DOCUMENT RESUME

ED 163 086 TM 008 227

AUTHOR	Rock, Donald A.
TITLE	The Identification of Population Moderators and Their Effect on the Prediction of Doctorate Attainment. GRE Board Professional Report GREB No. 69-6bP.
INSTITUTION	Educational Testing Service, Princeton, N.J.
PUB DATE	Feb 75
NOTE	41p.: Not available in hard copy due to print quality
AVAILABLE FROM	Graduate Record Examinations, Educational Testing Service, Princeton, New Jersey 08541 (free while supplies last)
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	Academic Achievement; *Achievement Tests; Age Differences; *Aptitude Tests; Chemistry; *College Entrance Examinations; *Doctoral Degrees; Doctoral Programs; Graduate Study; *Graduation; Higher Education; Mathematics; Predictive Ability (Testing); *Predictive Validity; Predictor Variables; Psychology; Sex Differences
IDENTIFIERS	*Graduate Record Examinations; Moderator Variables

ABSTRACT

The Graduate Record Examinations (GRE) Aptitude and Advanced Tests were evaluated as predictors of a dichotomous criterion of whether or not the candidate attained the doctorate within a specified length of time. More specifically, the project attempted to define subgroups for which the GRE tests have varying degrees of validity, and to provide biographical profiles of each subgroup as well as the optimal predictive equation for those subgroups. The GRE-Advanced tests were consistently found to be the best predictors of Ph.D. attainment. However, the predictive accuracy of the GRE-Advanced test varied considerably across graduate fields and in one case within a graduate field. That is, prediction on the whole was considerably more accurate in the "hard science" graduate areas of mathematics and chemistry than in psychology. Within the psychology area, there was a "U" shaped relationship between predictability and age. That is, the total sample regression equation led to greater predictive accuracy for the younger and the older age groups. The middle age group (25-26 years old) was not only less predictable, but the errors in prediction tended to lead to underestimation of their actual rate of Ph.D. attainment. (Author/ROF)

THE GRADUATE RECORD EXAMINATIONS

GRE

THE IDENTIFICATION OF POPULATION MODERATORS AND THEIR EFFECT ON THE PREDICTION OF DOCTORATE ATTAINMENT

Donald A. Rock

GRE Board Professional Report GREB No. 69-6bP

February 1975

This report presents the findings of a research project funded by and carried out under the auspices of the Graduate Record Examinations Board.

EDUCATIONAL TESTING SERVICE, PRINCETON, NEW JERSEY □ BERKELEY, CALIFORNIA □ EVANSTON, ILLINOIS

THE IDENTIFICATION OF POPULATION MODERATORS AND THEIR EFFECT ON THE PREDICTION OF DOCTORATE ATTAINMENT

Donald A. Rock

GRE Board Professional Report GREB No. 69-6bP

February 1975

THE IDENTIFICATION OF POPULATION MODERATORS AND THEIR EFFECT ON THE PREDICTION OF DOCTORATE ATTAINMENT

Donald A. Rock

Abstract

The immediate focus of this research project was to evaluate the potential of GRE aptitude and advanced tests as predictors of a dichotomous criterion of whether or not the candidate attained the doctorate within a specified length of time. More specifically the project attempted to: (1) define subgroups for which the GRE tests have varying degrees of validity, and (2) provide biographical profiles of each subgroup as well as the optimal predictive equation for those subgroups.

It was found that the GRE-advanced tests were consistently the best predictors of a criterion of Ph.D. attainment. However, the predictive accuracy of the GRE-advanced test varied considerably across graduate fields and in one case within a graduate field. That is, prediction on the whole was considerably more accurate in the "hard science" graduate areas of Mathematics and Chemistry than in Psychology. Within the psychology area there was a "U" shaped relationship between predictability and age. That is, the total sample regression equation led to greater predictive accuracy for the "younger" and the "older" age groups. The "middle" age group was not only less predictable but the errors in prediction tend to lead to underestimation of their actual rate of Ph.D. attainment. Thus, the "middle" age group was characterized by over-achievement.

THE IDENTIFICATION OF POPULATION MODERATORS AND THEIR EFFECT ON THE PREDICTION OF DOCTORATE ATTAINMENT[1]

Donald A. Rock

Introduction

Researchers seeking to demonstrate the validity of test scores for predicting graduate school performance have encountered a number of operational as well as logical difficulties. Reilly (1971) lists three major difficulties. These are: (1) the small samples usually available which in turn lead to unstable estimates of the parameters, (2) homogeneity of the sample itself due to previous selection with respect to ability and achievement variables, and (3) the establishment of an adequate criterion. Graduate grade point average (GPA), while being the more accessible criterion, has also been the most severely criticized. Lannholm et al. (1968), probably levels the most valid and serious criticism of GPA when he concludes that it represents only a limited aspect of graduate school performance. It is also subject to an understandable unwillingness on the part of faculty to discriminate among individuals all of whom are members of a highly selected population.

The most desirable criterion, of course, would be some measure of achievement as a scientist. Aside from the logical difficulties in arriving at any sort of agreement as to what is a relevant measure of scientific achievement, we are faced with the operational problem of time lapse which must occur before such data can be collected. An alternative criterion of a more intermediate nature is whether or not one has attained his or her doctorate within a reasonable period of time. Attainment versus non-attainment of the doctorate is appealing on logical grounds since: (1) it is one test of the effectiveness of the overall selection process i.e., the decision to admit a student to graduate education or to admit him to candidacy for a higher degree implies an expectation that his formal graduate education will be completed. The attainment of the doctorate degree is the primary indicator that such an expectation has been fulfilled; and (2) more often than not attainment of the doctorate is a necessary pre-requisite to gaining entry into the scientific-academic arena. From an operational viewpoint doctorate attainment is readily quantifiable and, of course, requires less time to mature than "on the job" measures of effectiveness. One criticism, however, is that it lacks sensitivity in the sense that it cannot take into account the various qualitative levels of performance among individuals attaining or not attaining the Ph.D. Although the latter criticism may well be valid, it was felt that the ease of quantification, and availability were sufficiently compelling reasons for its use in this study. It was also felt that if it was sufficiently lacking in sensitivity, this in turn would be reflected in the relative level of its predictability.

Creager (1965) examined the relationship between doctorate attainment and the GRE tests for NSF applicants and found modest but significant relationships between the tests and the criteria. Creager and Harmon (1966) report a study of the validity of both tests and quality ratings against a number of different criteria including doctorate attainment, publication counts, citation counts, on the job ratings, and a composite criterion. The report validities ranging from .11 to .39, the highest validities being for the composite criterion.

The immediate focus of this research project was to evaluate the potential of GRE aptitude and advanced tests as predictors of a dichotomous criterion of whether or not the candidate attained the doctorate within a specified length of time. More specifically, the project attempted to: (1) define subgroups for which the GRE tests have varying degrees of validity and (2) provide biographical profiles of each subgroup as well as the optimal predictive equation for those subgroups.

This initial or predictive phase outlined above was then supplemented by a second or "explanatory" phase. This second phase is essentially a model building process in which path analysis methods (Wright, 1934; Blalock, 1961) were used to trace the "causal" role of the GRE tests in the decision-making process at time of application as well as their direct and indirect impact on Ph.D. attainment. The use of path models rather than straight predictive models allows the model builder to include other salient variables in the model which may influence outcomes such as Ph.D. attainment but which would be operationally impractical in the more limited prediction model. For example, if we are interested in predicting Ph.D. attainment for any given individual at a particular university, it is the individual's characteristics which we would focus on and not environmental characteristics of that particular university. In this type of prediction model it is, of course, assumed that environmental characteristics are constant for all individuals in a department at a particular university.

However, in the interpretive or explanatory path model used in this study, we attempted to estimate the direct and indirect effects of the aptitude-achievement measures on Ph.D. attainment when analyzed in conjunction with graduate department quality indices (Cartter, 1964) and relevant variables available on the applicant.

Methodology

Approximately 1,000 NSF applicant records in the areas of Psychology, Chemistry, and Mathematics were collected from the merging of the Doctorate Records File and the Office of Scientific Personnel (OSP) tape of the National Science Foundation applicants. These file records indicated time to Ph.D. and, of course, whether or not they received one. Additional biographical information available in the Doctorate Records File and on the OSP tape included sex, age, marital status, number of dependents, number of NSF applications made, and awards received. The OSP records also provided Office of Education codes for the institution each applicant had chosen for graduate study.

Predictor information available from the OSP records included the GRE tests scores--verbal, quantitative, and advanced--as well as undergraduate grade point average and reference report average. The reference report average (Harmon, 1966) is a quantification of an overall rating of the reference letters submitted in behalf of an NSF applicant. The Doctorate Records File had additional information on father's and mother's education, but was too incomplete to be useful for this study.

