfaction with their university. Class Absenteeism refers to the number of classes students reported they did not attend for excusable reasons (e.g., health).

Figure 1.

Profile of cognitive and noncognitive dimensions for the five clusters of students.



Adapted from Schmitt et al. (2007).

CollegeBoard
inspiring minds™

10b-1555