Running Head: CRITERION-FOCUSED APPROACH

Criterion-Focused Approach to Reducing Adverse Impact in College Admissions

Ruchi Sinha, Frederick Oswald, Anna Imus, & Neal Schmitt

Michigan State University

Author Note

We acknowledge the financial support of the College Board, who sponsored this research.

Correspondence concerning this article should be addressed to Ruchi Sinha, Department of Psychology, Michigan State University, 348 Psychology Building, East Lansing, MI 48824-1116. E-mail: sinharuc@msu.edu.

Abstract

The current study examines how using a multidimensional battery of predictors (high school GPA, SAT/ACT and biodata), and weighting the predictors based on the different values institutions place on various student performance dimensions (college GPA, OCB and BARS), can increase the proportion of some ethnic subgroups often disadvantaged by the use of only traditional measures such as the SAT/ACT. The sample consisted of 836 students from 10 universities across the United States. Results show that meaningfully different proportions of groups would be admitted to universities when the predictors included noncognitive measures and the weights for the various components in the battery were based on performance dimensions other than first-year GPA. These dimensions should reflect institutional values.

Criterion-Based Approach to Reducing Adverse Impact in College Admissions

In the employment arena, a selection ratio for any sex, or racial/ethnic group which is less than four-fifths (4/5) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as prima facie evidence of adverse impact (Sec. 1607.3 D, Uniform Guidelines, 1978). Adverse impact is said to occur when a decision, practice, or policy has a disproportionately negative effect on a minority group. Adverse impact is defined as a substantially different rate of selection in hiring, promotion, or other employment decision which works to the disadvantage of members of a race, sex, or ethnic group. In the Uniform Guidelines on Employee Selection Procedures, a "substantially different" rate is typically defined using the $4/5^{th}$ Rule. The $4/5^{th}$ rule defines adverse impact as occurring when the selection rate of the minority group is less than 80% of the selection rate of the majority reference group. Although the Uniform Guidelines on Employment Selection has codified the $4/5^{th}$ rule, it can be considered a rule of thumb as the $4/5^{th}$ test does not involve probability distributions to determine whether the disparity is a “beyond chance” occurrence. Statistical significance tests of the differences in proportion hired are sometimes used to determine adverse impact in enforcement and litigation settings. All calculations using the $4/5^{th}$ rule constitute evidence of discrimination, not proof of discrimination. Although this rule has been codified in the Uniform Guidelines and is used by the U.S. Equal Employment Opportunity Commission (EEOC), Department of Labor, and Department of Justice in Title VII enforcement, it has not been applied to educational institutions.

While these Guidelines do not apply to educational institutions, similar concerns are common to educational admissions decisions. Understanding and finding solutions to reduce adverse impact in the college admission process can help to increase diversity in student bodies.

Educators in the U.S. higher education sector have long argued for the creation of racially and ethnically diverse student bodies. The underlying idea is to create the best possible educationally stimulating environment for all students, white and minorities alike. Until recently these arguments for diversity were not backed by strong research evidence however studies have now begun to provide both theoretical and empirical rationale to support the link between diversity and educational outcomes.

In a study by, Gurin, Dey, Hurtado, & Gurin (2002) they examined the effects of classroom diversity and informal interaction among members of different races such as, African American, Asian American, Latino(a), and White students on learning and democracy outcomes. In their multi-institutional data set, they found that the experiences students have with diversity meaningfully affect important educational outcomes. Diversity experiences at college explain an important amount of variance in learning outcomes such as, active thinking skills, intellectual engagement and motivation, and a variety of academic skills. They also influence other democracy outcomes such as, perspective-taking, citizenship engagement, racial and cultural understanding, and judgment of the compatibility among different groups. These positive effects of diversity were found to be quite reliable across the various outcomes, across the national and single institutional studies, and across the different groups of students.

Apart from empirical research that has been able to establish valid links between diversity experiences during the college years and a wide variety of educational outcomes (e.g., Astin, 1993a, 1993b; Chang, 1996), there are other sources that provide evidence for the benefits of diversity. Evidence for increasing diversity in student bodies at colleges and universities has come from three other sources: (a) Students: who have reported benefits from interacting with diverse peers (e.g., Orfield & Whitla, 1999); (b) Faculty: who have reported positive impact of diversity on student learning (e.g., Maruyama, Moreno, Gudeman, & Marin, 2000); and (c) Society: through both financial and nonfinancial gains for students and the larger community in terms of increase in income as a result of attending institutions which support diversity, better graduation rates for both whites and minorities and the attainment of advanced and professional degrees that prepare students to become leaders (e.g., Bowen & Bok, 1998; Bowen, Bok, & Burkhart, 1999; Komaromy et al., 1997). Although the educational sector is bound by no law related to adverse impact (like the employers are by the $4/5^{th}$ rule), addressing the issues related to adverse impact is of importance to educators as it can help to increase diversity in the student body and can promote desired educational outcomes, as well as avoid possible litigation based on discrimination (or reverse discrimination) in admissions.

The use of SAT or ACT scores in admissions will often mean that a lower proportion of some minority groups will be chosen, especially if the institution is highly selective. Moreover, despite this potential for adverse impact for institutions that rely on the SAT or ACT for making admissions decisions, it is widely recognized that these standardized tests are not psychometrically biased against lower scoring groups (Jensen, 1998; Schmidt, 1988). In fact, perhaps the first study of predictive fairness (Cleary, 1968) affirmed that use of cognitive tests (i.e., the Scholastic Aptitude Test) did not produce underprediction of the academic grades of minority college students. When institutions have sought to employ admissions strategies that reflected a direct concern for the admission of minority students, they have been challenged in the courts (e.g., Grutter v. Bollinger, 2002; Regents of the University of California v. Bakke, 1978). In addition, legislative proposals have been passed in three states (California, Michigan, and Washington) that prohibit any preferential selection in college admissions based on ethnic status or gender.

The Cleary (1968) study referred to above, employed an analysis of differential prediction that has been adapted as the preferred approach to the study of subgroup differences. This approach considers slope and intercept differences. Separate regression equations can be used to obtain standard errors for subgroups as well. This approach has been codified in various standards for test use (AERA, APA, and NCME, 1999; Society for Industrial and Organizational Psychology, 2003) and other publications (e.g., Young, 2001). The usual finding in educational and employment situations is that there is little evidence of slope differences (roughly equivalent to differences in subgroup validity coefficients) but that there is some evidence of intercept differences (a reflection that predicted mean differences in performance do not match predictor differences) usually indicating a slight over prediction of lower scoring minority group performance. In our study, the presence/absence of differential prediction will vary as a function of the outcome employed and is not central to the purpose of our paper. In our examples, the academic performance of minority students in hypothetical institutions that value outcomes other than college GPA will be over predicted. Interested readers may request subgroup statistics on study variables from the senior author. Traditionally, measures such as SAT/ACT scores and high school GPA have been commonly used to make college applicant selection decisions (Harackiewicz, Barron, Tauer, & Elliot, 2002; Oswald, Schmitt, Kim, Ramsay, & Gillespie, 2004). The national survey of undergraduate admission policies, practices, and procedures indicates that both four year public and private institutions report high school GPA (and grades in college preps, strength of curriculum) and admission test scores as their top most important factors in college admissions (Breland, Maxey, Gernand, Cumming, & Trapani, 2002, Hawkins & Clinedinst, 2008).

