Problems with current predictors of success for graduate students (Graduate Record Examinations and grades) include restriction of range artifacts, grade inflation, and the increasing diversity of examinees. A case is made as to why noncognitive variables can add to the validity of selecting graduate students. Legal, moral, ethical, and statistical arguments are presented and discussed. Key points made in the report include the lack of logic in reflecting diversity with a single measure and the necessity of including a range of measures to be fair to all applicants. (Contains 1 table and 55 references.) (Author/SLD)
Why We Should Use Noncognitive Variables

With Graduate Students

William E. Sedlacek

Research Report # 5-01
Why We Should Use Noncognitive Variables

With Graduate Students

William E. Sedlacek

Research Report # 5-01

Paper presented at a Symposium on Noncognitive Assessments for Graduate Student Admissions, Graduate Record Examination Board, Toronto, June 21, 2001
Why We Should Use Noncognitive Variables With Graduate Students

William E. Sedlacek

Research Report # 5-01

Summary

Problems with current predictors of success for graduate students (Graduate Record Examinations and grades) include restriction of range artifacts, grade inflation, and the increasing diversity of examinees. A case is made as to why noncognitive variables can add to the validity of selecting graduate students. Legal, moral, ethical and statistical arguments are presented and discussed. Key points made in the report include the illogic of reflecting diversity with a single measure, and the necessity of including a range of measures to be fair to all applicants.
The simplest response to the title of this paper is that the measures that are commonly employed in evaluating prospective graduate students *don't work*. That is, the Graduate Record Examination (GRE) and prior grades (GPA) appear to have little validity in assessing graduate student potential (Sternberg & Williams, 1997; Sedlacek, 1998; Bair & Haworth, 1999). I believe there are several reasons for this conclusion that I will address below.

**Restriction of Range**

The GRE attempts to measure what nearly all measures of academic aptitude attempt to measure: verbal and mathematical ability. However, because the GRE is used with the population for which it is intended (e.g. the most capable individuals in our society) there is a great restriction of range of scores on the measure; we do not get scores on people who did not attend baccalaureate programs and performed reasonably well in them (Darlington, 1998). Restricting the range of scores we can study depresses our measures of association. While estimates of the effects of restriction of range are possible, they may be artifacts if the valid variance is not there to begin with.

Additionally, we do not know how to distinguish among potential graduate students on measured verbal and math abilities. Throughout most of the last century test developers have been refining available measures, with some success, but I feel we have come to the limits of our ability to develop such measures. Even if the population of potential graduate students might be distinguishable on the constructs of verbal and math ability, I do not feel that our current methods of test development are likely to do the job. Thus, I feel there is no more practical valid variance to capture in the verbal and math areas.
As a predictor of success GPA causes us even more problems in restriction of range than the GRE. Studies from the College Entrance Examination Board (CEEB) documented the problem of grade inflation, which appears to exist at all levels of education (Rigol & Kimmel, 1997). Students are receiving higher and higher grades, which restricts the range of possible GPAs we have to study. That applicants to graduate school tend to be on the higher end of the GPA distribution restricts our range even further.

As if a lack of variability in predictor scores weren't bad enough, the criteria that we wish to predict are also highly restricted. We know that graduate students tend to get even higher grades than undergraduates, and are more likely to be bunched at the top of the grade distribution. The criteria of retention or completion of degree offer some alternative indices of success, but again may offer fewer individuals in the noncompleter category than among undergraduates. So, on restriction of range alone, we severely limit our ability to effectively study the utility of GRE or GPA, but there are other problems in studying predictors of graduate student success.

Increasing Diversity

That our baccalaureate level students have been increasing in diversity on dimensions of race, gender, sexual orientation, religion, age, disability and many other attributes has been well documented (McTighe Musil, Garcia, Hudgins, Nettles, Sedlacek, & Smith, 1999). These changes have a potentially profound impact on the kinds of attributes we should be measuring in predicting success in graduate school. As we begin the 21st century, do we need to examine alternatives to admissions concepts developed in the early 20th century? The answer is yes! Potential measures are already developed and validated on many student groups, but they have
not been as widely utilized as they might be, particularly with graduate students. I have discussed several reasons for this including what I call the "Three Musketeers" problem (Sedlacek, 1994).

