Predictive Validity of Applicants' Reference Information for Admission to a Doctoral Program in Educational Leadership

I Phillip Young
University of California-Davis

Admission to a doctoral program is both an individual and a faculty activity. Individuals must make decisions to apply and faculty must make decisions about acceptance. Decision making of these different stakeholders is addressed by only limited research within the professional literature.

For individuals as applicants, research suggests that their decision making can be influenced by the content of messages as communicated through recruitment brochures. To illustrate, Young, Galloway, and Rinehart (1991) manipulated the content of recruitment brochures to emphasize economic advantages associated with attending a doctoral program, psychological benefits enjoyed by students enrolling in a doctoral program, or criteria/hurdles necessary for admission to a doctoral program. These investigators found that doctoral program brochures describing criteria/hurdles were viewed as the most appealing to potential applicants and that those individuals were concerned most about how to "get in" when making decisions to apply for as well as to accept an offer to enroll.

Turning from recruitment to selection, graduate faculty members must make admission decisions about applicants. Admission decisions made by faculty are important because several studies suggest that only approximately 50% of the applicants admitted ever complete a doctoral program (Dorn & Papalewis, 1997; Marcus, 1997). Consequently, inadequate selection decisions on the part of faculty fail to serve well either students or doctoral degree programs.

Indeed, a major task for graduate faculty charged with administering a doctoral program in education is the selection of students from a larger applicant pool seeking admission. When fulfilling this gate keeping function relative to the admission process, graduate faculty members rely on several sources of information about potential candidates to guide their deliberation. These sources of information vary, however, in important ways and provide insights from different perspectives.

Some of the information pertains to perceptions about the past academic performance of potential doctoral candidates. Most notably among these indicators of past academic performance are grade point

averages as reflected on official transcripts. Grade point averages are considered generally both for undergraduate as well as for graduate coursework on the basic assumption that past performance is the best predictor of future performance (Schmitt & Chan, 1998).

In addition, many doctoral programs require potential applicants to submit scores from a standardized test. Most common among these tests are scores either from the Graduate Record Examination (GRE) (Educational Testing Service, 2004) or from the Millers' Analogy Test (MAT) (The Psychological Corporation, 1994). Scores from a standardized test (either the GRE or the MAT) provide faculty with a national norm reference source of comparison.

Research addressing grade point averages as well as scores from standardized tests indicates that faculty members afford great deference to these potential predictors when considering candidates (Norton, 1992). With respect to undergraduate grade point average, Creighton and Jones (2001) surveyed 450 different institutions of higher education and found that minimum grade point average requirements ranged from a low of 2.50 to a high of 3.00. In general, these purported cutoff levels were established by graduate schools at large rather than by individual doctoral programs located within a particular university.

For graduate grade point averages and for standardized test scores, only a single study exists relative to doctoral programs focusing on educational leadership. This study (Young, 2005) examined the predictive validity of these academic predictors for the selection of doctoral candidates and found that both graduate grade point average and test scores from standardized measures were valid predictors of future performance. Within this study, consideration was afforded both to the GRE and/or to the MAT as a predictor within the admission process.

Unexplored within this current research stream is any information about additional less objective potential predictors used within the student selection process. Most importantly overlooked within the admission research are data addressing reference information about potential doctoral candidates provided by external others purportedly knowledgeable about an applicant's potential for success. This void is particularly surprising because with few, if any exceptions, most all doctoral programs require reference information about potential candidates for admission consideration, and graduate faculty members consider this information, at least in part, to inform their selection decisions when choosing among applicants (Norton, 1992).

However, unlike grade point averages and standardized test scores, reference information can vary in several ways. These ways are as follow: (a) focus, (b) confidentiality, (c) medium, (e) format, and/or (f) content. Because all of these ways of variation have implications for validity assessments, each is discussed and controlled in subsequent statistical analyses.

Focus of reference information can be either personal or professional, and this difference is determined largely by the source from which the reference information is obtained. Personal reference information is obtained generally from significant others (e.g., community leaders, ministers etc.) not necessarily familiar with an applicant's academic potential in a doctoral program. On the other hand, profession reference information is obtained most likely from those knowledgeable about an applicant's academic potential (e.g., college professor, immediate supervisor, colleague etc.).