Standardized cognitive ability tests like the SAT in general have shown relatively high criterion–related validities (uncorrected > .30) with college grade point average (Hezlett et al., 2001; Oswald et al., 2004). However, it is also clear that there are sizable subgroup differences on these measures wherein African Americans score about one standard deviation below Whites, Hispanic Americans score about two-thirds standard deviations below Whites, and Asian Americans typically score higher on such tests of cognitive ability (e.g., Hunter & Hunter, 1984; Loehlin, Lindzey, & Spuhler, 1975; Schmidt, Greenthal, Hunter, Berner, & Seaton, 1977). These differences lead to substantial adverse impact on the lower scoring groups when cognitive ability tests are used to make selection decisions.

In the past several decades, much research and thought has gone into developing selection/admissions systems that can reduce adverse impact and maintain or increase their validity. These studies have affirmed the validity of these procedures in the prediction of first-year college GPA. They have also demonstrated smaller but statistically significant relationships with degree attainment, persistence and study habits (Hezlett et al., 2001; Kuncel & Hezlett, 2007). The various approaches to the reduction of adverse impact have been summarized by Sackett, Schmitt, Ellingson, and Kabin (2001). Their conclusion was that some of these approaches may reduce the impact of cognitive ability measures slightly, but that substantial subgroup differences would remain even when these procedures are employed. In this paper, we focus on the use of a criterion-based approach to the selection and use of predictor measures in college admissions. We first review the usual approach to admissions decisions, then describe the criterion-related approach, and then provide an example of such an approach in different scenarios.

Predictor-Based Approach to Reducing Adverse Impact

As mentioned above, a vast majority of research in the area of reducing adverse impact has been focused on understanding and refining tests and predictors (Sackett et al., 2001). Some of these studies have attempted to reduce adverse impact by using noncognitive measures to supplement tests of cognitive ability (De Corte, 1999; Potosky, Bobko, & Roth, 2005; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). A concern about low selection rates for minority students has led to an increased interest in the inclusion of standardized noncognitive predictors in admission decisions (Thomas, Kuncel, & Crede, 2004). There are many educational researchers that argue that the measurement of noncognitive variables and their use as predictors is needed to effectively predict college student success (Duran, 1986; Pantages & Creedon, 1978; Sedlacek, 2004). The national survey of undergraduate admission policies, practices, and procedures also indicates that measures of personal qualities of the applicants through the use of letters of recommendation, essays, interviews, portfolios, and the like are becoming more important (Breland, Maxey, Gernand, Cumming, & Trapani, 2002). Bial and Rodriguez (2007) emphasize the need to go beyond traditional admission measures in order to identify diverse students who might otherwise not be selected.

A large number of studies have examined the effects of different regression-weighted predictor composites on adverse impact and the validity of selection decisions. Bridgeman, Burton, and Cline (2003) demonstrated the impact on the composition of the admitted class when SAT I was substituted with SAT II subject tests. Such a predictor-focused approach is based on the assumption that combining predictors that have less or no adverse impact with tests of cognitive ability can lead to some reduction in overall adverse impact. These studies have generally considered the prediction of a single criterion construct, in educational admissions, usually first-year GPA. Hattrup and Rock (2002) call this approach the *predictor-based* *approach*. Such an approach emphasizes the choices and weights about predictors based on their correlation with an outcome measure. In cases in which multiple predictors are considered, predictors are chosen based on the magnitude of their regression weight which is based on the regression of GPA on the set of predictors. This approach does not consider how the predictor composite will be related to student performance profiles that include outcomes other than college GPA, usually first-year GPA.

The predictor-based approach can be criticized as it often conceptualizes performance as a unidimensional construct. More specifically, choices about the inclusion and weighting of predictors is done based on their relationship with a single criterion with no consideration as to whether and how the predictor composite might affect performance on different dimensions of performance (Hattrup & Rock, 2002). A strategy that includes the consideration of alternate predictors and weights them based on their effects on adverse impact may result in reduction in adverse impact but may also reduce predicted performance on the commonly used GPA criterion. It also generates the notion that some form of reverse discrimination has occurred and generates the legal cases and legislative actions referred to above. Finally, results from several studies have shown that the reduction in adverse impact from the use of predictor composites that include measures on which little or no subgroup differences are observed is not as substantial as one would think (Bobko, Roth, & Potosky, 1999; Sackett & Wilk, 1994). It is a mathematical truism that the adverse impact from a cognitive predictor cannot be reduced appreciably by adding many predictors with zero mean difference between majority and minority subgroups, in fact adding measures with reduced adverse impact can at times actually *increase* the overall adverse impact (Sackett & Ellingson, 1997).

Criterion-Focused Approach to Predictor Selection

Hough, Oswald, and Ployhart (2001) emphasize that practically speaking, predictors derive their importance from their ability to predict meaningful criteria. Understanding the criterion of interest should thus drive and determine the appropriate content of the predictor battery. A vast number of studies have defined college student success primarily in terms of college GPA, and of those studies, the majority focus only on first-year college GPA (Hughes & Douzenis, 1986; Kanoy, Wester, & Latta, 1989; Mouw & Khanna, 1993; Pettijohn, 1995; Ting & Robinson, 1998; Young & Sowa, 1992). However, a few notable studies have begun to investigate longer-term student success. For instance, Boyer and Sedlacek (1988) examined how the Non-Cognitive Questionnaire (the NCQ) predicted the GPA of international students over the course of two years. They found that self-confidence and the availability of a strong support system predicted GPA. Harackiewicz et al. (2002) found that achievement goals and ability predicted early success in college and over the longer term. Although they broadened the set of predictor variables considered, these two studies still focused on the usual outcome of first-year college GPA. Oswald et al. (2004) claim that when predicting college GPA, standardized tests and high school grade point average are the most effective predictors. However, they do suggest that motivational and background characteristics would be stronger predictors of college outcomes on which empirical research does not focus traditionally (e.g., leadership, community involvement, perseverance). Their argument is, of course, based on the assumption that these alternate criteria are relevant and important outcomes of college education, and many college mission statements do in fact make that claim.