The criterion of doctorate attainment required a judgment to be made concerning the time lapse to be allowed before assigning an individual to the attainment versus non-attainment category. It is, of course, rare that one completes a doctorate within three years after the baccalaureate. In the science fields the mean time lapse is approximately eight years (Creager, 1965), with greater deviations above the mean than below. If time were allowed for almost everyone to complete a doctorate, the study might well suffer on both operational as well as rational grounds. That is, not only would more of the people attaining doctorates have more time out and extensive study time (thus complicating the interpretation), but more persons of low measured ability would have achieved a doctorate under possibly lower standards of dissertation and course quality. From the viewpoint of efficient use of resources as well as cost of graduate education, it would seem to be desirable to select those individuals capable of successfully finishing the program in a reasonable amount of time. Conversely, too short a time lapse would eliminate many high quality people, possibly those very able persons who take on more ambitious dissertation projects and/or more difficult course offerings.

These considerations lead, for criterion definition purposes, to setting limited cutoff times for doctorate completion. The doctorate completion cutoff was June 1968. Since most of the subjects included applied for first-year fellowships in 1958-1961, they had seven to ten years from fellowship application time to attain their doctorate.

The matching and merging of the OSP and Doctorate Records File was completed at the Office of Scientific Personnel. The merged tape was then sent to the Educational Testing Service for analysis. Before analysis, additional information was compiled on the characteristics of the institution chosen for graduate study. Cartter's (1964) report on the quality of graduate departments furnished the quality indices which were then assigned to each candidate according to the ranking of the department which he attended. Additional institutional quality information was also collected from an Office of Education tape which included such information as: (1) proportion of faculty with doctorate, (2) per student expenditure, (3) number of books in the library, (4) income per student, and (5) faculty-student ratios. These particular "quality indices" suffer from the fact that they apply to the total institution and thus are not necessarily an accurate picture of the graduate school or, more specifically, the graduate department itself.

Within each major field, the sample was split into two random halves for validation and cross-validation purposes. The data was then analyzed using the moderated regression technique (Rock et al., 1967). This technique not only furnishes the researcher with the usual multiple regression validity information, but also searches for consistent biographical patterns associated with "types" of individuals who, in turn, are characterized by varying levels of predictability. For example, this type of analysis enables one to determine if any one subgroup—such as older NSF applicants—should have a different prediction equation than another subgroup. The moderated regression technique allows the researcher to hypothesize up to five moderators or grouping variables at one time. First it will form subgroups on each moderator singly and then it will form groups based on similar profiles based on combinations of moderators. For example, if the moderated regression technique were used with two possible moderators such as age and department quality index, it might identify a group of older individuals attending low quality graduate departments who are unpredictable with respect to Ph.D. attainment when GRE-test scores were used as predictors. Since this technique requires complete information, the sample sizes were reduced to 779[a], 845; and 643, for Psychology, Mathematics, and Chemistry, respectively.

Within each graduate discipline the moderated regression was run first on the validation sample. This analysis led to subgroups characterized by differential predictive accuracy.

Potential moderators which were analyzed with respect to their impact on accuracy of prediction were age, sex, marital status, university quality indices, and graduate department quality indices. Students attending the same institution were assigned that

---

[a] Analyses in Psychology were based on a total of 930 observations when department quality indices were not part of the analysis. This was due to the fact that a substantial number of cases had to be dropped when the quality indices were included.

particular institution's quality ratings as well as department ratings. Those moderators which grouped individuals into subgroups on the validation sample (which in turn were characterized by differential prediction) were selected for replication on the cross-validation sample. That is, groups were formed within the cross-validation sample which had similar profiles to the differentially predictable groups from the first sample. Then the appropriate group as well as total equations were applied to these corresponding groups in the cross-validation sample in an effort to determine if differential predictability was a stable characteristic of these various subgroups.

Results and Discussion

Psychology

Sex had little or no effect as a moderator, but proved to be a good predictor of Ph.D. attainment. As indicated in Table 1, sex has

---

Insert Table 1 about here

---

the highest single variable correlation with the criterion (-.45 in Sample 1 and -.34 in Sample 2) among all the potential predictors or moderators. The negative sign indicated that women are less likely to attain their doctorate in Psychology than are men. Further inspection of Table 1 indicates that the test variables (GRE-Verbal, Quantitative, and Advanced), one college quality index (department rating), reference average, and number of NSF applications have consistent (non-zero in both samples) relationships with the criterion. The department rating's relationship with the criterion carries a negative sign, since the quality code indices range from one to four, with one signifying the highest quality and four the poorest quality. The remaining institution quality indices appear to be too general and thus do not necessarily reflect the quality of the Psychology departments. The correlations between the department rating and the college quality indices range from a low of -.13 for percentage of faculty with the Ph.D., to a high of -.51 for income per student, indicating a large proportion of the variance in the department rating is not accounted for by the more general quality indices.

The relatively high correlation between the number of NSF applications made and Ph.D. attainment is somewhat artifactual, since a large percentage of the NSF applicants in this study were required to reapply for their grant every year. Many of those students who did not reapply may have either dropped out of the program or possibly felt that their past performance record would not be supportive of a grant extension. Thus applications made may be considered an intermediate progress report on the way to the Ph.D. in Psychology.

Among the predictor tests, Advanced Psychology appears to have the more consistent relationships with the criterion when considered across both samples. Undergraduate grade point average had a surprisingly low predictive validity with respect to Ph.D. attainment.

Insert Table 2 about here

Table 2 shows the results of the moderated regression analysis on Sample 1 and its cross-validation on Sample 2. The total column of Table 1 indicates the multiple correlations between the five predictors (GRE-Verbal, Quantitative, and Advanced, undergraduate grade point average, and reference report average) and the dichotomous criterion of whether or not they achieved their doctorate. Similar interpretation applies to the remaining columns, except that within each group there are now two correlations. The correlation based on the group equation indicates the strength of the relationship between the criterion and the five predictors when the weights for the predictors are based only on information unique to that group. This, of course, is the optimal prediction equation for that group. The total equation correlations are also presented within each group. The weights for the predictors here are, of course, based on the total sample and in general will not be optimal for any particular subgroup. They are of particular interest since they indicate the differential predictive accuracy that might occur within a subgroup when an overall equation is used, which is the normal procedure in a selection situation. Considering the restricted nature of the sample, i.e., NSF applicants, it would seem that the overall multiple of .33 in Sample 1 is quite respectable.

Of the biographical data for the Psychology students, only age level led to a consistent pattern of differential predictability; that is, the pattern from Sample 1 was replicated in Sample 2. It is interesting to note that there is a "U" shaped relationship between age and predictability. That is, the relatively young and the relatively older groups were considerably more predictable that the 25- and 26-year old applicants. However, the oldest group appears to be the one group we can most accurately predict, whether using the total or group equation. Although the oldest group appears to be the most predictable, they have the smaller probability of getting their doctorate. That is, while almost fifty percent of each of the remaining age groups did obtain their doctorate within the specified time, only twenty-eight percent of the older group did likewise.

Inspection of the means on Table 3 for the predictor scores for

---

Insert Table 3 about here

---

the various age groups indicated that both the older and the "middle" age groups had similar means, both of which are consistently lower than the youngest group. Thus, the "middle" age group, although the least predictable group, is similar to the most predictable group (the oldest group) with respect to ability as measured by test scores. Surprisingly, however, the "middle" age group (25- and 26-year olds) consistently produces a greater proportion of doctorates than either the younger or the older groups. The "middle group" is also characterized by a lower undergraduate grade point average than either of the remaining predictable groups. Since the "middle" age group tends to have lower predictor scores on the average, yet possesses the highest level of Ph.D. attainment, they are generally underpredicted when the overall equation is used. Thus they are what is commonly referred to as overachievers in the psychometric literature.

It may well be that the 25- and 26-year olds have overcome their somewhat mediocre ability-achievement credentials by a higher level of motivation and consequently have a higher rate of Ph.D. attainment. Unfortunately, we do not have the data to determine what, if any, other age-related characteristics are operating here. These findings of differential predictive accuracy as well as possible motivational differences point out the need for more biographical information about graduate applicants if we are to understand and/or infer the causal pattern underlying their differential performance.

Table 4 presents the results when the moderated regression was

---

Insert Table 4 about here

---

used to search for patterns among the institutional quality indices which might lead to consistent patterns of predictability and/or unpredictability. Unfortunately, since the moderated regression requires complete data, a number of cases had to be dropped from this part of the analysis because they were missing one or more institutional quality indices.

As in the case of the biographical variables, only one quality index to a replicable pattern of differential predictability. That is, there is a slight but seemingly consistent tendency for students attending "low quality" Psychology departments to be more predictable. This result certainly comes as no surprise, since the so-called "higher quality" schools are more selective of applicants with respect to the GRE test scores and thus attainment of the Ph.D. is likely to depend on some u[illegible] measured quality. It is, however, interesting to note that at the "lo[illegible] quality" Psychology departments, the probability of obtaining the Ph.D. is consistently less than at the "higher quality" departments.