Confidentiality of reference information about applicants can be either confidential or non-confidential. According to the Family Educational Rights and Privacy Act (1996) potential applicants can wave (confidential) or fail to wave (non-confidential) their right for access to reference information. However, independent of their choice (waved or not waved), all reference information must be afforded the same deference by decision makers relative to admission decisions.

Medium used to collect reference information about potential graduate students can vary by requiring either letters of recommendation or completion of a standardized reference form. Each medium has certain advantages as well as disadvantages for guiding decision making of graduate faculty within the admission process. A major advantage associated with letters of recommendation is the context rich information provided about particular applicants, while a major disadvantage associated with letters of recommendations is the lack of consistent content across candidates.

In contrast to letters of recommendations is the structured reference form used by many doctoral programs. The structured reference form limits both the amount of and the type of reference information about applicants afforded to faculty within the admission process (disadvantage). However, this medium does provide the same information about all candidates and facilitates comparison among applicants seeking admission (advantage).

When using a structured reference form to collect information about potential doctoral candidates, different response formats exists.

Response formats can be either criterion referenced or norm referenced. The former format requires a reference source to evaluate applicants on a graphic rating scale (e.g., 1-5 point Likert type scale) relative to a perceived level of performance, while the later format requires a reference source to rate applicants relative to the performance of other individuals they have known (e.g., as compared to lowest 5% etc.).

For a structured reference form using either response format, reference information sought within the admission process can differ in content. Content contained within a structured reference could vary according to psychological processes (e.g., motivation/maturity), academic proclivity (e.g., research interest/basic intellectual skills), and/or communication skills (e.g., written/verbal). Depending on the particular content assessed by a structured reference form, the utility of information imparted may well exhibit different degrees of validity for guiding graduate faculty members in their selection of doctoral candidates.

Advancements

To provide information addressing this void within the professional literature about the utility of reference information for guiding the decision making of graduate faculty is the focus of this study. More specifically, this study examines the utility of reference information for the selection of doctoral candidates from several perspectives. Most importantly, this study focuses on "real" as opposed to simulated selection practices by examining historical data used to make actual admission decisions.

As noted within the literature review, reference information for potential doctoral candidates can vary in many meaningful ways. Given these ways of variation, this study holds constant the focus by using only professional reference sources, the confidentiality of information by examining only those situations where applicants have waived their rights, the medium used to collect reference information by examining responses to a structured reference form, the format of the structured reference form to included a norm as opposed to a criterion anchored response format, and the content of the reference form by requiring reference sources to rate all potential doctoral candidates on the same measures (psychological processes, academic proclivity, and communication skills). In contrast to sources of variation held constant across potential predictors, manipulated within this study is the success of applicants for being admitted to a particular doctoral program (rejected or accepted).

Methodology

The population for this study is all applicants (N=306) applying to a particular doctoral program between the academic years of 1990 and 2000. This doctoral program focuses on educational leadership and is located in a Pacific coast state. As a total applicant pool, these candidates varied in diversity relative to sex (51%, female), chronological age (mean = 42 years old, SD = 9 years), undergraduate GPA (mean=3.09, SD = .43) and graduate GPA (mean= 3.65, SD =.28).

Procedure and Instrumentation

Admission to this particular doctoral program is fixed rather than rolling throughout the academic year and allows a comparison among all candidates within a given year at the same time. To be considered as a viable candidate for admission, all applicants must have on file by a specific date in early winter a completed dossier containing all transcripts, standardized test results, and reference information from three sources. Of these purported predictors, The focus of this study is on reference information.

Reference information was assessed via standardized form completed by three professional sources nominated by applicants. This form contains nine different areas for evaluating potential candidates: intellectual ability, knowledge of education, motivation, research ability, maturity, work habits, problems solving aptitude, oral communication, and written communication. For each area assessed, reference sources were requested to evaluate an applicant's standing on a norm based 5-point scale relative to others they have known (i.e., 1=lower 25%---5=top 5%) or to indicate that they are "unable to judge the applicant" in a particular area.

