Predictive Validity of Grade Point Averages and of the Miller Analogies Test for Admission to a Doctoral Program in Educational Leadership

I. Phillip Young
University of California Educational Research Center
University of California, Davis

A very important administrative task for most doctoral programs in educational leadership is the admission of students. Each academic year, applicants apply and graduate faculty must delimit this pool of potential candidates. The ultimate goal of faculty within the admission process is to select among the most able candidates.

To select among the most able candidates seeking admission, all doctoral programs in educational leadership rely on purported predictors of future academic performance. Some of these predictors are subjective, while others of these predictors are objective. Through considering information on both types of predictors, faculty members make decisions about who will be extended and who will be denied admission to a doctoral program.

Subjective predictors include reference information provided by applicants through sources of their own choosing, and in some instances individual interviews with potential candidates by faculty committees. Objective information about applicants has been obtained generally by several means. These means include measures of past academic performance as assessed by grade point averages and measures of future academic potential as assessed by standardized tests.

In most instances, the purported validity of all these predictors has been assumed rather than has been assessed at the

department/program level. This oversight is disappointing given the recommendation by a leading organization, "Departments using GRE scores for graduate admission, fellowship awards, and other approved purposes are encouraged to collect validity information by conducting their own studies" (Educational Testing Service, 2004). In keeping with this basic recommendation, only recently has empirical data begun to emerge in the professional literature about the predictive validity of many of the predictors used to select doctoral candidates and addressed at the department/program level.   To compliment these efforts, this manuscript continues this research stream by assessing the predictive validity of several predictors used to delimit an initial applicant pool of doctoral candidates at the department/program level. Particular predictors addressed in this manuscript are measures of past academic performance and of future academic potential. Past academic performance is assessed by grade point averages, while future academic potential is assessed by scores from the Miller Analogies Test (MAT) (Miller Analogies Test, 2004).

Related Literature

Selection of employees in general (Delli & Vera, 2004) as well as selection of doctoral students has been conceptualized as a process rather than as an event.   Underlying this process perspective is that applicants must apply, must be selected, and must enroll to consummate the matriculation process. For each stage of this process, research is needed to guide deliberate decision making both by applicants and by faculty.

Of these different stages encompassing the entire selection process, this study focuses on outcomes from the selection process based on the performance of applicants relative to objective academic measures used to delimit an initial applicant pool. This research, as reported here within, compliments the emerging research stream in several ways. Most importantly, it builds on other research as reported in this learned journal for an educational administration program and extents findings as addressed by different professional schools in law (Johnson, Davis, Sterling,

Jones, &Anderson, 1986) as well as in medicine (Young, 1995).

Published research, as reported in this learned journal, indicates that selection decisions of faculty members can be guided, at least in part, by a reliance on recommendations for potential doctoral candidates (Young, 2005a). Within this published research, recommendations for potential doctoral applicants have been deconstructed in several ways. That is, recommendations can be either personal or professional, can be either norm referenced as compared to other individuals or criterion referenced as compared to an external standard, can be either structured through a specific form for recommendations or unstructured by letters of recommendation, and/or can vary in specific content addressed by reference sources.

Outcomes from this research indicate that validity for subjective information about applicants varies according to each of the above mention specifications. Valid subjective information obtained via references is most likely acquired through professional sources, structured forms, norm referenced, and addressing specific content. Of the many content items considered by a single study, only information about "perceived research ability" and "perceived work habits" of perspective applicants as provided by professional reference sources were found to differentiate between those rejected and those accepted to a doctoral program in educational leadership. Beyond subjective information about potential doctoral candidates, other studies have explored objective information used to delimit an applicant pool for a doctoral program. Objective information for applicants' receiving attention in the professional literature is measures of past academic performance and of future academic potential. Past academic performance has been measured by grade point averages, and future academic potential has been assessed by results from standardized examinations.

According to Creighton and Jones (2001) as well as to Norton (1996), a great deal of difference is afforded to past academic performance when delimiting an initial applicant pool. Past academic performance is assessed generally from transcripts submitted by potential doctoral candidates. Separate indicators of

past academic performance are computed for undergraduate grade point averages (UGPA) and for graduate grade point averages (GGPA).

