Masters Theses & Specialist Projects   Graduate School

7-1980

A Study of the Validity of Undergraduate Grade Point Average and Graduate Record Examination Scores as Predictors of Success in the Selection of Masters Level Graduate Students in Psychology

Margaret Tandy O'Connor

Recommended Citation

A STUDY OF THE VALIDITY OF UNDERGRADUATE
GRADE POINT AVERAGE AND GRADUATE RECORD
EXAMINATION SCORES AS PREDICTORS
OF SUCCESS IN THE SELECTION OF
MASTERS LEVEL GRADUATE STUDENTS
IN PSYCHOLOGY

A Thesis

Presented to

the Faculty of the Department of Psychology

Western Kentucky University

Bowling Green, Kentucky

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Margaret Tandy O'Connor

July, 1980

A STUDY OF THE VALIDITY OF UNDERGRADUATE
GRADE POINT AVERAGE AND GRADUATE RECORD
EXAMINATION SCORES AS PREDICTORS
OF SUCCESS IN THE SELECTION OF
MASTERS LEVEL GRADUATE STUDENTS
IN PSYCHOLOGY

Recommended   **7 - 9 -80**
                          (Date)

Richard L. Miller,
                   Director of Thesis

Betsy Houston

John O'Connor

Approved   *July 24, 1980*
                (Date)

Elmer Gray
Dean of the Graduate College

Acknowledgements

I wish to express my appreciation to the members of my Thesis Committee, Rich Miller, Betsy Howton and John O'Connor. The suggestions made by Rich and Betsy have proven invaluable in helping to organize and present a particularly knotty problem in a somewhat manageable manner.

John's contributions began with suggesting the problem and continued throughout its completion. His guidance in the statistical analysis of the data is gratefully acknowledged. Suggestions made by Ray Mendell concerning statistical difficulties inherent to research of this type were enlightening and most helpful.

This thesis is dedicated with much love to John, who as Department Head, Committee Member and husband has appropriately encouraged, supported, advised, criticised and consoled me throughout the entire ordeal. His constant love, abiding faith and inexhaustible patience have never been more appreciated and undoubtedly have contributed immeasurably to my "success" in graduate school.

Table of Contents

LIST OF TABLES

A STUDY OF THE VALIDITY OF UNDERGRADUATE GRADE POINT AVERAGE
AND GRADUATE RECORD EXAMINATION SCORES AS PREDICTORS OF SUC-
CESS IN THE SELECTION OF MASTERS LEVEL STUDENTS IN PSYCHOLOGY

Margaret Tandy O'Connor   July 1980   72 pages

Directed by: R. Miller, J. O'Connor and B. Howten

Department of Psychology   Western Kentucky University

The purpose of this study was to assess the validity of
the undergraduate grade point average (UGPA), the grade point
average during the last two years of undergraduate work (LTYR),
the Graduate Record Examination Verbal plus Graduate Record
Examination Quantitative scores (GREM), the Graduate Record
Examination Analytical Test score (GREAN) and the Graduate
Record Examination Advanced Psychology Test score (GREAD) as
predictors of graduate grade point average (GGPA) in the Master
of Arts degree program in Psychology at Western Kentucky Uni-
versity. The validity of the Graduate Record Examination
Quantitative score (GREQ) as a predictor of grades received
in the two required research methods courses, Psychometric
Theory (GPAPT) and Experimental Design (GPAEXD) was also
investigated. The subjects were 55 students who had completed
at least two semesters of graduate work in the program. The
findings indicate that UGPA and GREM are both valid predictors
of GGPA and contribute about equally to the validity of the
prediction. The addition of GREAN appears to add to the pre-
diction but, for this sample, the addition of GREAD reduced
its validity. GREQ was found not to be a valid predictor of
grades received in the two required statistics courses.

vi

Introduction

During the last several years the selection of graduate
students for both M.A. and Ph.D. level training in Psychology,
as well as in other disciplines, has become a matter of
increased concern and extensive investigation. The numbers
of applicants have increased. The opportunities for graduate
study have become relatively limited. The qualifications of
those who seek to do graduate work are often very similiar.
The greater number of applications to be considered has
increased both the time and the emotional frustrations involved
in the selection process. In addition, there is little con-
sensus as to what data supplied by the candidate are most
predictive of that candidate's potential as a graduate student.
To some extent this judgment can be made only in light of the
expected outcomes of the program, but expectations often vary
from program to program even within the same psychology
department. Perhaps it is for this reason that consensus is
lacking as to the most useful data.