Of those 306 applicants seeking admission, three completed reference forms containing evaluations on the 1-5 scale for each area were obtained for 243 persons (for 63 applicants some reference source indicated "unable to judge the applicant on at least one criterion) These ratings were used to compute composite scores for each area assessed (n=9) by summing across reference sources (n=3) provided by each candidate. For every applicant, a total of nine composite scores were obtain that could range potentially from a low 3 to a high of 15 with higher scores being more positive, and descriptive statistics as well as correlations among these areas are found in Table 1.

Table 1
Descriptive Statistics and Correlations among Variables

	Mean	S.D.								
Intellectual Ability	13.36	1.88	1.00							
Educational Knowledge	13.22	2.13	.76	1.00						
Motivation	14.09	1.70	.79	.78	1.00					
Research Ability	13.17	2.63	.57	.64	.59	1.00				
Maturity	13.68	1.89	.77	.80	.84	.61	1.00			
Work Habits	13.90	1.76	.79	.81	.90	.63	.87	1.00		
Problem Solving	13.51	2.00	.79	.80	.80	.64	.81	.81	1.00	
Verbal Ability	13.33	2.11	.75	.79	.75	.60	.80	.76	.83	1.00
Writing Ability	13.02	2.25	.74	.77	.72	.67	.75	.76	.78	.85	1.00

**All correlations are significant at the 0.01 level (2-tailed).
N=243

As can be observed in Table 1, high inter-correlations exist among the different areas. All ratings are highly favorable (see means) as well as related (see correlations), and these findings could be expected because applicants chose their own reference sources (Petress, 1999). Indeed, an initial principal component analysis of this correlation matrix indicates the emergence of only a single factor as assessed by a scree test based on plotted eigenvalues.

Given the emergence of a single latent structure among the reference items, two separate approaches exist for additional analyses concerning predictive validity. One approach involves computing factor scores for each individual as suggested by the factor analysis, while the other approach involves treating each reference item for individuals as a separate predictor variable. Although both approaches utilize a common statistical procedure (regression analysis) for assessing predictive validity, each serves a different purpose.

The factor analytic approach is theoretical in orientation, assumes the configuration of reference items as fixed, and seeks to assess the value

of a latent variable. In contrast, the individual reference item approach is empirical, assumes the configuration of items may be changed in the future, and assesses the unique contribution of particular items within a linear combination. Because the purpose of this study is to assess the predictive validity of particular reference items and not the utility of a specific latent variable as calculated via factor scores, the later approach is followed in the subsequent regression analysis.

Statistical Analysis

To assess the predictive validity of specific reference information within the admision process for these potential doctoral candidates rejected or admitted, a logistic regression analysis was performed. The dependent variable in this regression analysis contains two levels where one level includes those rejected from the doctoral program (N=107), while the other level includes those accepted into the doctoral program (N=136). All potential predictor variables, as contained in Table 1, were entered in a blockwise fashion, and results of the regression analysis are found in Table 2.

Table 2
Statistically Significant Variables entering the Logistic Regression

	B	S.E.	Wald	df	Sig.	Exp(B)
Research Ability	.377	.083	20.844	1	.000	1.458
Work Habits	-.309	.109	8.065	1	.005	.734
Constant	-.429	1.065	.162	1	.687	.651

Results contained in Table 2 indicate that a statistically significant linear combination exists for differentiating between those candidates accepted and those candidates rejected. Overall, approximately 14% of the variance between these different groups (rejected or accepted) was accounted for as measured by Nagelkerke's R^2. Contributing to this R^2 is only two out of nine of the reference items, research ability and work habits.

Interestingly, the direction of the signs associated with the different regression weights is inconsistent. For research ability the sign is positive, while for work habits the sign is negative. Perceived capability (research ability) rather than assessed initiative (work habits), as provided by reference sources for applicants, is found within the linear

equation to be more controlling in terms of magnitude for regression weights as well as differing in direction (see signs for regression weights).

Conclusion and Discussion

Graduate faculty members rely on various sources of information about candidates when choosing among potential applicants for admission to a doctoral program (Norton, 1994). Some information sources are based on past academic performance (grade point averages) (Creighton & Jones, 2001), while other information sources are based on potential academic performance. Included among the later potential information sources are empirical results from standardized examinations (Young, 2005) and subjective opinions of professional reference sources.