With respect to future academic potential of perspective doctoral candidates, most doctoral programs require applicants to submit results from a standardized examination. In practice, these results are obtained through scores either on the GRE or on the MAT. Of these two measures for potential academic performance of perspective doctoral candidates in the area of educational leadership, research, to date, has focused only on the GRE (Young, 2005b). Yet to be addressed in this emerging research stream is comparable information about the MAT, and the focus of this manuscript is to fill partially this void in current knowledge.

This manuscript does so in several specific ways. First, a compensatory model of decision making is used to assess the predictive validity of past academic performance and of future academic potential for admission to a particular doctoral program in educational leadership. The compensatory model assumes that high scores on one predictor can offset low scores on another predictor and is in contrast to a multiple hurdles model advocating a specific cut score on each predictor in isolation (for a discussion of decision models see Heneman & Judge, 2006). As such, the compensatory model considers the unique contribution of each academic predictor in light of all academic predictors in combination through a linear equation taking into consideration the intercorrelation among predictor variables used to delimit an initial applicant pool.

Second, to assess the predictive validity of past academic performance and of future academic potential for perspective doctoral applicants, field data are collected over a ten year period. Within this timeframe, actual applicants seeking admission to a doctoral program are classified according to their outcome status. That is, perspective doctoral candidates are classified as rejected, as accepted but not graduating, or as graduating.

Methodology

The time frame for this study is 10 years (1991-2001), and the population for this study is 102 applicants seeking admission to a particular doctoral program in educational leadership located within a Pacific coast state and satisfying the admission requirement by taking the MAT. This program serves, largely but not exclusively, 120 public school districts and attracts a diverse population with approximately 50% of those taking the MAT being female and the average age being 43.0 years (SD = 9.0). With respect to their performance on academic predictors, the average UGPA is 3.10 (SD = .39), the average GGPA is 3.64 (SD = .27), and the average MAT is 53% (SD = 30.9).

Academic Predictors. As part of the admission process, all applicants were required to submit evidence of their past academic performance and of their future academic potential. Past academic performance was assessed for undergraduate grade point averages (UGPAs) and for graduate grade point averages (GGPAs) as depicted by official transcripts. Potential academic performance was assessed by standardized scores from the MAT.

Grade point averages were assessed on a traditional 4-point scale with lower numbers reflecting a less satisfactory performance. Standardized test scores from the MAT could range from "1" to "100" with higher scores denoting a superior performance. Only standardized scores were used for the MAT because applicants took this instrument in different years with different norm groups across the time span covered in this study.

Statistically Analysis

To assess the predictive validly of past academic indicators and future academic performance as explored in this study for delimiting an applicant pool, a discriminant analysis was used. The classification variable consists of three mutually exclusive categories as required by a discriminant analysis. These mutually exclusive categories are as follow: (a) coded "1" are those applying but rejected [n= 48], coded "2" are those accepted but failing to graduate [n= 19], or coded "3" are those graduating [n= 35].

Descriptive information broken down by group membership on all academic predictors is found in Table 1. An

examination of these data by group membership indicates a sharp departure relative to representation among the groups (rejected=48, accepted=19, and graduated=35). Following the recommendation of know authorities for using a discriminant analysis, a test of the covariance matrices revealed that these data meet the homogeneity of variance assumption (Box's M, f=.84, p=.61) even though differences exist in membership numbers across levels of the classification variable.

Table 1: Descriptive Statistics by Group Classification for Academic Predictors

Classification		Mean	Std. Deviation	Range	N
Rejected	MAT %	43.8750	29.1676	93	48
	UGPA	3.0360	.4440	1.91	48
	GGPA	3.6283	.2769	1.00	48
Accepted	MAT %	64.6842	28.5658	93	19
	UGPA	3.1447	.2657	.82	19
	GGPA	3.5211	.3058	1.00	19
Graduated	MAT %	60.0286	31.2847	98	35
	UGPA	3.1749	.3651	1.42	35
	GGPA	3.7174	.2292	1.00	35
Total	MAT %	53.2941	30.8720		102
	UGPA	3.1039	.3917		102
	GGPA	3.6389	.2735		102