The Three Musketeers Problem

The rallying cry of "all for one and one for all" is one that we use often in developing what we think of as fair and equitable admissions measures. Commonly, our interpretation of how to handle diversity is to hone and fine-tune our measures so they are equally valid for everyone (Berk, 1982; Sackett, Schmidt, Ellingson & Kabin, 2001; Helms, 1992; Sedlacek, 1985). However, if different groups have different experiences and different ways of presenting their attributes and abilities, it is unlikely that we could develop a single measure, test item etc. that could be equally valid for all. If we concentrate on results rather than intentions, we could conclude that it is important to do an equally good job of selection for each group, not that we need to use the same measures for all to accomplish that goal. We want equality of results, not process. Therefore, we should seek to retain the valid variance that exists across diverse groups in our measures, rather than attempt to eliminate it.

Sternberg’s (1985, 1986) work on intelligence might prove instructive here. He suggested that there are three kinds of intelligence. Componential intelligence is the ability to interpret information in a hierarchical and taxonomic fashion in a well-defined and unchanging context. People who do well on standardized tests such as the Scholastic Assessment Tests (SAT) or the GRE have this type of intelligence. Experiential intelligence involves the ability to interpret information in changing contexts; to be creative. Standardized tests do not appear to measure this type of intelligence. Sternberg called his third type of intelligence contextual; it has to do with the ability to adapt to a changing environment; the ability to handle and negotiate the system.
If Sternberg's types of intelligence are applied to what is typically done in admissions in higher education, there is a heavy concentration on componential intelligence. Applicants who do not have traditional White middle or upper-middle class, mostly male-oriented experiences in the society may be less likely to show their abilities through componential intelligence than traditional applicants. These students will be called nontraditional here and include various racial-cultural groups, international students, women, gay, lesbian and bisexual students, athletes, students with learning disabilities or physical disabilities and older students. The list is intended to be illustrative not exhaustive.

Noncognitive Variables

Noncognitive variables have been defined in a number of ways in the literature. Some have seen them as extracurricular or nonacademic activities while others have used the term to describe motivational and personality variables (Sackett et al, 2001; Willingham, 1985) In this paper I am defining noncognitive variables as those that appear to reflect Sternberg's experiential or contextual intelligence, and may pick up on the abilities and potentials of nontraditional students as defined above (Sedlacek, 1996). The Noncognitive Questionnaire (NCQ) was developed to assess attributes that are more predictive of success in higher education for nontraditional students than are standardized tests (see Table 1) Work in assessing nontraditional variables with the NCQ supports the idea that nontraditional people often tend to show their abilities through experiential and contextual intelligence. (Ancis & Sedlacek, 1997;Boyer and Sedlacek, 1982; Fuertes & Sedlacek, 1995, Fuertes, Sedlacek & Liu, 1994; O' Callaghan & Bryant, 1990; Sedlacek, 1989, 1991, 1996a, 1996b, 1998a, 1998b, 1999; Sedlacek and Adams-Gaston, 1992; Ting, 1992; Tracey and Sedlacek, 1984, 1985, 1987, 1988, 1989; White and Sedlacek, 1986). Much of this is out of necessity because nontraditional people must learn to be
"multicultural" and examine issues from different perspectives. They must be able to negotiate a system that was not designed for them. Having long-range goals, a self-concept that includes how the system views you and an ability to handle racism are some of the scales on the NCQ.

Institutional racism is defined as the negative consequences that accrue to a member of a given group because of the way a system or subsystem operates in the society (e.g., college admissions) regardless of any other attributes of the individual (Sedlacek & Brooks, 1976). All "isms" (e.g., sexism, ageism) are included under the generic term "racism". The various "isms" take different forms but share a common basis. Thus, if there is a concentration on componential intelligence in admissions, less valid assessments will be done for nontraditional persons than for those with more traditional experiences in the system. This would be an example of institutional racism; unintended perhaps, but no less a serious problem.