Table 5 presents results when grouping was done on both age and the

---

Insert Table 5 about here

---

departmental quality index. Although the pattern of predictability is less clear cut, there remains a tendency for the older students attending "lower quality" institutions to be more predictable. In this four-way break-out, the sample sizes are rather small and the resulting instability of the parameter estimates makes any further interpretations of these results rather tenuous.

In order to determine the utility of age and departmental quality as potential predictors, they were incorporated into stepwise prediction equations in both their linear form and as bi-linear cross products with the remaining predictors. In no form did they consistently lead to an increment in prediction over the original five predictors (GRE-Verbal, Quantitative, and Advanced, UGGPA, and reference report average).

It would appear that for NSF applicants in Psychology, the utility of age information lies primarily in separating out those individuals for whom: (1) we have varying degrees of confidence in their predicted or expected achievement (in this case, Ph.D. attainment), and (2) motivational levels may differ.

The results also suggest that where there was differential prediction, the overall equation used within the groups was not noticeably inferior to the unique group equation with respect to predictive accuracy. This suggests that different weightings of the same predictor variables for different types of people (older versus younger, etc.) does not appear to be the answer. That is, some individuals seemed to be more or less predictable regardless of whether you use overall weights or their own unique weights. It is possible that entirely different predictor measures must be developed for the "unpredictable" people. This, of course, is beyond the scope of this study.

Mathematics

Table 6 presents the single variable validity coefficients for the

---

Insert Table 6 about here

---

predictors and potential moderators or grouping variables. In general it appears that the criterion of Ph.D. attainment in Mathematics is considerably more predictable from achievement-aptitude measures than was found to be the case in Psychology. Of particular interest in Table 6 are the correlations of .38 and .44 for the Advanced Mathematics test against the criterion for Sample 2 and 2, respectively. The GRE-Verbal and Quantitative, as well as undergraduate grade point average, have respectable although lower relationships with the criterion. Institutional quality indices such as income per student, faculty per student, and departmental quality index, also demonstrate stronger relationships with Ph.D. attainment in Mathematics. It may well be that the successful completion of the Ph.D. program in Mathematics depends upon the assimilation of a relatively structured body of knowledge which in turn leads to more accurate assessments of any one individual's standing with respect to this body of knowledge.

Table 7 shows the differential prediction by age groups. The

Insert Table 7 about here

---

multiple correlation between the five predictors (GRE-Verbal, Quantitative, and Advanced, undergraduate grade point average, and reference report average) and Ph.D. attainment is a quite respectable .40 in Sample 1 and cross-validates to a surprising .44 in Sample 2. Further inspection of Table 7 indicates that there is little or no consistent differential prediction by age group. Unlike Psychology, there is a linear relationship between age and Ph.D. attainment. That is, the older the NSF applicant, the less likely he is to attain his doctorate within the cutoff time of this study. As in Psychology, the "middle" and "older" NSF applicants had similar aptitude-achievement test scores, and when considered as a whole had consistently lower test scores than the younger candidates. The one exception to the above findings was the Advanced test, where the "older" NSF candidates were not only lower than the younger candidates, but were also one-half standard deviation below the "middle" age candidates.

Table 8 presents the multiple correlations within groups based on

---

Insert Table 8 about here

---

departmental quality indices. As with age, there does not seem to be any consistent pattern of differential predictability.

Table 9 shows the multiple correlations within groups based on the

---

Insert Table 9 about here

---

departmental quality index and age. With one exception, there appears to be little differential predictability within these groups. Somewhat surprisingly, the young who attend "lower quality" departments appear to be characterized by greater predictability than the remaining groups. In general, the mean ability-achievement scores for this group are below that of both the "high quality" young and the "high quality" old, but slightly above those of the "low quality" old group. Because the "low quality" young group size is so small, any further interpretation is probably unwarranted. As one would expect, the criterion means in Table 9 indicate that the young applicants who attend institutions with "high quality" departments are much more likely to attain the doctorate than are the older NSF candidates who attend institutions characterized by "low quality" Mathematics departments. When age was included as a predictor, no increment was found in predictive accuracy above that which resulted from the use of the original five predictors.

Chemistry

The single variable validity coefficients for the chemistry measures are similar both in level and pattern to those of the Mathematics NSF applicants. As in Mathematics, the GRE-Advanced test is the one best pre-

---

Insert Table 10 about here

---

dictor in both samples. However, among the Chemistry NSF applicants, undergraduate average, reference report average, and age demonstrate somewhat higher relationships with Ph.D. attainment than do their counterparts for the Mathematics applicants. In general, the level of correlations found in Chemistry yield additional support for the hypothesis that the so-called "hard sciences" may provide a more measurable domain with respect to criteria of success as well as measures of past achievements or aptitudes. It is also quite possible that it is easier to specify the necessary skills which are prerequesite to success (Ph.D. attainment, in this case) in the "hard sciences."

Table 11 presents the differential prediction results by age groups

---

Insert Table 11 about here

---

for the Chemistry NSF applicants. As in the case of the Mathematics applicants, there appears to be little consistent differential prediction. Prediction for Sample 1 is relatively strong considering the somewhat restricted nature of the sample. Surprisingly enough, the cross-validated multiple correlations increased from .39 in Sample 1 to .53 in Sample 2.

A considerably larger proportion of the NSF applicants in Chemistry do attain their Ph.D. than the Psychology and Mathematics NSF applicants.

Table 12 presents the differential prediction within groups based on

Insert Table 12 about here

the rated "quality" of their Chemistry departments. As was the case with age, there is little or no consistent differential prediction. Table 13 shows the within group multiple correlations when both quality indices

Insert Table 13 about here

and age are used as moderators. Once again, no consistent pattern of differential predictive accuracy was evident. It appears that in the two "hard science" areas of Mathematics and Chemistry, the assimilation of knowledge in their particular area as measured by the Advanced section of the GRE is the one best predictor of Ph.D. attainment, regardless of age group membership or quality of the institution of attendance.

Age was included as a predictor and unlike Mathematics or Psychology, it did add significantly to the prediction. It was the second variable after the GRE-Advanced section to enter the equation. In an effort to gain some insight into this relationship, the correlation between age and whether or not the student attended on a part-time basis was examined. This correlation was effectively zero (.02). Thus, the "older" students in Chemistry are no more likely to attend on a part-time basis than the other age groups.

The significant partial regression weight associated with age indicates that after the ability-achievement variables were controlled, there remained a significant amount of variance in age which was related to Ph.D. attainment. It would appear that additional biographical information might prove helpful in untangling this relationship.

Path Models

In an effort to interpret the interrelationships among the multiple predictors and departmental quality ratings, as well as to estimate the relative size of their direct and indirect effects on Ph.D. attainment, path models were analyzed. The general implication in path analysis is that verbal theories concerning the relationships between certain variables can be translated into "causal" models involving mathematical equations, if the direction of causality can be specified. In path analysis, the direction of causality between any two or more variables may be specified on logical grounds or on the basis of temporal sequence. The "causal" effect of one variable or another may have two estimable components: the direct effect (i.e., unmediated by any third intervening variable) and the indirect effect which, of course, must take into consideration intervening variables. In general, when data is collected longitudinally, many of the measures because of their temporal nature make the direction of causality explicit.

Within the three graduate disciplines, the same hypothetical path model was developed and tested. Figure 1 presents the traditional pictorial presentation of the model to be tested. The arrows going in

---

Insert Figure 1 about here

---

one direction specify the direction of causality. Arrows between two variables going in both directions signify that the direction of "causality" could not be determined on rational or temporal grounds.

The path coefficients b_{ij} are standardized partial regression coefficients and are the unknowns to be estimated. The relative size of any given weight (b_{ij}) may be interpreted as the relative direct influence of the i^{th} variable on the j^{th} variable. "Direct" in the sense that it is defined as that influence which remains after all other independent variables in the causal equation have been controlled. For example, if b_{y2} is twice as large as b_{y3}, then quality rating of the department is twice as important an estimator of Ph.D. attainment as scores on the GRE-Advanced test in determining whether or not one attains the Ph.D. If direction of causality is not known, the path coefficients are replaced by the simple correlation coefficient.

Inspection of Figure 2 shows the computed Psychology path coeffi-

---

Insert Figure 2 about here

---

cients indicating the "direct" contributions of the variables to the various dependent variables in the model. For example, the path coefficient symbolizing the influence of undergraduate grade point average (GPA) on the reference report average, suggest that UGGPA is almost twice as important as the individual's GRE-Advanced score in influencing the ratings derived from the reference letters. This suggests that the reference letters are often written without knowledge of the GRE-Advanced scores of the NSF applicant. Although UGGPA has a greater influence on decision-making concerning reference ratings, the GRE-Advanced test scores have by far the largest single influence on Ph.D. attainment.