Expanding the Criterion Domain for College Student Performance

Paralleling the development of multidimensional conceptualizations of job performance (e.g., Borman & Motowidlo, 1997; Campbell, McCloy, Oppler, & Sager, 1993) is the need to understand and expand the criterion domain of college student performance into a multidimensional model. SAT and ACT scores are predictive of first-year college GPA, but this is a narrowly defined conceptualization of college student performance. Broadening the criterion domain of college success potentially does two important things: First, it measures college success in a manner that is more comprehensive and relevant to university stakeholders, including university administrators, legislators, parents and the students themselves. Second, it justifies the inclusion of a broader range of predictor measures that are noncognitive in nature that could reduce adverse impact by improving the selection ratio for some of the minority subgroups, such as African Americans and Hispanics. Admissions personnel are now realizing that there are advantages associated with expanding their definition of merit (student success) and developing broader admission strategies in order to succeed in recruiting a more diverse student body (Gratz v. Bollinger, 2003).

Taber and Hackman (1976) were among the first researchers who identified multiple dimensions of student success. They identified 17 academic and nonacademic dimensions which could classify college students as successful and unsuccessful. Some of these dimensions were intellectual perspective, curiosity, communication proficiency, and ethical behavior. Studies have shown that those college students who excel in these behavioral domains tend to achieve greater success in their overall college experience as reflected in their scholastic involvement, accumulated achievement record, and their graduation (Astin, 1984; Willingham, 1985).

Oswald et al. (2004) identified 12 dimensions of college student performance that overlap with Taber and Hackman’s (1976). Their approach to expanding the domain of college student performance was to examine converging themes from college mission statements, the educational research literature, interviews with university staff, and from other sources concerning the goals of higher education. These 12 dimensions, as indicated in Table 1, deal with intellectual behaviors (Knowledge, Learning, and Artistic), interpersonal behaviors (Multicultural, Leadership, Interpersonal, and Citizenship), and intrapersonal behaviors (Health, Career, Adaptability, Perseverance, and Ethics).

Insert Table 1 about here

A criterion-focused approach to selecting predictors emphasizes the importance of the multidimensional nature of performance and the role of organizational/institutional values in determining the relative importance of the various types of performance relevant behaviors (Hattrup & Rock, 2002; Hattrup, Rock, & Scalia, 1997). Depending on the nature of the relevant criteria, use of this approach may reduce adverse impact. Hattrup and Rock also compared the predictor-based and criterion-based approach to reducing adverse impact and found that predicted performance is sacrificed when the weights used in forming predictor composites are incongruent with the values placed on the criterion dimensions.

The Present Study

The aim of the present study is to see how adverse impact is affected if we use a criterion-based approach, with predictor composites comprised of predictors that are differentially weighted based on the value associated with corresponding criterion dimensions. Weighted criterion composites reflecting university outcomes in addition to GPA were regressed on a set of cognitive and noncognitive measures to determine the weights of those predictors when making admissions decisions. If one uses predictor weighting strategies and both cognitive and noncognitive predictors with the sole intention of reducing adverse impact, one may slightly reduce adverse impact but also lower the predicted student performance on various criterion dimensions (Hattrup & Rock, 2002). Thus, it is very important to first identify the relevant criterion dimensions; second, to weigh them according to their importance and finally use this composite multidimensional criterion to determine the regression-weights for the predictor battery.

In the present study, our criterion composite combines an expanded criterion domain of college student performance, where we include not only college GPA but also measures of organizational citizenship behaviors (OCBs) and a self-rated behaviorally anchored rating scale (BARS) that is a unit-weighted sum across the 12 performance dimensions listed in Table 1. OCBs reflect employee behaviors that help the organization; they have been the focus of a great deal of organizational research in the last ten years (Motowidlo, 2003; Organ, 1997), and they conceptually and operationally translate readily into college-student behaviors (e.g., participating in student government, leading clubs and extracurricular activities, volunteering in the college community).

To present empirical results that differentially weigh the GPA, OCB and BARS criterion dimensions, we simulated a range of universities that differ in the value they place on the different criterion dimensions. The relative value placed on the different dimensions of college student performance was then used as the means of computing an appropriate criterion composite. This composite was then regressed on high school GPA, ACT/SAT scores, and a biodata measure to determine the regression weights associated with these predictors. Adverse impact implications were examined in light of the proportion of members admitted from different subgroups using these regression weights.

In order to predict such a multidimensional performance criterion we have developed a composite of broader predictors that include both cognitive and noncognitive measures. We have thus expanded the predictor battery to include not only tests of cognitive ability (SAT/ACT) and achievement scores (high school GPA) but also included a biographical measure (biodata). The biodata measure included items addressing a variety of interests, background experiences, and motivational characteristics of the students. The biodata measure was developed to measure the 12 dimensions listed in Table 1. The study described in this paper represents an example of how this criterion-focused approach to selecting and weighting different admissions tests or information can affect the proportion of members of different subgroups admitted to a university class.

Method

Sample

The sample consisted of 836 freshman students from 10 universities across the United States. We deliberately sampled from participating universities that were diverse in terms of region of the country; one was from the Southwest, two were historically Black colleges in the Southeast, five were Big Ten Midwestern universities, one was from the Southeast, and one was a highly selective private mid-western school. These institutions included Winston-Salem State University, California State University-Fullerton, Indiana University, Michigan State University, Ohio State University, Spellman College, University of Iowa, University of Chicago, University of Michigan, and Virginia Polytechnic Institute and State University. The institutions differed in terms of their selectivity ranging from highly selective to minimally selective. The average age of our participants was just over 18 years; in fact 99% of our sample was either 18 or 19 years of age. Sixty-five percent of the sample was female, 97% were U.S. citizens, and 96% indicated that English was their native language. Regarding the ethnicity breakdown, this sample was 75% Caucasian, 9% African American, 6% Hispanic and 10% Asian. All students provided responses to our paper-and-pencil measures in the first few days or weeks of their college career (e.g., around freshman orientation) by participating in group sessions supervised by admissions officers or other staff members at the university. Detailed instructions were provided to students and staff. Responses were recorded on machine-scorable answer sheets and were mailed to the researchers. These staff members were paid for their help, and students were paid ($40 per student) for their participation in the initial survey that included the predictor measures used in the study. The complete data collection effort at each college or university took approximately two hours.