It is the later of these informational sources pertaining to potential academic performance of aspiring doctoral candidates that is examined in this study. This study focuses on the perceptions of designated reference sources provided by applicants as a means for differentiating between those accepted and those rejected for a particular doctoral program focusing on educational leadership. Findings from this study indicate that reference information provided by reference sources varies in utility relative to differentiating between those rejected and those accepted for a particular doctoral program.

The percentage of variance accounted for within this study is not inconsequential. For example, similar percentages have been reported both for law schools (Young, 1995) and for medical schools (Zeleznik, Hojat, & Veloski, 1987) using more well developed normed measures (LSAT & MCAT, respectively) than reference forms. No doubt, valid information in the real world setting is hard to come by, and small increments are better than no increments.

Indeed, of all the potential reference items (n=9, see Table 1) used by this particular doctoral program, only two items are found to differentiate between those rejected and those accepted when assessed by a linear combination via logistic regression analysis. Accounting for unique variance between these two groups (rejected and accepted) given all the reference items considered are perceptions about research potential and work habits. Collectively, research ability and work habits account for 14% of the variance between classification levels (rejected or accepted).

Actual contributions of these items are different is value (see regression weights) and direction (see signs). Controlling is research

potential over work habits (see Table 2). Applicants are more likely admitted if they are perceived to have high research ability, and high work habits can not offset this valued attribute within the admission process.

In view of the inter correlations among research ability and work habits with other reference information as noted in Table 1, these findings suggest that this particular doctoral program collects redundant information from reference sources. Because the acquisition of reference information is time consuming as well as costly, these data suggest that the reference form used by this program should be modified. More specifically, those reference items pertaining to intellectual ability, educational knowledge, motivation, maturity, problem solving, verbal ability and writing ability should be replaced in the current reference form used by this particular doctoral program.

Along this same line of reasoning, these data indicate that other doctoral programs in educational leadership should include, at minimum, within the reference acquisition process items assessing professional references' perceptions of a candidate's research ability and work habits. This recommendation is speculative, at best, because a random sample of programs was not employed in this study. However, other researchers have suggested that admission studies institutionally specific are more useful than admission studies conducted across institutions, at least for law schools (Ramsey, 1961).

Foremost, is that this study is based on "real" as opposed to experimental data examining the doctoral admission process (see Young et. al., 1991). Also, these data cover an entire decade of admission decisions involving multiple faculty and hundreds of reference sources as well as applicants. During the time period covered by this study, content of the reference form was constant for faculty, reference sources, and applicants. Finally, reference information for all applicants within a given year was considered collectively (fixed admission process) rather than sequentially throughout an academic year (variable admission process).

References

Creighton, T.B. & Jones, G.D. (August, 2001). *Selection of 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 , TX.

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.

Marcus, M.B. (1997, April). Half a doctor, *US News and World Reports.*

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

Petress, K.C. (1999). Letters of recommendation: The motive and content. *College Student Journal*, December.

Ramsey, R.R. (1961). Validity patterns at two law schools. *Educational and Psychological Measurements, 21*, 489-491.

Schmitt, N. & Chan, D. (1998). *Personnel selection: A theoretical approach.* Thousand Oaks, CA: Sage.

The Psychological Corporation (1994). *Technical manual for the Miller analogies test.* San Antonio, TX.

Young, I.P. (2005). *Predictive validity of specific academic indicators for a doctoral program focusing on educational leadership.* Paper accepted for presentation at the American Educational Research Association's Annual Meeting: Montréal, Canada.

Young, I.P., Galloway, C.M., & Rinehart, J. (1990). The effects of recruitment brochure content and gender of the reactor for doctoral programs in educational administration. *Educational Administration Quarterly, 26* (2), 168-182.

Young, J.W. (1995). A comparison of two methods for improving the prediction of law school grades. *Education and Psychological Measurement, 5*(4), 558-571.

Zeleznik, C., Hojat, h., Veloski, J. (1987). Predicative validity of MCAT as a function of undergraduate institution. *Journal of Medical Educational Education, 62*, 163-169.