Based on these findings relative to sample size requirements (3:1) for a valid discriminant analysis as suggested by Tatsuoka (1970) and on homogeneity of covariance matrices as suggested by Stevens (2002), a discriminant analysis was performed. Within this discriminant analysis two discriminant functions surfaced that are statistically significant (X^2=16.95, df=6, P ›.01; X^2=6.30, df=2, P ›.05). These later discriminant functions account for 64% and 36% of the relative variance and reflect canonical correlations of .32 and .25 for each function, respectively. An examination of the structure matrix coefficients and of the standardized canonical discriminant function coefficients indicates that each function is influenced largely by a single

academic predictor (see Table 2). Structural matrix coefficients indicate that the major correlation among each potential academic predictor and each discriminant function is the MAT percentile score for function one (.86) and UGGPA for function two (.96). However, because structure matrix coefficients fail to consider the redundancy of contributions for particular discriminating variables, an examination is made of the canonical discriminant coefficients.

Table 2: Standardized and Structured Matrix Coefficients for Academic Predictors

Variables	Structure Matrix Coefficients		Canonical Dis. Coefficients	
	Function 1	**Function 2**	**Function 1**	Function 2
Mat %	**.857**	**.383**	**.823**	.290
UGPA	**-.271**	**.957**	**.406**	.007
GGPA	.390	.408	-.503	.926

When examining the canonical discriminant coefficients derived with these data and when controlling for redundancy of information among potential discriminating variables, other implications are suggested with these data. For the first discriminant function, the percentile score from the MAT is the single most important influence relative to unique contribution (.82). However, graduate grade point average (.93) instead of undergraduate grade point average (.007) emerges as the most important unique contributor for the second discriminant function as noted by the canonical discriminant function coefficients.

To assess the satiability of these results, a "hold out one" process is used (Lachenbruch, 1967) whereby multiple iterations are computed, and within each iteration computed, a single (albeit different) case is held out to compute independent classification scores for the group of applicants. For example, case A is omitted from the first iteration in the data analyses, and a classification statistic is computed for case A. In the second iteration, case A is included in computing total classification weights, while case B is held out of the computation process for generating separate classification statistics for case B, and this process of holding one

out continues throughout the data analyses across all individuals.

Results of the classification analysis indicate that 52% of the individuals are classified correctly based on the norm group. For the hold one out group, 50% of the participants in this analysis are classified correctly. Collective, these results indicate very little shrinkage and suggest the stability of the equations.

To assess how classification levels vary on the different discriminant functions, group centroids are plotted. Contained in Figure 1 are data depicting each classification level according to group centroids. These data indicate that those rejected differ from those accepted or those graduating and that those accepted vary little from those graduating.

Conclusions

Information about the predictive validity of academic measures used to delimit initial applicant pools for doctoral programs in educational leadership is well warranted within the professional literature. From an overall perspective, this information serves applicants as well as faculty within the decision making process. Inadequate selection decisions fail to serve either party well because inadequate selection decisions result in a poor use of resources by both parties, and some authorities indicated that only approximately 50% of those admitted in the past will ever graduate (Dorn & Pappa-Lewis, 1997).

Figure 1: Group Centroids relative to Discriminant Functions for Academic Predictors

Canonical Discriminant Functions

Function 1
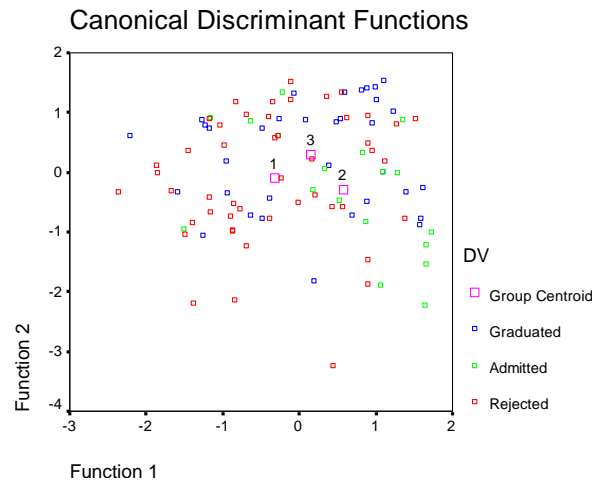
Somewhat reinforcing from findings of this study is that all the academic predictors used to delimit an applicant pool for a specific doctoral program in educational leadership portray at least some validity (see structural matrix coefficients in Table 2). However, when viewed from the lenses of a compensatory model certain academic predictors carry more weight that others given the intercorrelation among academic predictor variables (see canonical discriminant coefficients in Table 2). No doubt, through using a compensatory model for delimiting an applicant pool, better selection decisions can be made than by affording individual weight to these predictors as implied by a multiple cutoff model of decision making.