We collected criterion measures (i.e., OCB and BARS) at the end of the students' first academic year via a web-based survey of all student participants in the original survey. Students were recruited via e-mail, and participation was voluntary. Each student was sent the original request and up to two reminders. Students who returned the survey were awarded a $20 gift certificate from Amazon.com. College GPA was obtained from the registrar's office at the participating universities with the students' permission.

Measures

Ability (SAT/ACT), high school GPA and the biodata measure were used as predictor measures.

Ability. Information about participants' SAT scores and ACT scores were collected from the admissions offices at the participants' home institutions. We then converted all of the participants' admission test scores to a single scale by converting ACT (composite) test scores to equivalent SAT scores using a conversion table from www.collegeboard.com (Dorans, Lyu, Pommerich, & Houston, 1997). The converted ACT and SAT scores in this study were correlated .90 (N = 367). When a participant took both the ACT and SAT, the converted ACT score(s) was averaged with the raw SAT score(s).

High school GPA. Information about participant's high school GPA (HSGPA) was collected from the admissions offices at the participants' home institutions. Ninety-one percent of the reported high school GPA's were on a scale of 0.00 to 4.00. No high school GPAs were provided on a scale of 0-100, as occasionally may be found. A very small percentage of the respondents had reported high-school GPAs above 4.00. The reported high school GPA's above 4.00 were converted to 4.00 in our dataset.

Biodata. Biographical data (biodata), reflecting information about an individual's background and life history, were collected in the initial data collection phase. Some of the information collected in the biodata instrument is contained within college applications, but it is often provided by students in an essay or short answer format and tends to be used by admissions officers in an intuitive or implicit manner (e.g., in interpreting the extracurricular activity lists and resumes that applicants provide). By contrast, we undertook the development of a biodata inventory with directly quantifiable answers, which is a more systematic and standardized way to obtain similar information, and which would therefore allow for more explicit and consistent methods for admissions officers to incorporate this information in making college-admissions decisions or in providing guidance with respect to major or course choices. A description of the development and some initial validation results for this measure is contained in Oswald et al. (2004).

The biodata inventory contained standard multiple-choice questions about a student's previous experiences (e.g., number of student organizations in which the student took a leadership role). Participants completed 126 biodata items reflecting 11 of the 12 dimensions of college student success proposed by Oswald et al. (2004). Because of low alpha reliability, the Interpersonal Skills scale was not used. The psychometric quality of these biodata scales is reported elsewhere (Oswald et al.; Schmitt et al., 2007). In the present study we used a composite score for the biodata measure. The biodata composite score can be interpreted as an overall life and academic experiences construct. The biodata composite score summarizes an applicant's life experiences on the dimensions listed in Table 1.

Several criteria reflecting a multidimensional view of student performance were collected.

First-year college GPA. First-year college GPA was collected from all participants' respective institutions in the summer of 2005 for those students who completed the first year of college. Because admissions policies differed across the schools in our sample, admitted students varied in their average SAT/ACT score, which also differed across schools; thus, we corrected first-year college GPA using a procedure that College Board employs in assessing the validity of the SAT in similar instances. Specifically, we first standardized the GPA variable within university. We then regressed the standardized grades across universities on the ability measure (i.e., the SAT/ACT composite) along with a set of nine dummy variables representing the 10 colleges and universities. The coefficients for the dummy variables indicate the differences in grades that would be expected for students with comparable SAT scores at the various universities. Grades for students at each school were then adjusted by that school's regression coefficient such that students at universities with higher average SAT scores received a relatively higher adjusted college GPA, and conversely, students at universities with lower average SAT scores received a relatively lower adjusted college GPA. Thus the between-school differences in students' SAT/ACT scores were removed when examining relationships with college GPA

BARS-Behaviorally anchored rating scale for multiple dimensions of college performance. Students' self-reported performance on 12 dimensions of college student success was measured using BARS. The 12 dimensions (mentioned in Table 1) served as a guide in developing a behaviorally anchored rating scale (BARS). For each of the 12 BARS items, a dimension name and its definition were presented along with two examples of college-related critical incidents and various behavioral anchors that reflected three levels of performance on a seven-point scale, which ranged from *unsatisfactory* to *exceptional*.

We conducted a principal-axis exploratory factor analysis (EFA) that yielded a large first factor that accounted for 32% of the variance and four times as much variance as the second factor (e.g., see Oswald et al., 2004). A confirmatory factor analysis (CFA) of these BARS ratings using LISREL 8.51 (Jöreskog & Sörbom, 2001) yielded support for a single-factor model, $\chi^2 = 122.71$, $p < .01$ (54, N = 641); root-mean-square error of approximation (RMSEA = .05); comparative fit index (CFI = .95); and nonnormed fit index (NNFI = .93). Thus, we decided to create the composite as an overall and broad measure of college student performance, using the mean of the 12 BARS items. Empirically, the BARS composite score appears to measure a large general performance construct that represents a student's performance across several domains. Students with high scores on this composite rated themselves highly across the BARS-rated dimensions, and conversely students with low scores rated themselves low across BARS dimensions. The alpha coefficient associated with this self rating was .74.

Organizational citizenship behaviors (OCB). OCBs are defined as behaviors that are discretionary, not directly or explicitly recognized by the formal reward system, and that in the aggregate promote the effective functioning of the organization/institution. In the college context, OCBs refer to behaviors that are not critical to student's degree attainment but are important in promoting the welfare of the student's institution (Organ, 1997). To assess student citizenship behavior, we adapted the measure created by Moorman and Blakely (1995). We adapted this measure by selecting content from three of the five subscales included in the original instrument and altering the items such that they reflect an academic, rather than organizational, setting. The three subcategories of citizenship behaviors that we considered relevant to university settings were interpersonal helping, loyalty, and individual initiative. This measure consisted of a series of 15 five-point Likert-type scales with responses varying from "strongly disagree" to "strongly agree." Example items included "Gone out of your way to make new students feel welcome at school," "Defended your school when other students tried to criticize it," and "Participated in student government or other clubs that try to make your school a better place." Alpha for this scale in the present study was .85.

Simulated Schools

The purpose of this study was to demonstrate how differential weighting of criterion dimensions influences predictor regression weights and eventually influences the ethnic composition of the student body. For this purpose we simulated 10 schools that differentially valued (weighted) the three criterion dimensions of college student success. The different weighting schemes for the three criterion dimensions (College GPA, OCB and BARS) based on school type is shown in Table 2. The weights were chosen to operationalize a wide range of possible value systems that universities might place on this set of criteria. For example, School A was simulated to be a school that primarily values academic achievement and thus was assigned a weight of 1.00 for college GPA and zero for OCB and BARS. On the contrary we simulated another school, which was school G where GPA, OCB and BARS were given equal weights. The aim of the present study is to demonstrate the effects of a criterion-based approach on the adverse impact ratios for minority groups. The simulated schools present only a subset of all possible weighting schemes and are meant to be exemplars of possible weighting schemes. Some of these schools may strike some readers as unrealistic, but our purpose is to provide a range of scenarios for purposes of illustration. We also think even those that seem like "extreme weighting schemes" are often employed, particularly when universities seek athletic excellence, artistic excellence, etc., as well as academic ability.