The correlation coefficients indicating the relationships between both the GRE-Advanced test scores and UGGPA on the departmental quality ratings, indicate that the GRE-Advanced test scores of the students within the departments have a considerably stronger relationship with the quality rating than did grades in undergraduate school. At this point it might be instructive to point out that the arrows (indicating direction of influence) were drawn in both directions; that is, quality of department may influence both UGGPA and GRE-Advanced test scores or vice versa. Influence in one case may be interpreted as that good quality departments tend to attract or select students with both high GRE scores and undergraduate grade point averages. Another interpretation might be that the department rating is a function of the abilities of the students who choose to attend that particular institution. The interpretation chosen here is that high quality departments put a greater emphasis on either attracting or selecting students with high test scores than they do on attracting students with high grade point averages.

Figures 3 and 4 give the path coefficients for the Mathematics and

Insert Figures 3 and 4 about here

Chemistry data, respectively, and yield essentially the same results as the Psychology data. As one might expect from the discussion of the prediction results, the GRE-Advanced test scores reflecting knowledge of subject area have a greater influence on Ph.D. attainment in the "hard sciences" than in Psychology. This is particularly true in the Mathematics area where the importance of the GRE-Advanced test is approximately five times that of the other predictors. A less obvious difference in the relative importance of effects occurs between Chemistry and the remaining two "disciplines." That is, in Chemistry, undergraduate grade point average was both more important in influencing decisions with regard to reference letters, but also had a greater influence on Ph.D. attainment than in the other two disciplines.

Figure 5 presents the reduced path model, i.e., the pictorial

Insert Figure 5 about here

representation of only those "causal" effects which had an average path coefficient exceeding .10. Average is defined here as simply the arithmetic mean of corresponding path coefficients across the three graduate disciplines. The criterion for deleting effects not exceeding .10 is purely arbitrary but seemed to be a reasonable cut-off for this study. Werts (1967) also suggests this cut-off as a practical criterion for eliminating negligible effects.

The reduced model makes clear the central nature of the GRE-Advanced test with respect to the prediction of Ph.D. attainment, as well as its influence on earlier decisions such as reference letter averages and finally quality ratings. The GRE-Advanced test was selected as the one test to be included in the path models, since it consistently had the highest 0-order correlation with the criterion among the test scores. In general, path models are more interpretable when only one measure is used to represent any given measurable domain, since the inclusion of many similar measures (measures of the same domain) lead to high co-linearities and thus unstable estimates of the path coefficients.

Conclusions

It was found that the GRE-Advanced tests were consistently the best predictors of a criterion of Ph.D. attainment. However, the predictive accuracy of the GRE-Advanced test varied considerably across graduate fields and in one case within a graduate field. That is, prediction on the whole was considerably more accurate in the "hard science" graduate areas of Mathematics and Chemistry than in Psychology. Within the Psychology area there was a "U" shaped relationship between predictability and age. That is, the total sample regression equation led to greater predictive accuracy for the "younger" and the "older" age groups. The "middle" age group was not only less predictable but the errors in prediction tend to lead to underestimation of their actual rate of Ph.D. attainment. Thus, the "middle" age group was characterized by overachievement.

Path analysis models were applied to the intercorrelation matrices in an effort to determine the effects of each of the predictor variables on the criterion as well as on other intermediate decision-making processes. The resulting path coefficients supported the notion that the GRE-Advanced test was measuring a domain of knowledge which was critical to Ph.D. attainment within the specified time limit.

Footnote

1. The author wishes to thank Dr. Lindsey R. Harman for his assistance in making the data available for this study and his helpful suggestions during the early part of the data analysis.

References

Blalock, H. M., Jr. Causal inferences in nonexperimental research. Chapel Hill: University of North Carolina Press, 1961.

Cartter, A. M. (Ed.) American universities and colleges. (9th ed.) Washington, D. C.: American Council on Education, 1964.

Creager, J. A. Predicting doctorate attainment with GRE and other variables. Technical Report No. 25, 16 November 1965. Research on Fellowship Selection Techniques Supported by the National Science Foundation. Office of Scientific Personnel, National Academy of Sciences--National Research Council, Washington, D. C.

Creager, J. A., & Harmon, L. R. On-the-job validation of selection variables, April 15, 1966. (Multilithed report, 70 pp.)

Fourteen years of research on fellowship selection. A summary. Lindsey R. Harmon. Progress report on research supported by the National Science Foundation. July 20, 1966. Office of Scientific Personnel, National Academy of Sciences, National Research Council. Publication 1420. Washington: National Academy of Sciences, National Research Council, 1966.

Lannholm, G. V., Marco, G. L., & Schrader, W. B. Cooperative studies of predicting graduate school success. Graduate Record Examinations Special Report 68-3. Princeton, New Jersey: Educational Testing Service, 1968.

Reilly, R. R. Critical incidents of graduate student performance. Graduate Record Examinations Board Technical Memorandum No. 1. Princeton, New Jersey: Educational Testing Service, 1971.

Rock, D. A., Barone, J., & Linn, R. L. A moderated stepwise prediction system. Educational and Psychological Measurement, 1967, 27, 709-713.

Werts, C. E. The study of college environments using path analysis. National Merit Scholarship Corporation, Research Reports, Vol: 3, No. 4, 1967.

Wright, S. The method of path coefficients. Annals of Mathematical Statistics, 1934, 5, 161-215.

TABLE 1

VALIDITY COEFFICIENTS, MEANS, AND STANDARD DEVIATIONS OF THE PREDICTORS FOR SAMPLES 1 AND 2 (PSYCHOLOGY)

	Sample 1 N=380			Sample 2 N=398		
	r	$\bar{X}$	σ	r	$\bar{X}$	σ
Criterion	-	1.40	.49	-	1.39	.49
Sex	-.45*	1.41*	.49*	-.34	1.34	.48
Age	-.05	23.37	3.70	-.03	23.68	4.15
No. of Books	.06	9.87	.57	.02	9.85	.67
Income/Student	.15	8.75	1.77	.05	8.71	1.97
Faculty/Student	.13	6.36	3.04	.07	6.45	2.98
Percent with Ph.D.	-.02	7.19	3.06	.01	7.13	3.11
Dept. Rating	-.14	1.67	1.14	-.16	1.66	1.29
GRE-V	.12	63.52	8.25	.19	63.47	9.29
GRE-Q	.33	59.89	11.34	.14	60.96	10.82
GRE-Adv.	.19	60.98	8.90	.24	60.87	9.05
Ref. Average	.16	43.86	8.36	.14	43.93	8.49
UGGPA	.02*	241.70*	40.10*	.02	236.78	42.96
App. Made	.31*	1.63*	1.11*	.33	1.64	1.10

* N=462

TABLE 2

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON AGE

(PSYCHOLOGY)

		1 Younger	2 Middle	3 Older	Total
		N=355	N=61	N=46	N=462
	Within-Group Multiple R's				
PRIMARY	Group Equations	.33	.33	.47	
	Total Equations	.31	.27	.43	.33
SAMPLE	Moderator Means (Age)	21.82	25.48	32.54	
	Criterion Means (1=No PhD; 2=PhD)	1.41	1.52	1.28	1.41
		N=354	N=77	N=37	N=468
	Within-Group Multiple R's				
	Group Equations	.23	.15	.25	
CROSS-	Total Equations[a]	.27	.14	.30	.24
VALIDATION SAMPLE	Moderator Means (Age)	21.87	25.69	35.00	
	Criterion Means (1=No PhD; 2=PhD)	1.40	1.46	1.32	1.41

[a] Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

TABLE 3

PREDICTOR MEANS AND STANDARD DEVIATIONS FOR THE PSYCHOLOGY AGE GROUPS

Sample 1

Predictors	Group 1 N=355		Group 2 N=61		Group 3 N=46	
	$\bar{X}$	σ	$\bar{X}$	σ	$\bar{X}$	σ
GRE-V	64.17	8.45	62.41	7.84	62.52	9.36
GRE-Q	61.17	11.17	57.44	9.64	56.78	11.48
GRE-Adv.	62.16	8.36	58.56	8.48	57.76	10.39
Reference Reports	44.33	8.15	44.34	8.70	43.09	8.89
UGGPA	242.67	39.17	236.18	44.21	241.43	42.41
Moderator Mean (Age)	21.82		25.48		32.54	

Sample 2

Predictors	Group 1 N=354		Group 2 N=77		Group 3 N=37	
	$\bar{X}$	σ	$\bar{X}$	σ	$\bar{X}$	σ
GRE-V	64.90	9.04	60.19	9.19	61.[illegible]	9.62
GRE-Q	62.23	10.48	58.26	10.28	56.22	11.74
GRE-Adv.	62.15	8.33	58.52	9.69	56.38	10.34
Reference Reports	44.52	8.59	41.77	9.56	40.16	8.53
UGGPA	240.69	41.44	225.92	45.23	234.08	45.86
Moderator Mean (Age)	21.87		25.68		35.00	

TABLE 4

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON DEPARTMENT QUALITY INDICES (PSYCHOLOGY)