It is not that componential intelligence is not important to nontraditional people; it may be that experiential and contextual abilities may be prerequisite (Westbrook and Sedlacek, 1988). If someone is struggling with racism in the system, time and energy may not be available to show componential talents. The point illustrates that there is a need to think of measures differentially in order to achieve equitable assessments for all. There is probably a classic oxymoron here in thinking that one can assess diversity of experience with a single measure. The arguments presented here are positive and proactive. Lowering standards of admission is not being advocated. The suggestion is to develop and use the most valid measures one can for all groups that can be operationally defined. Reliability of the NCQ has been estimated, and generally runs in the .80s employing several different methods (Tracey & Sedlacek, 1984; Sedlacek, 1996).
The noncognitive measures are also valid for traditional students, particularly against a retention or graduation criterion. Nearly all the references noted above show some validity for traditional applicants. For example, the handling racism dimension becomes “handling the system” for those not experiencing systematic discrimination. While the noncognitive dimensions discussed here are useful for traditional applicants, they are critical for nontraditional applicants.

There is a lack of evidence on the predictive validity of noncognitive variables for graduate students. However, they have shown validity for a wide range of students in many fields and there have been enough studies on medical and other health profession students (Webb, Sedlacek, Cohen, Shields, Gracely, Hawkins, & Nieman, 1997; Sedlacek & Prieto, 1990, Bandalos & Sedlacek, 1989; Helm, Prieto, & Sedlacek, 1997) to suggest that they would likely show some validity for graduate students.

Other Methodological Problems

Graduate students tend to matriculate in isolated, decentralized pockets of experience. Issues relating to success or failure of students may vary within institutions across fields, and across institutions within fields (Bair & Haworth, 1999). Thus, a student in physics may have a much different experience than one in art at the same institution, and students in art may have vary in their environments at different institutions.

Also, numbers of students in a given program, at a given institution, may be quite small, thus complicating a study of predictors of success.
How Should ETS Proceed?

A presumptive question, but one that is the point of the conference. First, stop trying to increase the validity of the GRE as it is currently focused. I do not believe there is any practical variance yet to tap that would deal with the problems noted above. Second, begin a research program on developing noncognitive variables as predictors of graduate student success. The program should employ multiple methods, as well as explore multiple predictors and criteria. In that research program, I would attempt to improve measures of the noncognitive variables shown in Table 1. I would also study additional variables that show promise.

Creativity

Sternberg (1985, 1986) stressed creativity as one of the key elements of experiential intelligence. However, creativity appears to have been under-studied considering its potential importance in understanding human behavior. Guilford (1950) in his American Psychological Association presidential address of 1950 called for more research on creativity. However, relatively little work in the area has been done since.

Sternberg & Lubart (1996) identified a number of reasons why creativity has not been studied more. One reason noted was that creativity is often seen as a mystical power rather than a measurable attribute. Another reason was seen as a focus on pragmatic approaches to generating creative ideas without understanding how creativity works. "Brainstorming" was suggested as an example.

Sternberg & Lubart also discussed the marginalizing of creativity research through considering it as a special case of a more general phenomenon rather than a potentially unique area. Additionally, considering creativity as the province of a single discipline within psychology has made the study of creativity rather obscure. The authors call for a
multidisciplinary study of creativity within psychology. I would suggest that the concept of multidisciplinary research be expanded to include researchers in many areas including those concerned with higher education and test development. Also, I would specifically include racial, cultural, and gender-related variables in the research program.

Sternberg & Lubart noted the difficulties in measuring creativity and felt that confluence theory was a potentially fruitful area for research on creativity. The basis of confluence theory is that a combination of cognitive and personality elements must combine to have creativity; elements such as "connects ideas", "sees similarities and differences" has "flexibility" and is "unorthodox." All of these elements would seem to be relevant to one or more of the noncognitive variables discussed above.