Insert Table 2 about here

Criterion Weighting Strategy

In forming a weighted criterion composite, it is important to distinguish between nominal weights and effective weights. *Nominal weights* are the multiplicative weights applied directly to each variable being combined into a linear composite. For instance, with three criterion variables, one could apply weights of 4, 2 and 1, multiplying scores for the first variable by four, and scores for the second variable by two (multiplying the third score by one is unnecessary), and then adding them together. Alternatively, one could apply unit weights, meaning simply adding scores for each variable together without weighting, assuming such component variable has the same standard deviation.

A problem is that nominal weights, as just described, do not always translate into the effective weights that are desired. *Effective weights* reflect the relative contribution each variable makes to the variance of the linearly weighted composite score. They are not the same as the nominal weights. When variables are standardized and uncorrelated, then effective weights are proportional to the *square* of the nominal weights. For instance, with nominal weights of 4, 2 and 1, the effective weights are 16/21, 4/21, and 1/21 (i.e., the square of the nominal weights is in the numerator, and the sum of the squared weights, 16 + 4 + 1 = 21, is in the denominator). These effective weights are proportions that reflect the contribution of each variable's scores to the composite variance, and as proportions they sum to one. Nominal and effective weights are equal only in one case, where the variables to be combined are standardized and uncorrelated, and the weights are unit weights; otherwise nominal and effective weights always differ.

As seen from the example just provided, effective weights are straightforward to compute when variables are standardized and uncorrelated. However, when variables are not standardized and correlated, then computing effective weights is not as straightforward (see Guion, 1998, pp. 346-348). When each variable is unstandardized, then each variables' standard score is essentially weighted by its standard deviation, such that a variable with a larger SD contributes more to the composite score variance. When variables are correlated (regardless of whether they are standardized), then the contribution of the covariances to the composite cannot be ignored.

In the present study we know that our three different criterion variables were correlated (see Table 3). Therefore we did not simply take proportions of the squared nominal weights as a reflection of the effective weights. Rather than applying nominal weights to each criterion variable and then determining the effective weights, we approached the problem in reverse: We first decided on the desired effective weights—or the proportions we wanted each variable to contribute to the composite variance (i.e., based on the value placed on the three criterion dimensions mentioned in Table 2). Then we computed the nominal weights needed to result in the desired effective weights, with the caveat that the set of nominal weights had to be nonnegative (see Adams & Wilmut, 1981). This approach was accomplished by using an optimization program that can solve a system of equations with nonlinear constraints (nonlinear because variances have a quadratic form). Many such programs are available for this relatively simple problem; we solved this problem in Mathematica (Wolfram, 1999).

Insert Table 3 about here

Results

Intercorrelations between the predictor variables and the criterion dimensions are presented in Table 3. As can been seen from Table 3, SAT/ACT (ability) and high school GPA (HSGPA) are more strongly correlated with college GPA than with organizational citizenship behavior (OCB) and BARS. Similarly biodata is more strongly correlated with OCB and BARS than it is with college GPA. Our study is not predicated on a hypothesized pattern of predictor-criterion relationships, but most of these relationships are in line with what we might have expected. Although ability and HSGPA were negatively correlated with OCB, the correlation was weak and almost negligible. In the college context, OCBs refer to behaviors that are not critical to student's degree attainment or academic performance but are important in promoting the welfare of the student's institution (Organ, 1997). There is no theoretical basis for us to expect that OCBs will be positively related to achievement test scores or measures of academic achievement like HSGPA and college GPA.

Given the effective weights for GPA, OCB and BARS criteria are applied, and criteria are combined into a composite, Table 2 contains the regression weights for the predictors based on each school type. Regression weights change in the expected direction; for instance, when college GPA is given more weight in computing the criterion composite, we get higher regression weights for the ability and HSGPA predictors. Similarly, when a school places a greater value on OCB and BARS, the regression weights for biodata as a predictor increase. These changes are congruent with the pattern of convergent and discriminant relationships evident in the correlation table (see Table 3), where there is a stronger relationship between college GPA, HSGPA, and SAT/ACT measures (vs. OCB with SAT/ACT and HSGPA) and there is a strong relationship between OCB, BARS and biodata (vs. college GPA with biodata).

Table 4 contains results of these different weighting strategies for the admission of members of different subgroups. The table indicates the percentage of minorities selected under each scenario, using the regression weights in Table 2 and selection ratios of .25, .50 and .75 (i.e., ranging from being relatively selective to being relatively inclusive). Adverse impact ratios (AI) for the different schools were also calculated across the three different selection ratios. AI values were calculated by dividing the selection rate for the minority group by the selection rate for the majority group.

Insert Table 4 about here

School A was simulated to be a school that values college GPA as the sole criterion for college student success. In such a school we can see that there is adverse impact using the four-fifths rule for African Americans at all selection ratios (.25, .50 and .70). There is adverse impact for Hispanics at the 25% and 50% selection ratio. There was no adverse impact for Asians in such a school. In fact, Asians would be selected at a rate substantially larger than that of White applicants reflecting the fact that Asians scored higher than Whites on the ACT/SAT and high school GPA predictors. By contrast, School B was simulated to be a school that gives 80% weight to college GPA, 10% to Organizational Citizenship Behaviors and 10% to BARS. In such a school we can see that there was no adverse impact for Hispanics at any selection ratio. Using the $4/5^{th}$ rule there is indication of adverse impact for African Americans in School B, but there is an increase in percentage of blacks selected compared to School A.

The results in Table 4 demonstrate the impact of expanding the criterion domain and using weighted predictor composites which include both cognitive and noncognitive measures on the admission of minority students. Results for the remainder of the hypothetical schools reflect an intermediate impact on the admission of members of various subgroups. This is expected as the weights for various criterion dimensions reflect a balance between GPA, OCB, and BARS. Note that Whites are not considered an 'affected group'; however, across many of the scenarios just described, they are disadvantaged relative to certain minority groups by the four-fifths rule, in cases where AI ratios are greater than 1/.80 = 1.25.