		1 Hi-Quality	2 Med-Quality	3 Low-Quality	Total
PRIMARY SAMPLE		N=271	N=49	N=60	N=380
	Within-Group Multiple R's				
	Group Equations	.33	.55	.34	
	Total Equations	.32	.50	.27	.35
	Moderator Means (Quality Index)	1.00	2.54	4.00	1.68
	Criterion Means (1=No PhD; 2=PhD)	1.44	1.40	1.26	1.40
CROSS-VALIDATION SAMPLE		N=282	N=56	N=60	N=398
	Within-Group Multiple R's				
	Group Equations	.12	.06	.22	
	Total Equations[a]	.14	-.15	.31	.15
	Moderator Means (Quality Index)	1.00	2.52	4.00	1.67
	Criterion Means (1=No PhD; 2=PhD)	1.44	1.32	1.26	1.39

[a]Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

TABLE 5

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON DEPARTMENT QUALITY INDICES AND STUDENT AGE

(PSYCHOLOGY)

		1 Hi-Quality Young	2 Low-Quality Somewhat Older	3 Low-Quality Young	4 Hi-Quality Older	Total
		N=266	N=39	N=52	N=23	N=380
	Within-Group Multiple R's					
PRIMARY	Group Equations	.33	.50	.48	.47	
	Total Equations	.32	.38	.25	.43	.35
SAMPLE	Moderator Means					
	(Quality Index)	1.07	3.69	3.56	-1.04	1.68
	(Age)	1.12	2.41	1.00	3.00	1.35
	Criterion Means (1=No PhD; 2=PhD)	1.45	1.36	1.21	1.33	1.40
		N=290	N=40	N=49	N=19	N=398
	Within-Group Multiple R's					
	Group Equations	.12	.22	.10	-.02	
SECONDARY	Total Equations[a]	.12	.27	.05	.05	.15
	Moderator Means					
	(Quality Index)	1.07	3.77	3.59	1.26	1.66
SAMPLE	(Age)	1.14	2.42	1.00	3.00	1.34
	Criterion Means (1=No PhD; 2=PhD)	1.43	1.42	1.22	1.21	1.39

[a]Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

TABLE 6

VALIDITY COEFFICIENTS, MEANS, AND STANDARD DEVIATIONS OF THE PREDICTORS FOR SAMPLES 1 AND 2 (MATHEMATICS)

	Sample 1 N=423			Sample 2 N=422		
	r	$\bar{X}$	σ	r	$\bar{X}$	σ
Criterion	-	1.54	.50	-	1.54	.50
Age	-.17	22.35	2.06	-.18	22.71	2.50
No. of Books	.04	9.92	.40	.10	9.90	.48
Income/Student	.20	9.35	1.35	.17	9.24	1.53
Faculty/Student	.15	7.98	2.67	.13	7.57	2.88
Percent with Ph.D.	.06	7.55	2.88	.13	7.61	2.86
Dept. Rating	-.18	1.60	1.09	-.29	1.65	1.13
GRE-V	.27	62.95	10.96	.32	62.63	11.33
GRE-Q	.27	72.67	9.51	.26	71.54	10.14
GRE-Adv.	.38	65.93	15.39	.44	64.93	15.94
Ref. Average	.23	42.60	9.38	.27	42.59	9.69
UGGPA	.21	252.60	40.22	.24	248.77	43.13
App. Made	.52	2.32	1.52	.50	2.37	1.61

TABLE 7

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON AGE

(MATHEMATICS)

		1 Younger	2 Middle	3 Older	Total
		N=358	N=31	N=34	N=423
	Within-Group Multiple R's				
PRIMARY	Group Equations	.35	.66	.48	
	Total Equations	.35	.50	.38	.40
SAMPLE	Moderator Means (Age)	21.63	24.45	27.59	22.35
	Criterion Means (1=No PhD; 2=PhD)	1.58	1.41	1.32	1.54
		N=325	N=58	N=39	N=422
	Within-Group Multiple R's				
	Group Equations	.43	.35	.28	
CROSS-	Total Equations[a]	.43	.44	.39	.44
VALIDATION	Moderator Means (Age)	21.60	24.84	28.74	22.71
SAMPLE	Criterion Means (1=No PhD; 2=PhD)	1.59	1.43	1.38	1.54

[a]Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

TABLE 8

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON DEPARTMENT QUALITY INDICES

(MATHEMATICS)

		1 Hi-Quality	2 Med-Quality	3 Low-Quality	Total
PRIMARY SAMPLE		N=317	N=52	N=54	N=423
	Within-Group Multiple R's				
	Group Equations	.34	.37	.60	
	Total Equations	.34	.19	.59	.40
	Moderator Means (Quality Index)	1.00	2.75	4.00	1.60
	Criterion Means (1=No PhD; 2=PhD)	1.61	1.40	1.33	1.54
CROSS-VALIDATION SAMPLE		N=308	N=49	N=65	N=422
	Within-Group Multiple R's				
	Group Equations	.37	.34	.26	
	Total Equations[a]	.39	.45	.24	.44
	Moderator Means (Quality Index)	1.00	2.61	4.00	1.65
	Criterion Means (1=No PhD; 2=PhD)	1.62	1.40	1.29	1.54

[a]Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

TABLE 9

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON DEPARTMENT QUALITY INDICES AND AGE

(MATHEMATICS)

		1	2	3	4	Total
		Hi-Quality Young	Low-Quality Young	Hi-Quality Old	Low-Quality Old	
		N=106	N=24	N=211	N=82	N=423
PRIMARY SAMPLE	Within-Group Multiple R's					
	Group Equations	.20	.52	.40	.43	
	Total Equations	.17	.38	.39	.40	.40
	Moderator Means					
	(Quality Index)	1.00	3.54	1.00	3.34	1.60
	(Age)	20.64	20.88	22.91	23.56	22.35
	Criterion Means					
	(1=No PhD; 2=PhD)	1.69	1.42	1.57	1.35	1.55
		N=114	N=17	N=194	N=97	N=422
SECONDARY SAMPLE	Within-Group Multiple R's					
	Group Equations	.31	.73	.36	.29	
	Total Equations[a]	.35	.65	.36	.31	.44
	Moderator Means					
	(Quality Index)	1.00	3.29	1.00	3.42	1.65
	(Age)	20.78	20.65	23.07	24.61	22.71
	Criterion Means					
	(1=No PhD; 2=PhD)	1.75	1.53	1.56	1.31	1.55

[a]Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

TABLE 10

VALIDITY COEFFICIENTS, MEANS, AND STANDARD DEVIATIONS OF THE PREDICTORS FOR SAMPLES 1 AND 2 (CHEMISTRY)

	Sample 1 N=322			Sample 2 N=321		
	r	$\bar{X}$	σ	r	$\bar{X}$	σ
Criterion	-	1.75	.43	-	1.74	.44
Age	-.29	22.10	1.81	-.28	22.26	1.82
No. of Books	.08	9.82	.59	.01	9.84	.55
Income/Student	.12	9.44	1.12	.14	9.24	1.34
Faculty/Student	.08	7.89	2.66	.04	7.37	2.89
Percent with Ph.D.	-.05	7.11	2.99	.12	7.07	3.12
Dept. Rating	-.14	1.36	.79	-.22	1.39	.86
GRE-V	.15	59.66	10.69	.23	58.40	10.75
GRE-Q	.28	69.26	10.70	.34	67.96	10.70
GRE-Adv.	.33	67.41	11.81	.48	66.27	12.31
Ref. Average	.30	41.48	9.89	.33	42.23	9.15
UGGPA	.27	246.93	44.55	.36	247.93	42.82
App. Made	.42	2.43	1.48	.39	2.34	1.48

TABLE 11

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON AGE

(CHEMISTRY)

		1 Younger	2 Middle	3 Older	Total
		N=114	N=176	N=32	N=322
PRIMARY SAMPLE	Within-Group Multiple R's				
	Group Equations	.29	.47	.21	
	Total Equations	.28	.46	.03	.39
	Moderator Means (Age)	20.78	22.16	26.44	22.10
	Criterion Means (1=No PhD; 2=PhD)	1.82	1.75	1.47	1.75
		N=93	N=188	N=40	N=321
CROSS-VALIDATION SAMPLE	Within-Group Multiple R's				
	Group Equations	.28	.57	.15	
	Total Equations[a]	.31	.58	.55	.53
	Moderator Means (Age)	20.78	22.13	26.30	22.26
	Criterion Means (1=No PhD; 2=PhD)	1.78	1.79	1.40	1.74

[a]Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

TABLE 12

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON DEPARTMENT QUALITY INDICES

(CHEMISTRY)

		1 Hi-Quality	2 Med-Quality	3 Low-Quality	Total
PRIMARY SAMPLE		N=253	N=52	N=17	N=322
	Within-Group Multiple R's				
	Group Equations	.39	.44	.35	
	Total Equations	.38	.38	.27	.39
	Moderator Means (Quality Index)	1.00	2.25	4.00	1.36
	Criterion Means (1=No PhD; 2=PhD)	1.77	1.69	1.53	1.74
CROSS-VALIDATION SAMPLE		N=254	N=44	N=23	N=321
	Within-Group Multiple R's				
	Group Equations	.49	.47	.03	
	Total Equations[a]	.50	.50	.63	.53
	Moderator Means (Quality Index)	1.00	2.27	4.00	1.39
	Criterion Means (1=No PhD; 2=PhD)	1.77	1.70	1.43	1.74