Identity

Another potential research area that may prove useful in admissions concerns identity theory. Expanding and refining our definitions and measurements of self-concept in this way would seem particularly valuable. The validity of self-concept as a predictor of educational performance for nontraditional persons was discussed above. Self-concept, as measured by the NCQ, contains elements of seeing oneself as a nontraditional person in a traditional situation. There is evidence that race and gender identity are important aspects of how one may show ability. Helm (1992) has developed identity models for Blacks, Whites and people of color. Neville, Heppner & Wang (1997) found racial identity attitudes of African American college students related to perceived stressors and coping styles. Ossana, Helms & Leonard (1992) have studied "womanist" identity as an important part of self-esteem and Jones (1997) found that race varied in its importance in the identity development of women in college, depending on their
race. Also, Frankenburg (1993) discussed the relevance of exploring the concept of “whiteness” in the identity development of White women.

Steele’s (1997) work on “stereotype threat” supports the importance of the psychological set with which examinees approach a test. If African Americans are told that they do not usually do well on a test, they do less well than if a more positive set is given. That African Americans don’t do as well as Whites on standardized tests is well documented in the professional and popular literature (Sedlacek, 1998; Lehman, 2000). Therefore, for African Americans, the act of taking a high-stakes test probably involves dealing with the racism that has been involved in creating the stereotype threat in the first place. Hence, the variance that is being measured when an African American takes the GRE, is likely in part, handling racism, one of the noncognitive variables I am proposing.

If nontraditional individuals, as I have defined them here, can approach an evaluation with a feeling of empowerment, or expected success, they may be employing a noncognitive skill that would give a more accurate prediction of their potential. One of the consequences of including noncognitive variables in admissions programs would likely be to reduce the racism that has developed within our assessment systems. I believe we should not only seek the best predictors of success, but we should, as a testing and evaluation industry, try to reduce racism in our procedures wherever we can.

Other aspects of self-concept have been explored, which might be worth further study. For example, there has been work on religious identity (Fowler 1981; Suthkaran & Sedlacek, 2001) and lesbian, gay, and bisexual identities (D’Augelli, & Patterson, 1995; Mohr, Israel, &
Sedlacek, 2001). Again, my intent is not to list all possibilities, but to suggest some areas that might prove fruitful.

Multiple Methods

Noncognitive variables can be assessed and researched using several techniques.

Questionnaires

The Noncognitive Questionnaire (NCQ) has been shown to have validity and reliability in assessing the eight noncognitive variables noted above. Several forms of the NCQ have been developed and employed in different admissions contexts. The questionnaire can be administered on-line. The Gates Millennium Scholars program assesses the eight noncognitive variables shown in Table 1 using short answer questions in awarding scholarships to students of color. The use of a noncognitive questionnaire has been presented by the plaintiffs in Castaneda v. The University of California Board of Regents, as the preferred method of increasing minority student enrollment at The University of California-Berkeley. A version of the NCQ has been shown to have validity in selecting traditional and nontraditional students to health programs at a western state community college (Noonan, Sedlacek & Suthakaran, 2001). Students are evaluated in project ACCESS at Prarie View A&M University using the NCQ, and they report greater validities for the measure than for grades or test scores.

Interviews

It is feasible to interview applicants using noncognitive variables. The key is to train interviewers to identify how applicants may show high or low scores. The Louisiana State University Medical School in New Orleans has employed noncognitive variables in their admissions program through interviews starting in the late 1980s. In the 10 years since the use of noncognitive variables was introduced, enrollment of students of color doubled to 21 percent
with an 87 percent retention rate. During this period, admissions committee members were trained to interview concerning the eight noncognitive variables shown in Table 1 using simulated cases. More than 80% of the admissions committee members felt the noncognitive variables were useful in admissions and 92% thought the training helped them identify the noncognitive variables in applicant interviews (Helm, Prieto & Sedlacek, 1997). The admissions committee thought self-concept (97%), realistic self-appraisal (95%), leadership (84%), support person (83%), and handling racism (81%) were the most useful indicators of “minority” student success. Sixty one percent felt grade point average and 57% felt Medical College Admission Test scores were useful for minority admissions.