Discussion

The main purpose of this research was to demonstrate the use of a criterion-focused approach to the selection and use of predictor measures in college admissions and how it can reduce adverse impact. A vast majority of research in the area of adverse impact uses a predictor-focused approach which does not incorporate the criteria to be predicted and only assumes that combining predictors that have no adverse impact with tests of cognitive ability can reduce overall adverse impact. The present study moves beyond this approach by first broadening both the predictor and criterion domain with a range of relevant predictor and outcome variables that are both cognitive and noncognitive in nature. Doing so serves a dual purpose: to represent and predict performance more broadly *and by doing so*, to reduce adverse impact by including criteria and predictors that are less cognitive in nature. Rarely does the literature on multidimensional performance and on adverse impact make this connection explicit.

In our simulations, we obtained widely varying college-admissions results from a variety of schools that reflect very different values placed on different student performance dimensions. Most of the previous studies that have taken a criterion-focused approach to reducing adverse impact have focused on a single criterion and have relied on Monte Carlo simulation data. The present study is different in that it is based on actual college-relevant predictor and criterion data that have been collected over a period of one year; it also has focused on the challenge of combining multiple criteria into a composite that reflects the values that an institution places on the performance of the student body it seeks to admit.

Applications of our Findings

We have demonstrated how a criterion-focused approach can reduce adverse impact. This approach assumes the importance of the multidimensional nature of college student performance and the role of organizational/institutional values in determining the relative importance of the different types of performance relevant behaviors (Hattrup & Rock, 2002; Hattrup et al., 1997). In terms of the applied nature of our findings, the main learning point for academic and educational institutions is that adverse impact in the college admission selection decision can be reduced if colleges/universities use a battery of cognitive and non-cognitive predictors that are weighted according to the values institutional stakeholders place on an expanded performance criterion of student success. The results from this study illustrate that by expanding the predictor and criterion domain one cannot only enhance prediction of college student outcomes based on the values of the institution but that one can also reduce adverse impact by increasing the selection rate for some of the minority subgroups, such as African Americans and Hispanics. Various admission models (e.g., eligibility-based models, performance based models, student capacity to benefit model etc.) have been identified in the educational literature that reflect the philosophical perspective (values) of the educational institutions (Rigol, 2002). These admission models provide the context in which admission decisions are made. These models play a critical role in the definition of success and the identification of relevant predictors of success (admission criteria). Although a detailed discussion of which admission model is most conducive for reducing adverse impact is beyond the scope of this study, we recommend that future researchers and practitioners who seek to apply this approach must review and apply the illustrated strategy consistent with the stated mission and admission model.

The mission statements of many colleges and universities reflect student performance dimensions in terms of goals relating to intellectual behaviors (Knowledge, Learning, and Artistic), interpersonal behaviors (Multicultural, Leadership, Interpersonal, and Citizenship), and intrapersonal behaviors (Health, Career, Adaptability, Perseverance, and Ethics; see Oswald et al., 2004). Many times colleges and universities value these goals and outcomes in students but fail to consider them when selecting students. The results from this study support the idea that colleges/universities need to first explicitly decide what they value in terms of student success and then form the weighted performance criterion composite. These decisions regarding the weighting of different criterion dimensions of student performance determine the predictor weights which in turn affects the admission rates for different subgroups. This suggested approach of designing the admission process is also beneficial for accreditation purposes. There are several standards for excellence that have been laid down by accrediting agencies (e.g., The Middle States Commission on Higher Education—a recognized agency that is part of the

Council for Higher Education Accreditation). One such standard expects the college admission process (recruitment, admission criteria and selection) to be congruent with the institutional mission statements and goals. Another standard expects the assessment of student learning and performance to be a reflection of the institute's educational and mission goals (The Middle States Commission on Higher Education Report, 2009). The College Board report on the best practices in admission decisions (Rigol, 2002) also suggests that the criteria for success in college must be developed in the context of the mission of the institutions.

Limitations

Some limitations of the research should be noted. First, we simulated only a limited range of schools with different weighting scenarios. There could be other schools that might place different values on the student performance dimensions. The range of simulated schools was selected to illustrate some of the different weighting schemes and how they affect subgroup admission rates at different selection ratios. Practitioners who wish to apply the criterion-focused approach to reducing adverse impact must determine the precise set of weights for various student outcomes based on what their institution values and considers important for successful performance. Justifying any particular set of weights is not an easy task; for instance, there are likely to be limited data informing the issue; there are multiple stakeholders in the college admissions process; and there may be a host of different-yet reasonable perspectives on how different predictor and criterion constructs should even be defined and measured (regarding the latter, the present authors faced such a challenge in their college admissions research).

Second, the data used for this illustration are unique to the set of students and schools in our study. Although the between-school differences in students' SAT/ACT scores were removed when examining relationships with college GPA's, we were unable to correct for the differences in GPA's across disciplines. While the overall sample was relatively large, some subgroup sizes were much smaller. Generalizability of the predictor-criterion relationships to any other set of students/schools is certainly an important concern, though the observed relationships on the whole were consistent with the broader literature on student admissions (Hezlett et al., 2001).

Third, self-ratings of nonability measures certainly have a tendency to be inflated in high-stakes admissions situations; in the present study data, the sample was not high-stakes (students were already admitted into their universities) and there was considerable reliable variance in the biodata scores (i.e., they are not all high) leaving the potential for predictive relationships, as we have demonstrated in previous studies (e.g., Oswald et al., 2004). This is no guarantee however, that such relationships may be found in data where student-applicants are highly motivated to be admitted into a given college or that relationships between biodata and OCBs/BARS were not inflated due to a common method bias. However, these two sets of data were collected at time points separated by more than a year. The time lag in the measurement of predictors and criterion not only attenuates the effects of halo but also of normative implicit theories students may have about predictor-criterion relationships. Note that we would expect other colleges predictive validity data to have the same lag between predictor measures at the point of admission and subsequent performance measures taken during the span of students' college careers. We also suggest that future research and practical implementation of these measures utilize appropriate checks. For example, inflation in biodata can be checked by asking respondents to report names of individuals who can verify information reported on biodata forms as well as by providing a warning regarding lie detection (e.g., Schmitt & Kunce, 2006; Schrader & Osburn, 1977). Future measures of citizenship behavior must use both self-report as well as objective measures as they both have the capacity to capture different and important aspects of discretionary behaviors.

Finally, we would like to reiterate that the primary purpose in this paper was to present the general case for differentially weighting criteria to judge its effects on applicant diversity. As such, the observed relationships are in one sense illustrative; we recognize and acknowledge the limitations of our design, but they will not materially change the nature of this illustration. For any specific case or application, we certainly would recommend that other institutions rely on their own data (subjective & objective measures) to assess these research outcomes.