[a]Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

TABLE 13

DIFFERENTIAL PREDICTION WITHIN GROUPS BASED ON DEPARTMENT QUALITY INDICES AND AGE

(CHEMISTRY)

		1 Hi-Quality Young	2 Low-Quality Young	3 Hi-Quality Old	4 Low-Quality Old	Total
		N=244	N=18	N=48	N=12	N=322
PRIMARY SAMPLE	Within-Group Multiple R's					
	Group Equations	.38	.80	.25	.84	
	Total Equations	.37	.40	.19	.25	.39
	Moderator Means					
	(Quality Index)	1.13	3.50	1.17	3.67	1.36
	(Age)	21.47	21.45	24.72	25.27	22.10
	Criterion Means					
	(1=No PhD; 2=PhD)	1.82	1.72	1.52	1.33	1.75
		N=230	N=26	N=56	N=9	N=321
SECONDARY SAMPLE	Within-Group Multiple R's					
	Group Equations	.43	.36	.54	.23	
	Total Equations[a]	.44	.57	.58	.52	.53
	Moderator Means					
	(Quality Index)	1.10	3.69	1.16	3.56	1.39
	(Age)	21.54	21.69	24.93	25.67	22.26
	Criterion Means					
	(1=No PhD; 2=PhD)	1.81	1.69	1.54	1.22	1.74

[a]Correlation between observed and predicted score when predicted scores were obtained using regression weights from the total primary sample.

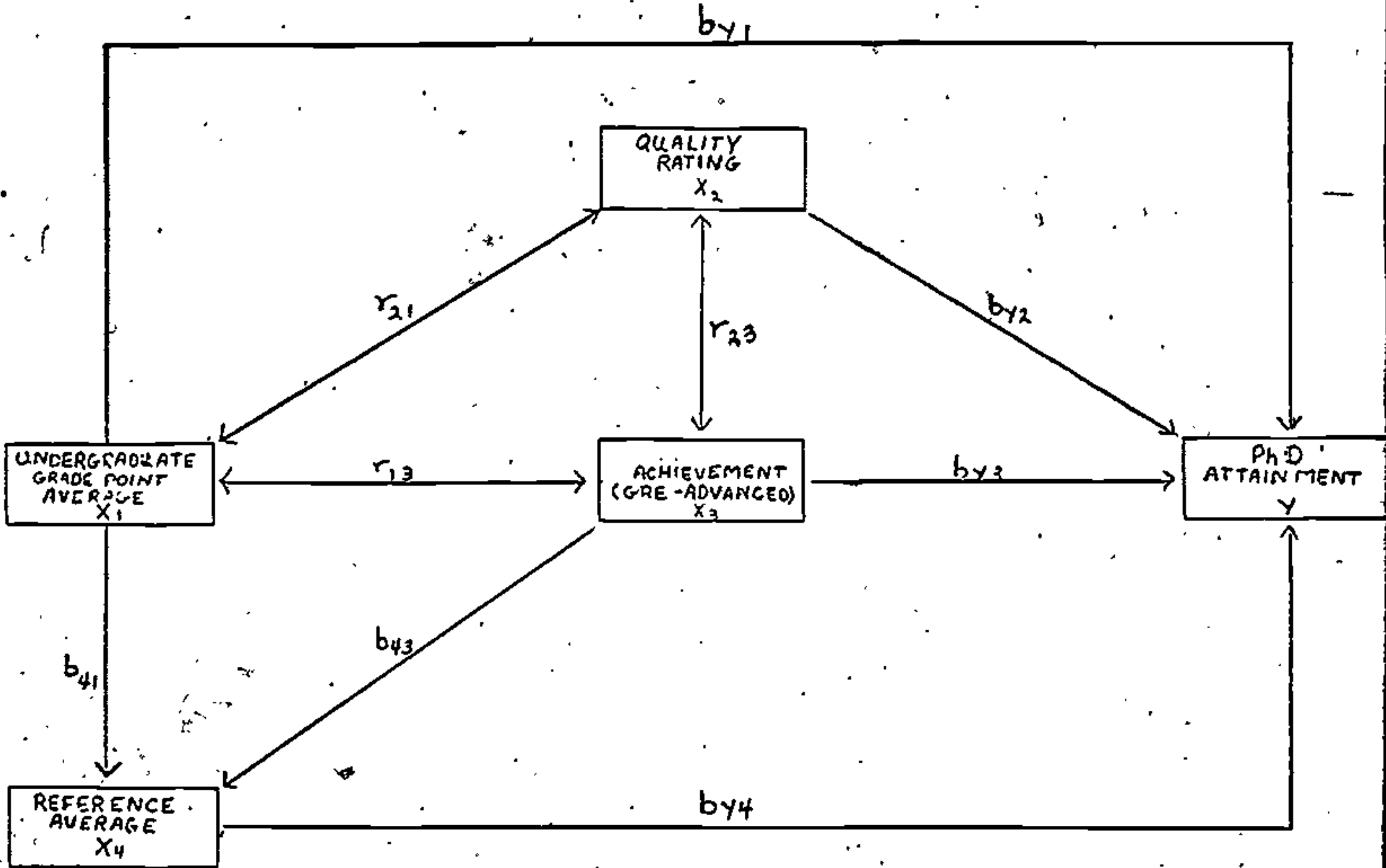
by1
QUALITY RATING X2
r21
r23
by2
UNDERGRADUATE GRADE POINT AVERAGE X1
r13
ACHIEVEMENT (GRE-ADVANCED) X3
by3
PhD ATTAINMENT Y
b43
b41
REFERENCE AVERAGE X4
by4

Fig. 1.