The University of Maryland Medical School employs interviews to assess applicants on the noncognitive variables shown in Table 1. They have defended their validity in an ongoing lawsuit that has challenged their fairness (Farmer v. Ramsay).

Portfolios

The use of portfolios provides yet another way to assess noncognitive variables (LaMahieu, Gitomer & Eresch, 1995). Portfolios have been commonly used in the arts to demonstrate the work of applicants for admission.

The School of Design at North Carolina State University in Raleigh has required an additional admissions procedure beyond the general one employed for all undergraduates. They have traditionally required a portfolio containing design-related materials produced by the applicant. Administrators and faculty at the school wished to broaden the content of the portfolio to contain information on noncognitive variables, such as how they had overcome obstacles, how they saw themselves and what were their goals. They felt this would give them better information on which to judge their applicants, particularly those of color. Faculty evaluators
were trained in identifying examples of high and low scores on noncognitive variables. Such training is important in order to avoid one of the potential problems in portfolio assessment; that middle-class students may benefit most from such assessments (Koretz, 1993).

The University of California, Irvine included a Personal Achievement Profile along with SAT or ACT scores, grades and specific courses completed as part of its admission profile. It included, among other things, the noncognitive variables of leadership, community service and creative achievement. After applicants were screened on their academic credentials about 60% of the admissions were determined. The additional 40% of the admissions were selected based on the Personal Achievement Profile. Using a double-blind procedure admissions staff trained in reviewing the profiles made the judgments. No interviews or letters of recommendations were employed and the entering class cut across a number of dimensions.

Essays

With appropriate training, it is possible to have raters score essay material on noncognitive variables. For example, in the Gates Millennium Scholar program, readers were able to score applications with high reliability on the noncognitive variables shown in Table 1. A normal distribution of scores for over 16,000 applications of students of color was achieved. This sample includes several hundred graduate students, and will include more as the recipients move on to advanced academic work. Validity studies are underway relating those scores to academic and nonacademic outcomes.

Conclusions

Based on the logic, research and suggestions above there appear to be a number of reasons to proceed with some plans to assess, study and potentially implement programs to employ noncognitive variables in admitting graduate students. As argued above, our previous
efforts in designing the more traditional verbal and math ability tests may have run their course. It appears that if noncognitive variables were to be employed in admitting graduate students, more diversity along many dimensions could be achieved, without directly selecting on race, gender or other attributes. Noncognitive variables may provide a solution to many of the legal, moral, ethical and practical problems presented in balancing the validity of assessment measures with achieving diversity and fairness in selection. A number of possible noncognitive variables have been studied and shown validity with a variety of other groups of applicants. Some time, creativity and resources should be spent on the effort of studying them, and possibly including them, in the selection of graduate students.

References


Table 1

NONCOGNITIVE ADMISSIONS VARIABLES

William E. Sedlacek


II. REALISTIC SELF-APPRAISAL, especially academic. Recognizes and accepts any deficiencies and works hard at self-development. Recognizes need to broaden his/her individuality.

III. UNDERSTAND AND DEALS WITH RACISM. Realist based upon personal experience of racism. Is committed to fighting to improve existing system. Not submissive to existing wrongs, nor hostile to society, nor a "cop-out." Able to handle racist system. Asserts school or organization role to fight racism.

IV. PREFERS LONG-RANGE GOALS TO SHORT-TERM OR IMMEDIATE NEEDS. Able to respond to deferred gratification.

V. AVAILABILITY OF STRONG SUPPORT PERSON to whom to turn in crises.

VI. SUCCESSFUL LEADERSHIP EXPERIENCE in any area pertinent to his/her background (gang leader, church, sports, noneducational groups, etc.)

VII. DEMONSTRATED COMMUNITY SERVICE. Has involvement in his/her cultural community.

VIII. KNOWLEDGE ACQUIRED IN A FIELD. Unusual and/or culturally-related ways of obtaining information and demonstrating knowledge. Field itself may be non-traditional.
NOTICE

Reproduction Basis

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").

EFF-089 (3/2000)