Fourth, this study is not intended to provide a definitive solution to the problem of adverse impact in college admissions. Nor does this study advocate any particular weighting scheme for the criterion dimensions. Results from the present study are a mere demonstration of how values placed on different dimensions of college student performance can result in differential weighting of predictors which in turn can influence the demographic profile of students selected.

References

Adams, R. M., & Wilmut, J. (1981). A measure of the weights of examination components, and scaling to adjust them. *The Statistician, 30*, 263-269.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC. American Psychological Association.

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

Astin, A. W. (1993a). *What matters in college? Four critical years revisited*. San Francisco: Jossey-Bass.

Astin, A. W. (1993b). Diversity and multiculturalism on the campus: How are students affected? *Change*, *25*, 44-49.

Bial, D., & Rodriguez, A. (2007). Identifying a diverse student body: Selective college admissions and alternative approaches. *New Directions for Student Services, 118*, 17-30.

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

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

Bowen, W. G., & Bok, D. (1998). *The shape of the river: Long-term consequences of considering race in college and university admissions*. Princeton, NJ: Princeton University Press.

Bowen, W. G., Bok, D., & Burkhart, G. (1999). A report card on diversity: Lessons for business from higher education. *Harvard Business Review, 77*, 139-151.

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

Breland, H. R., Maxey, J., Gernand, R., Cumming, T., & Trapani, C. (2002). *Trends in College Admission 2000: A report of a national survey of undergraduate admission policies, practices, and procedures.* ACT, Inc., Association for Institutional Research, The College Board, Educational Testing Service, and National Association for College Admission Counseling.

Bridgeman, B., Burton, N., & Cline, F. (2003). Substituting SAT II: Subject Tests for SAT I: Reasoning Tests: Impact on admitted class composition and quality. *Research in Higher Education, 44*(1), 83-98.

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

Chang, M. J. (1996). *Racial diversity in higher education: Does a racially mixed student population affect educational outcomes?* Ph.D. dissertation. Los Angeles: University of California.

Characteristics of Excellence in Higher Education. (2009). The Middle States Commission on Higher Education Report. Philadelphia, PA, Middle States Standards.

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

De Corte, W. (1999). Weighting job performance predictors to both maximize the quality of the selected workforce and control the level of adverse impact. *Journal of Applied Psycholog*y, *84*, 695-702.

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

Duran, R. P. (1986). Prediction of Hispanics' college achievement. In M. A. Olivas (Ed.), *Latino college students* (pp. 221-245). New York: Teachers College Press.

Gratz v. Bollinger, 2003 – The Oyez Project, 539 U.S. 244 (2003), available at: <http://www.oyez.org/cases/2000-2009/2002/2002_02_516/>.

Grutter v. Bollinger, 288 F.3d 732, 737 (6th Cir. 2002).

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

Gurin, P., Dey, E. L., Hurtado, S., & Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. *Harvard Educational Review, 72*, 330-366.

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

Hattrup, K., & Rock, J. (2002). A comparison of predictor-based and criterion-based methods for weighting predictors to reduce adverse impact. *Applied H.R.M Research*, *7*, 22-38.

Hattrup, K., Rock, J., & Scalia, C. (1997). The effects of varying conceptualizations of job performance on adverse impact, minority hiring, and predictor performance. *Journal of Applied Psychology*, *82*, 656-664.

Hawkins, D. A., & Clinedinst, M. (2008). *State of College Admission 2008 Report*. The National Association for College Admission Counseling. NACAC.

Hezlett, S. A., Kuncel, N. R., Vey, M. A., Ahart, A. M., Ones, D. S., Campbell, J. P., & Camara, W. (2001). 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., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment, 9,* 152-194.

Hughes, T. M., & Douzenis, C. (1986). Predictor and performance variables in a performance-based education course. *Journal of Psychology, 120*, 143-147.

Hunter, J. E., & Hunter, R. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin, 96*, 72-98.

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

Jöreskog, K. & Sörbom, D. (2001). *LISREL 8.50*. Chicago, IL: Scientific Software.

Kanoy, K. W., Wester, J. L., & Latta, M. (1989). Predicting college success of freshmen using traditional cognitive, and psychological measures. *Journal of Research and Development in Education, 22*, 65-70.

Komaromy, M., Grumbach, K., Drake, M., Vranizan, K., Lurie, N., Keane, D., & Bindham, A. B. (1997). The role of Black and Hispanic physicians in providing health care for underserved pPopulations. *New England Journal of Medicine*, *334*(20), 1305-1310.

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

Loehlin, J. D., Lindzey, G., & Spuhler, J. M. (1975). *Race differences in intelligence*. San Francisco: Freeman.

Maruyama, G., Moreno, J. F., Gudeman, R. W., & Marin, P. (2000). *Does diversity make a difference? Three research studies on diversity in college classrooms.* Washington, DC: American Council on Education and American Association of University Professors.

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

Motowidlo, S. J. (2003). Job performance. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology,* (Vol. 12, pp. 39-53). Hoboken, NJ: Wiley.

Mouw, J. T., & Khanna, R. K. (1993). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal, 27*, 328-336.

Orfield, G., & Whitla, D. (1999). *Diversity and legal education: Student experiences in leading law schools*. Cambridge, MA: Harvard University, The Civil Rights Project.

Organ, D. W. (1997). Organizational citizenship behavior: It's construct cleanup time. *Human Performance, 10,* 85-97.

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

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

Pettijohn, T. F., II. (1995). Correlations among college students' grade point averages and American college test scores. *Psychological Reports, 76*, 336-338.

Potosky, D., Bobko, P., & Roth, P. (2005). Forming composites of cognitive ability and alternative measures to predict job performance and reduce adverse impact. *International Journal of Selection and Assessment, 13*(4), 304-315.

Regents of the University of California. v. Bakke, 438 U.S. 265 (1978).

Rigol, G. W (2002). *Best Practices in Admission Models.* New York: The College Board.

Sackett, P. R. & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology, 50*, 707-722.

Sackett, P. R., & Wilk, S. L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist. 49*, 929-954.

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

Schmidt, F. L. (1998). The problem of group differences in ability test scores in employment selection. *Journal of Vocational Behavior, 33*, 272-292.

Schmidt, F. L., Greenthal, A. L., Hunter, J. E., Berner, J. G., & Seaton, F. W. (1977). Job samples vs. paper-and-pencil trade and technical tests: Adverse impact and examinee attitudes. *Personnel Psychology, 30*, 187-197.

Schmitt, N., & Kunce, C. (2002). The effect of required elaboration of answers to biodata questions. *Personnel Psychology, 55*, 569-587.

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

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency using various predictor combinations. *Journal of Applied Psychology, 82*, 719-730.