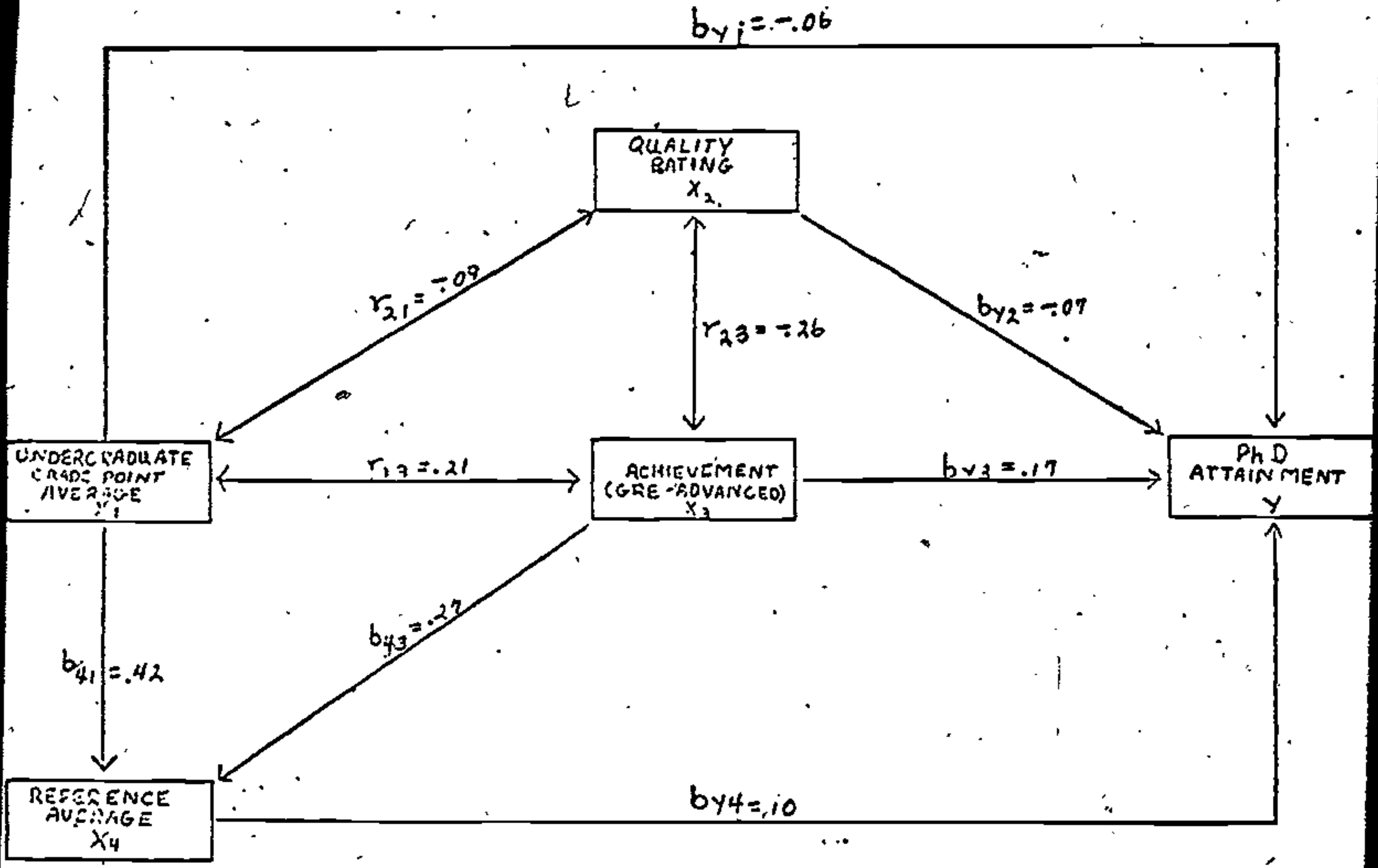

Fig. 2. Psychology Path Coefficients

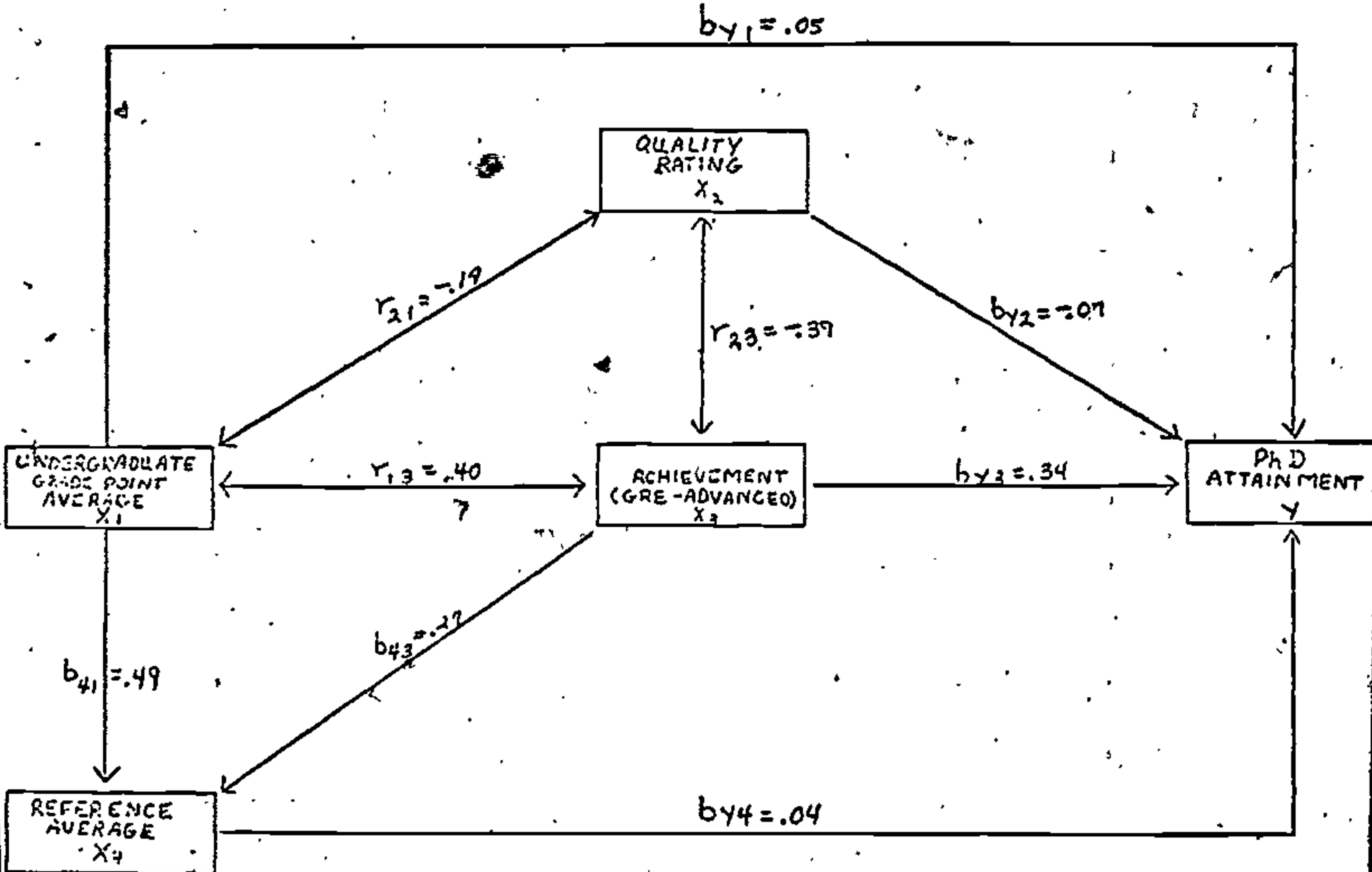

Fig. 3. Mathematics Path Coefficients

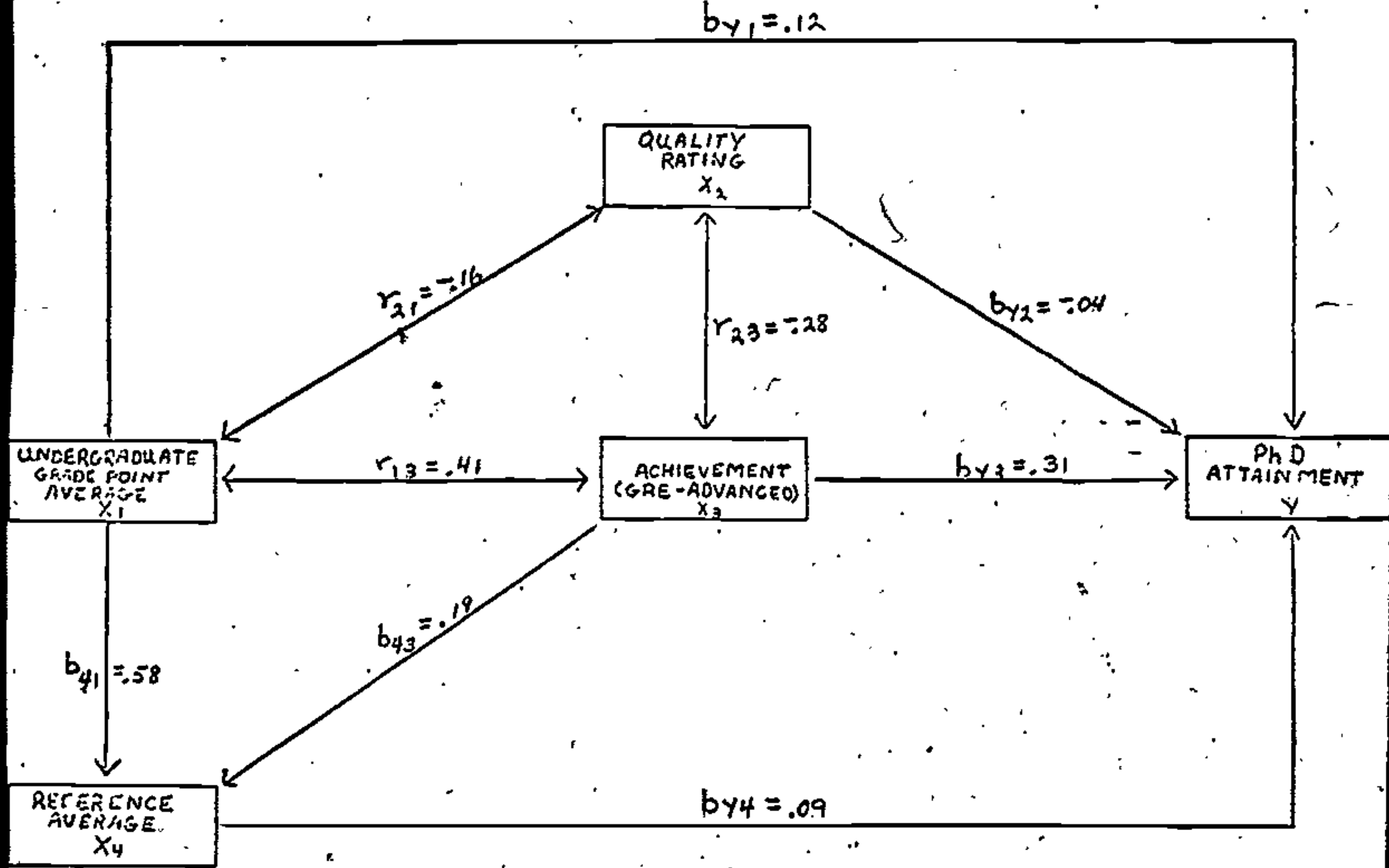

Fig. 4. Chemistry Path Coefficients

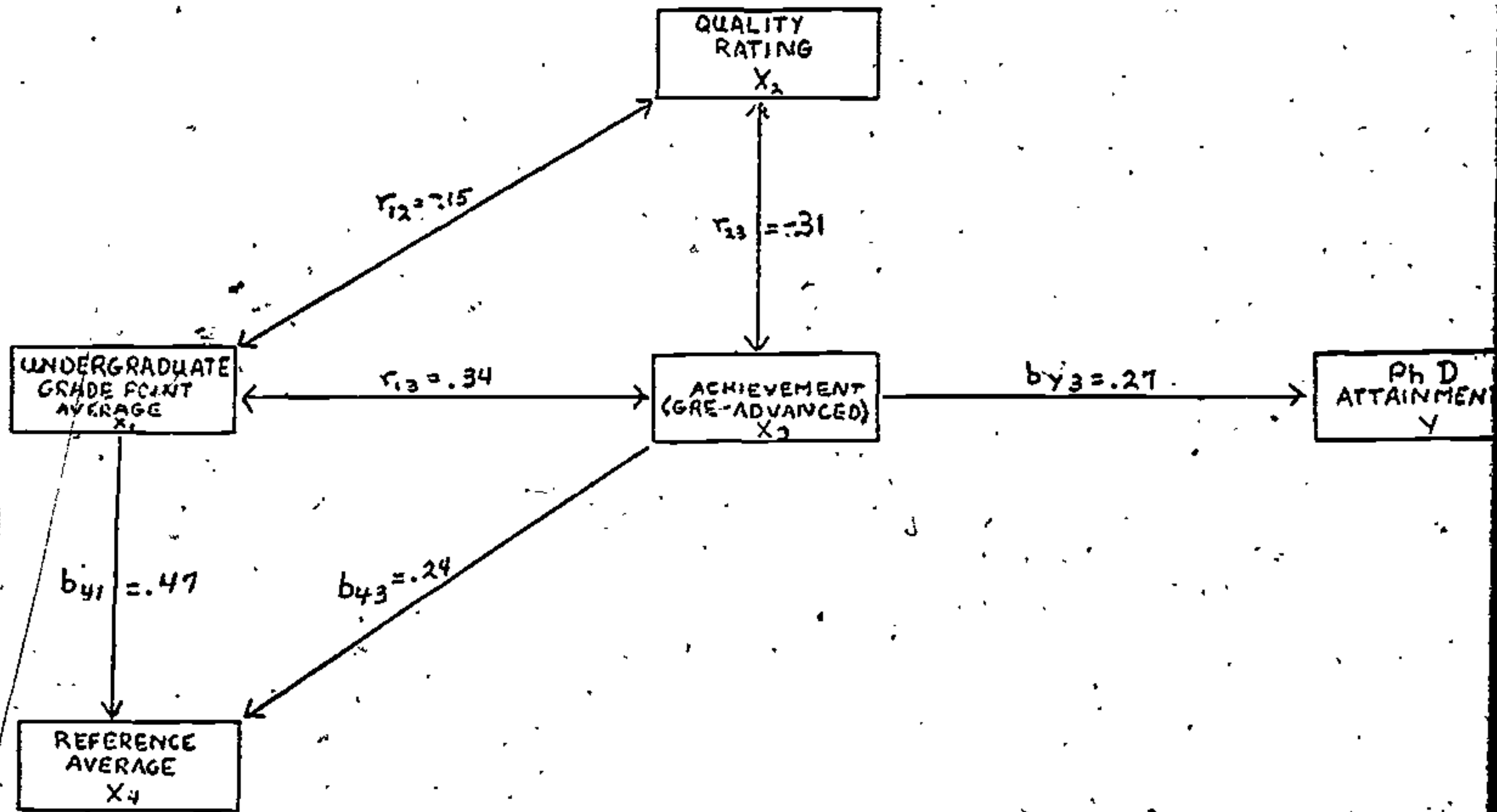

Fig. 5. Reduced Model Average Path Coefficients

Appendix A

Supplementary Analysis of Within School Validity Coefficients

Following the release of the non-technical report[1] on this project, questions as to the value of the data presented to the chairman of a particular department were raised. These were primarily concerned with the effect of between-school differences on the correlations between the predictor variables and the criterion of degree completion. It was suggested that the results would be more useful to department chairmen if these between-school differences were removed and the correlations recomputed. The resulting correlations would give a more direct answer to the question: "Within a given institution, how much help are age, GRE-V, GRE-Q, GRE-Advanced, Reference Average, and Undergraduate Grade Point Average in predicting completion?"

To remove between-school differences, all variables describing individuals were measured as deviations from their respective school means. We were then able to pool the data across schools and thus arrive at a single estimate which may be considered a kind of weighted average of the within-school validity coefficients.

Table 1a presents both the 0-order validity coefficients as well as the correlations based on deviations from school means for the

- - - - - - - - - - - - - - -

Insert Table 1a about here

- - - - - - - - - - - - - - -

Mathematics NSF applicant samples. The deviation correlations are somewhat smaller than the simple 0-order validity coefficients as one might expect since in a loose sense they are partial correlations with the school effect partialled out. When the deviation correlations are compared with the 0-order correlation, the predictors of Ph.D. attainment maintain approximately the same rank order with respect to their accuracy of prediction. That is, the GRE-Advanced remains the best single predictor of doctorate attainment. The second best predictors were GRE-Verbal or -Qualitative, closely followed by either Undergraduate Grade Point Average or Reference Average. This general rank ordering was consistent across both mathematics samples.

---

[1] The prediction of doctorate attainment is Psychology, Mathematics, and Chemistry. Graduate Record Examinations Board, Preliminary Report, Educational Testing Service, Princeton, New Jersey. August 1972.

Table 2a presents similar data for the two samples of Chemistry NSF applicants. Once again, the single best predictor is the GRE-Advanced

- - - - - - - - - - - - - -

Insert Table 2a about here

- - - - - - - - - - - - - -