Many authors have concluded that the results of studies
designed to answer these questions may not be generalizable
from one program to another even within the same college
(Bean, 1975; Willingham, 1974; Lannholm, 1968; Robertson and
Hall, 1964) and that predictor-criterion relationships should
be assessed not only by individual institutions but by

1

departments and perhaps programs, as well. Willingham states, "It should be emphasized also that validity studies at particular schools and departments give varying results. Such variability is exacerbated by the small samples often used, but real variations do occur. It is important to undertake local studies in order to justify selection procedures and utilize available information to maximum benefit."(p.276)

The rationale of this study, then, is the justification of present selection procedures and the utilization of available information to the maximum benefit of both institution and applicant. A prescribed, valid, selection procedure should make the decisions of the members of the Selection Committee more efficient, more predictive and less frustrating. At the same time the use of a valid predictor set would help to insure that every applicant received fair consideration and that those selected had at least the potential for success.

In addition, the identification of a valid predictor set may help to focus attention upon the criteria of success in the graduate programs and how these criteria relate to success in the profession of Psychology. This study has been an attempt to explore the validity of the predictor variables currently in use by the Psychology Department of Western Kentucky University and to determine their comparative contributions to the effective prediction of success in the various programs in psychology.

Review of Related Literature

The selection of students for graduate programs has become increasingly complex. While there are now relatively fewer opportunities for graduate training, the numbers of applicants and the expense involved in processing these applications have greatly increased over the past several years. The current legal and social emphasis upon the right of every applicant to be fairly considered lends added importance to the traditionally crucial selection process.

Decisions of selection committees are not only crucial to the individual applicant but to the educational institution as well. The investment of time, money and effort by the institution and its faculty demands that the applicants selected for acceptance into a particular program of training be those who are best suited to that specific program and, hence, those most likely to succeed in, and contribute to, their chosen fields. The acceptance of a student who cannot, or does not, succeed is costly both to the institution and to the student. The rejection of one who may have made a contribution to the field, called the Type II error of the selection system by Kelsey and Dobson (1977), is costly to the profession, the institution and to society.

In an effort to minimize such errors, a number of graduate institutions have conducted research investigating the

3

relationships of various selection criteria to the success or
failure of students in their programs. The number of predic-
tor variables which have been used in these investigations is
vast and the definitions of success varied.

Among predictors, undergraduate grade point average and
Graduate Records Examinations scores (GREs) appear to be most
widely used. Letters of recommendation quality of undergradu-
ate institution; research experience as an undergraduate; age;
sex; Miller's Analogy Test scores; various written professional
examination scores such as the National Teacher's Examination
scores; self, peer and faculty ratings; time to completion of
degree or other specified program requirements; comprehensive
examination scores and selection committee ratings have also
been used (Willingham, 1974; Hirschberg and Itkin, 1978;
Mehrabian, 1969).

The most commonly used criterion variable appears to be
graduate grade point average (GGPA). Attainment of degree,
time to attainment of degree and faculty ratings are also
frequently used (Willingham, 1974).

Obviously the availability of data, and the program it-
self, define, in some measure, what variables may appropriately
be used. Additionally, the purpose of the study determines
the suitability of the various predictors. For example, peer
and faculty ratings gathered at the end of one semester or one
year of graduate work have been shown to be valid predictors
of success (Hirschberg and Itkin, 1978; Wiggins and Blackburn,
1969). Such predictors are not useful in selecting students,

however, as students must already have been selected before they may rate one another or be rated by their faculties.