Schrader, A. D., & Osburn, H. G. (1977). Biodata faking: Effects of induced subtlety and position specificity. *Personnel Psychology, 30*, 395-404.

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

Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of personnel selection procedures*. Bowling Green, OH: Society for Industrial and Organizational Psychology.

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

Thomas, L. L., Kuncel, N. R., & Crede, M. (2004*). Non-cognitive predictors of academic performance: The case of the Non-Cognitive Questionnaire (NCQ)*. Jossey-Bass Higher and Adult Education Series. San Francisco: Jossey-Bass.

Ting, S. M. R., & Robinson, T. L. (1998). First year academic success: A prediction combining cognitive and psychosocial variables for Caucasian and African American students. *Journal of College Student Development, 39*, 599-610.

Uniform Guidelines on Employee Selection Procedures. (1978). *Federal Register, 43*, 38290-38315.

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

Wolfram, S. (1999). *The Mathematica Book* (4th Ed). Cambridge, MA: Wolfram Media/Cambridge University Press.

Young, B. D., & Sowa, C. J. (1992). Predictors of academic success for Black student athletes. *Journal of College Student Development, 33*, 318-324.

Young, J. W. (2001). *Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis* (College Board Research Report No. 2001–6). New York: College Board.

Table 1

Dimensions of College Student Performance

Intellectual Behaviors

Knowledge, learning, and mastery of general principles (Knowledge)

Continuous learning, and intellectual curiosity (Learning)

Artistic appreciation and curiosity (Artistic)

Interpersonal Behaviors

Multicultural appreciation (Multicultural)

Leadership (Leadership)

Interpersonal skills (Interpersonal)

Social responsibility, citizenship and involvement (Citizenship)

Intrapersonal Behaviors

Physical and psychological health (Health)

Career orientation (Career)

Adaptability and life skills (Adaptability)

Perseverance (Perseverance)

Ethics and integrity (Ethics)

Note. These 12 dimensions were developed in Oswald et al. (2004).

Table 2

Changes in Regression Weights for the Different Simulated Schools Based on Effective Weights Applied to the Three Criterion Dimensions of College-Student Success

School Name	Weights on College GPA	Weights on OCB	Weights on BARS	Changes in Regression Weights (SAT/ACT, HSGPA, BIODATA)
School A	1	0	0	.42, .33, .05
School B	0.8	0.1	0.1	.11, .09, .08
School C	0.6	0.2	0.2	.08, .08, .11
School D	0.4	0.3	0.3	.06, .06, .13
School E	0.4	0.4	0.2	.05, .06, .12
School F	0.4	0.2	0.4	.06, .06, .13
School G	0.33	0.33	0.33	.04, .05, .13
School H	0.2	0.4	0.4	.03, .04, .15
School I	0.2	0.6	0.2	.02, .04, .13
School J	0.2	0.2	0.6	.03, .04, .15

Note. The last column indicates the changes in regression weights for the following predictor variables- SAT/ACT, High School GPA, Biodata (in that order) for each of the simulated schools.

Table 3

Intercorrelations between Predictor Variables and Criterion Variables

	1	2	3	4	5	6
Predictors						
1. SAT/ACT	1					
2. HSGPA	.48*	1				
3. Biodata	.17*	.20*	1			
Criteria						
4. College GPA	.59*	.55*	.18*	1		
5. OCB	-.12*	-.06	.29*	-.04	1	
6. BARS	.06	.10*	.47*	.15*	.38*	1

Note. * Correlation is significant at $p < .01$ (2-tailed).

Table 4

Percentage of Minorities Selected and Adverse Impact (AI) Ratios Based on the Different Schools Types for the Three Selection Ratios (.25, .50 and .75)

	Percentage Selected				AI Ratios		
School A							
	.25	.50	.75		.25	.50	.75
White	25	51	77	White	__	__	__
Asian	49	70	93	Asian	1.95	1.36	1.20
Black	3	21	40	Black	.11	.40	.51
Hispanic	16	40	64	Hispanic	.62	.78	.83
School B							
	.25	.50	.75		.25	.50	.75
White	25	50	77	White	__	__	__
Asian	45	76	93	Asian	1.83	1.50	1.21
Black	5	18	42	Black	.22	.35	.55
Hispanic	20	47	67	Hispanic	.81	.92	.87
School C							
	.25	.50	.75		.25	.50	.75
White	25	50	76	White	__	__	__
Asian	40	76	94	Asian	1.59	1.51	1.24
Black	7	21	48	Black	.28	.41	.63
Hispanic	29	49	69	Hispanic	1.16	.98	.91

School D							
	.25	.50	.75		.25	.50	.75
White	24	50	76	White	__	__	__
Asian	43	72	90	Asian	1.77	1.46	1.18
Black	7	27	55	Black	.28	.55	.72
Hispanic	29	51	69	Hispanic	1.19	1.03	.91

School E							
	.25	.50	.75		.25	.50	.75
White	24	50	75	White	__	__	__
Asian	43	71	91	Asian	1.79	1.42	1.20
Black	8	26	58	Black	.34	.52	.76
Hispanic	31	51	69	Hispanic	1.29	1.03	.91

School F							
	.25	.50	.75		.25	.50	.75
White	25	50	75	White	__	__	__
Asian	43	70	90	Asian	1.75	1.40	1.19
Black	5	26	56	Black	.22	.52	.74
Hispanic	29	51	71	Hispanic	1.18	1.02	.94

School G							
	.25	.50	.75		.25	.50	.75
White	24	50	75	White	__	__	__
Asian	42	69	90	Asian	1.76	1.38	1.19
Black	12	29	56	Black	.52	.58	.75
Hispanic	31	53	73	Hispanic	1.31	1.07	.97

School H							
	.25	.50	.75		.25	.50	.75
White	24	49	75	White	—	—	—
Asian	42	67	88	Asian	1.75	1.37	1.18
Black	11	33	62	Black	.46	.67	.82
Hispanic	31	53	73	Hispanic	1.30	1.08	.98

School I							
	.25	.50	.75		.25	.50	.75
White	24	49	74	White	—	—	—
Asian	41	65	88	Asian	1.71	1.32	1.19
Black	12	34	64	Black	.52	.69	.87
Hispanic	33	56	78	Hispanic	1.40	1.13	1.05

School J							
	.25	.50	.75		.25	.50	.75
White	24	49	75	White	—	—	—
Asian	41	65	88	Asian	1.69	1.32	1.18
Black	11	37	59	Black	.46	.75	.79
Hispanic	31	53	78	Hispanic	1.29	1.08	1.04

Note. White group is always the reference group across schools, and thus the AI ratio (deleted above) is always 1.0.