test. Unlike the Mathematics results, there seems to be no consistent rank ordering for the "next best" predictors across both Chemistry samples. It does, however, seem that Undergraduate Grade Point Average is a more accurate predictor of Ph.D. attainment for Chemistry NSF applicants than for Mathematics NSF applicants. Similar conclusions resulted from the path analysis approach[2] described earlier. It is interesting to note that age tends to maintain its predictive accuracy after the school means are removed. This suggests that older NSF applicants (age measured at entry) appear to be less likely (or they take longer) to attain their doctorate in Chemistry.

Conclusion

The supplementary analysis using individual scores as deviations from school means does tend to reduce somewhat the size of the validity coefficients but the same patterns of predictive accuracy which were found when the data were pooled were maintained. Even though the validity is reduced the resulting coefficients are at a generally acceptable level for academic admissions purposes.

These results suggest that further analysis does not seem to be warranted.

---

[2] Similar to the goals of the deviation procedure, path analysis attempts to control for school effects by incorporating departmental quality indices in a pre-specified "causal model." The resulting path coefficients are more closely akin to part correlations, however, rather than partial correlations.

Table 1a

CORRELATION COEFFICIENTS OF SELECTED STUDENT AND SCHOOL VARIABLES WITH DOCTORATE ATTAINMENT IN MATHEMATICS

	SAMPLE I N=423				SAMPLE II N=422			
	r^a	$\bar{X}$	σ	r^b	r^a	$\bar{X}$	σ	r^b
AGE	-.17	22.35	2.06	-.10	-.18	22.71	2.50	-.15
GRE-V	.27	62.95	10.96	.19	.32	62.63	11.33	.27
GRE-Q	.27	72.67	9.51	.23	.26	71.54	10.14	.17
GRE-ADV	.38	65.93	15.39	.31	.44	64.93	15.94	.36
REFERENCE AVERAGE	.23	42.60	9.38	.16	.27	42.59	9.69	.19
UNDERGRADUATE GRADE POINT AVERAGE	.21	2.52	.40	.15	.24	2.49	.43	.20

r^a is the simple product moment correlation between the selected variables and Ph.D. attainment across all schools.

r^b is the same as r^a except all individual scores are deviations from the school means. This may be considered a kind of weighted average of the within school correlation.

Table 2a

CORRELATION COEFFICIENTS OF SELECTED STUDENT AND SCHOOL VARIABLES WITH DOCTORATE ATTAINMENT IN CHEMISTRY

	SAMPLE I N=322				SAMPLE II N=321			
	r^a	$\bar{X}$	σ	r^b	r^a	$\bar{X}$	σ	r^b
AGE	-.29	22.10	1.81	.27	-.28	22.26	1.82	-.23
GRE-V	.15	59.66	10.69	.15	.23	58.40	10.75	.15
GRE-Q	.28	69.26	10.70	.28	.34	67.96	10.70	.24
GRE-ADV	.33	67.41	11.81	.31	.48	66.27	12.31	.39
REFERENCE AVERAGE	.30	41.48	9.89	.26	.33	42.23	9.15	.27
UNDERGRADUATE GRADE POINT AVERAGE	.27	2.47	.45	.22	.36	2.48	.43	.35

r^a is the simple product moment correlation between the selected variables and Ph.D. attainment across all schools.

r^b is the same as r^a except all individual scores are deviations from the school means. This may be considered a kind of weighted average of the within school correlation.