The present study is concerned with the efficient selection of potentially successful students for masters degree programs in psychology. Therefore, the focus of this review will rest upon pre-selection predictors (those data available to selection committee members prior to their acceptance or rejection of an applicant to their programs) and how these predictors have been found to relate to the various criteria of success.

In view of the number of predictor variables which have been investigated and the diversity of the findings, it seems appropriate, in the interest of simplicity, to look at each predictor and its relationship to the various criterion variables separately. In fact, few studies actually recommend the use of any single predictor alone.

With this in mind, predictor variables will be considered in the following manner: 1) Cumulative undergraduate grade point average (UGPA) and combinations of various undergraduate grades in specified courses; 2) Graduate Record Examination Test scores; 3) Other objective test scores; 4) Quality of undergraduate institution; 5) Personal characteristics of the applicant; 6) Combinations of variables and weighted composites. Next, special notes on methodology will be considered and finally, suggestions for improvement of the prediction process will be reviewed.

Cumulative Undergraduate Grade Point Average and Combinations
of Various Undergraduate Grades in Specified Courses

Willingham (1974), in a review of forty-three studies,
published between 1952 and 1972, states, "By far the most
common predictors used in studies of success in graduate
school are undergraduate average and GRE scores."(p.275) In
fact, both cumulative undergraduate grade point average (UGPA)
and various combinations of grade point average in selected
undergraduate courses have been used as predictors. The find-
ings are often contradictory and the criteria often differ so
that interpretation is somewhat difficult. However, the liter-
ature does appear to suggest that the grades received by a stu-
dent in his undergraduate career are at least helpful in pre-
dicting his success in graduate training, though perhaps less
so when used alone. This brings to mind the comment of
Hirschberg and Itkin (1978) that ". . .it is not surprising
that grades predict grades and ability tests predict perfor-
mance when the time lag is not great between the measurement
of the two."(p.1087)

An examination of the studies reviewed by Willingham
suggests that both UGPA and GRE scores appear to be moderately
good predictors of future success, although according to
Willingham, GRE scores appear to be somewhat more valid than
UGPA when these measures are used alone as single predictors.
These observations are apparently supported by a number of
studies but consensus is far from unanimous.

Stordhal (1970), for example, found UGPA to be the best single predictor of graduate academic performance in a Master of Arts degree program at Northern Michigan University. His criterion measure was graduate grade point average. Ayers (1971) suggests that the use of UGPA appears justified for use in predicting success in a masters degree program in education.

On the other hand, Bean (1975) found UGPA not to be significantly related to any of the criteria of success used in a study of success in a masters degree program in educational psychology. However, an investigation of predictors of success for a sample of special education graduate students at the same institution, published three years later, found UGPA to be the best predictor of graduate achievement for this population with a validity coefficient of .40 (Hartogson, Trainer and Chansky, 1978). Willingham (1974) has suggested that such a coefficient may be considered a "moderate" predictor of success in graduate school.

Heritage (1978), in a study of the predictive validity of admissions criteria for masters students in a reading program, found that UGPA did not differ significantly for graduated and not-graduated groups. The graduated group was made up of students who had attained the degree while the not-graduated group was composed of those who had dropped out of the program.

Robertson and Hall (1964) found that cumulative undergraduate grade point average failed to correlate significantly with faculty ratings of psychology graduate students in a

Ph.D. program when UGPA was used as a single predictor of success.

Lannholm (Aug. 1968) reports upon a cooperative study involving six disciplines representing one or more graduate departments from ten schools. It was found that UGPA failed to consistently predict success when departmental ratings and level of academic achievement of students were used as measures of success in the various programs.

Finally, Hirschberg and Itkin (1978) and Swanson and others (1969) agree that UGPA is an ineffective predictor of the liklihood of completion of degree. Hirschberg and Itkin's population consisted of Ph.D. students in psychology; Swanson and others refer to masters degree students.