These findings, as well as existing findings addressing academic predictors, echo the relative importance of standardized test results over past grade point averages. Those that excelled on standardized test scores were more likely to be admitted and more likely to graduate than those performing less well on these measures. At first glance, this would seem to be a self fulfilling hypothesis.

That is, only those admitted had high scores and only those denied had low scores. However, an inspection of the basic data as contained in Table 1 for the range as well as for the standard deviation of particular values on academic predictors assessed in

this study suggests this is not the case. Some of those admitted as well as some of those graduating had lower scores than some of those denied.

Functionally, the above findings illustrate that all selection procedures are prone to errors. Both false negatives (inappropriately rejected) and false positives (inappropriately accepted) will always exist in the applied setting when delimiting an initial applicant pool. However, errors of both types can be reduced through appropriately derived linear equations based on valid predictors (see classification results).

Finally, this study, like all studies, suffers from certain limitations. Most importantly, these equations were assessed for a particular doctoral program, and outcomes will vary across programs, especially specific weights. Until further research is conducted, any generalization beyond these finding should be made with extreme causation, and these findings should serve as a beginning and not as an ending for this very important administrative process used to delimit an initial applicant pool for a doctoral program in educational leadership.

References

Creighton, T.B. & Jones, G.D. (August, 2001). Selection or self selection? How rigorous are our selection criteria for educational administration programs. Paper presented at the Conference of National Professors of Educational Administration. University of Houston: Houston , Texas.

Delli, D.A. & Vera, E.M. (2004) Psychological and contextual influences on the teacher selection interview: A model for future research. *Journal of Personnel Evaluation in Education, 17* (2), 137-155.

Dorn, S.M. & Papalewis, R. (1997). Improving doctoral retention. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Educational Testing Service (2004). Guidelines for the use of GRE scores. Retrieved October 21, 2004 from

http://www.gre.org/scoreuse.html.

Heneman, H.G. & Judge, T.A. (2006). *Staffing organizations* (6th ed.), Middleton WI: Mendota House.

Johnson, D.G., Davis, G., Sterling, M., Jones, R.F., & Anderson, J. (1986). Predicting academic performance at a predominantly Black medical school. *Journal of Medical Education, 61* (8):629-639.

Lachenbruch, P.A. (1967). An almost unbiased meth0d of obtaining confidence intervals for the probability of misclassification in discriminant analysis. *Biometrics, 23*, 639-645.

Miller Analogy Test. (2004). *Technical manual*. Harcourt Assessment.

Norton, M.S. (1994). *Student recruitment and selection in educational administration programs. Arizona State University*. (ERIC Document Reproduction Service No. ED. 366 087).

Stevens, J.P. (2002). *Applied multivariate statistics for the social sciences (4th ed.)*. Lawrence Erlbaum, Mahwah:NJ.

Tatsuoka, M.M., (1970). Discriminant analysis. *Institute for personality and ability test.* Champaign, IL..

Young, I.P. (2005a). Predictive validity of applicants' reference information for admission to a doctoral program in educational leadership. *Educational Research Quarterly. 29* (1), 16-25.

Young, I.P. (2005b, July). Predictive validity of applicants' reference information for admission to a doctoral program in educational leadership. Paper presented at the National Conference of Professors of Educational Administrations, Washington D.C.

Young, J.W. (1995). A comparison of two adjusted methods for improving the prediction of law school grades. *Educational and Psychological Measurement, 55* (4) 558-571.