While combinations of undergraduate grades in specified courses or for specific periods of time, have not so frequently been studied, the results appear to be somewhat more consistent. Sticker and Huber (1976) identify undergraduate grade point average in psychology courses as the single best predictor of success in a Ph.D. program in clinical psychology. Robertson and Nielsen (1961) found undergraduate grade point average in math and sciences courses to correlate at the .05 level of significance with faculty ratings of students in the Psychology Department at the University of Florida. However, neither grade point average during the last two years of undergraduate work nor undergraduate grade point average in psychology courses was found to correlate significantly with these ratings.

When Mehrabian (1969) factor analyzed thirteen admissions variables, a principal component solution was found which yielded six factors with eigenvalues greater than unity. (The eigenvalue is a measure of relative importance.) These factors reportedly accounted for seventy-five per cent of the total variance. Varimax rotation of these factors yielded six criteria groupings, one of which was a Grade Point Average factor. This factor is defined by overall and last two years' undergraduate grade point average. In further analysis it was found that last two years' GPA had a stronger relationship to graduate school performance than did overall UGPA.

Of the twelve studies cited above, six are in support of the UGPA, or some portion thereof, as being useful in the selection of potentially successful graduate students. However, there does not appear to be any clear line of division. Half of those studies used samples of psychology Ph.D. candidates. Half used M.A. candidates, mostly from various areas of education. Three of the studies using Ph.D. candidates and three studies using M.A. candidates support the use of the UGPA as a predictive measure. The same numbers in each group fail to support the use of this variable. To further complicate the issue, among psychology Ph.D. populations, three are in favor and three against the use of the UGPA as a predictor of success in graduate training. Neither is there consensus as to the utility of the various Graduate Record Examinations Tests scores as single predictors of success.

Graduate Records Examinations Tests Scores

As mentioned earlier, Willingham (1974) sees GRE scores
as somewhat more valid as predictors of success then UGPA when
these measures are used alone. He further states that Advanced
GRE scores appear to be the most generally predictive of over-
all success while GRE Verbal (GRE-V) scores appear to be more
predictive of success in the more verbally oriented disciplines
and GRE Quantitative (GRE-Q) scores the better predictor of
success in those scientific fields where quantitative ability
is desirable. If there is agreement as to the qualities of
a single predictor, or set of predictors, it probably is at
this point but the overall utility of these scores is still
in question.

Robertson and Nielson (1961) found GRE-V scores to cor-
relate at the .05 level with undergraduate psychology grade
point average. Bean (1975) found that graduate grade point
average (GGPA) was correlated significantly with GRE-V and
that GRE-Q correlated significantly with grades in two gradu-
ate research methods courses. Mehrabian (1969) found GRE
Advanced scores to relate strongly and consistently with
graduate success as measured by grades received and depart-
mental ratings.

Lannholm (March, 1968), reviews thirty-eight studies in
which GRE scores were used as predictors of success in gradu-
ate school. He concludes that the Verbal Ability score is
most highly related to performance in subjects of a more
descriptive nature while the Quantitative ability score is

usually more useful in the physical sciences. He also sees Advanced test scores as useful predictors. With specific reference to graduate training in psychology, Lannholm concludes that the Advanced Psychology test was a "somewhat" better predictor of graduate performance when it was used but it was not used in every study.

Heritage (1978) found GRE-V scores to correlate significantly with degree attainment. Graduated groups (those who had attained the degree) differed at the .05 level from not-graduated groups (those who had dropped out of the M.A. in reading program) on this measure only.

Of those variables obtained before graduate admission, Hirschberg and Itkin (1978) found only the various GRE scores and UGPA predicted first year graduate grades. GRE Advanced was found to be the best of the GRE scores in overall prediction, while GRE-V predicted graduate grades in content courses and GRE-Q was found to predict success in the required first year statistics course.

Stordhal (1970) found GRE-V to be significantly correlated with GGPA when it was used independently of UGPA. His sample of students who had taken any one Advanced GRE test was insufficient to allow valid assessment of its relationship to graduate achievement but he strongly suggests this assessment be made.

Robertson and Hall (1964) report a study conducted at the University of Florida. The objectives of this study are reported to be a test of the findings of the previous study

by Robertson and Nielsen (1961) on an enlarged sample of students, as well as a comparison of the predictive success of the Miller Analogies Test, GRE scores and UGPA. In addition, the MAT, GRE and UGPA were correlated with comprehensive examinations and peer ratings and a selection index based upon a weighted combination of GRE, MAT and GPA was correlated with faculty ratings and comprehensive examination scores. The findings of this study support the use of the Mean GRE scores as a predictor of a criterion of success such as faculty ratings. Neither the MAT nor the UGPA correlated significantly with faculty ratings in either the previous or the present study when used alone.

The search for reliable and valid predictors of graduate school success led Ewen (1969) to use the GRE Psychology Advanced Test (GRE-P) as an unobtrusive measure of motivation for students who obtain higher scores on GRE Aptitude Tests. He reasoned that, given equal ability, those students who are more motivated and who possess those personal qualities likely to produce success in graduate study, are more likely to prepare for the Advanced Test and, given the nature of the examination, preparation is likely to lead to higher scores. Therefore, capable students who receive higher scores on the GRE-P are more likely to succeed than capable students who receive lower scores on this measure. In a study designed to test this hypothesis, Ewen used combinations of GRE-V, GRE-Q, GRE-P and MAT scores as well as UGPA in various combinations as predictor variables. Criterion variables were percentage

of A grades in graduate school and degree attainment. From the statistical analysis it is apparent that restriction of range was operating on both GRE-V and GRE-Q. Interestingly, it was less apparent on GRE-P and did not appear to be operating on MAT. Correlation of GRE-P with percentage of A's produced a validity coefficient of .44. Correlation with the graduation criterion produced a coefficient of .66. Ewen cautions against generalizing these results to other, less capable, populations but does feel that the results suggest the possibility that, for students high in verbal and quantitative ability, the Psychology Achievement Test may be an unobtrusive measure of motivation and that its use could serve to improve the prediction of success in graduate school in psychology.

On the other hand, there are those whose findings fail to support the use of the GRE scores as predictor variables. Borg (1963), for example, concludes that neither Verbal nor Quantitative Graduate Records Examinations scores, used alone, is of value as a predictor of success in the graduate program in education at Utah State University. His sample consisted of students enrolled in M.A. level programs. Predictor variables were GRE-V and GRE-Q scores. The criterion variable was graduate grade point average at the completion of at least fifteen quarter hours of work subsequent to the Bachelor's Degree.

Willingham (1974) has suggested that validity coefficients of .40 may be considered "moderate" and .20 "modest." He

further suggests that conditions of selection for graduate programs are often such that the use of even "modest" predictors may be appropriate. Borg reports a validity coefficient of .36 for GRE-V and .37 for GRE-Q. Both coefficients exceed Willingham's "modest" validity figure and approach his definition of "moderate." Borg, however, reports that with a GRE-V score cut-off of one-half standard deviation below the mean, 72 percent of the unsuccessful students and 27 percent of the successful ones in this study would have been eliminated. The total number of successful students eliminated would have been 41 while the total number of unsuccessful students eliminated would have been 21. This finding leads Borg to conclude that such a process is of doubtful value in a setting such as he has described.

Madaus and Walsh (1965) report a study of the predictive efficiency of the GRE Aptitude Tests for various departments in the graduate school of a New England university. GRE scores were found to be, from a practical standpoint, inefficient predictors of success in this graduate school when subjected to regression analysis. However, when departments were taken singly and had relatively large N's, the correlations between the GRE variables and GGPA were statistically significant and were of a magnitude similar to those reported in other studies of success in graduate work utilizing the GREs as the independent variables.

Newman (1968) concluded that the Aptitude and Psychology portions of the GRE were of little practical value in graduate

student selection. His criterion of success was the scholastic grade point average of the sixty-six graduate students who were studying for advanced degrees in psychology. It is interesting to note that his report of the variance accounted for by these scores (4.45 per cent) is very similar to that reported by Madaus and Walsh (1965).

Similarly, Eckhoff (1966) concludes that in a Stepwise Multiple Regression analysis, using UGPA, MAT scores and the Advanced Education portion of the GRE, the GRE Advanced test added very little to the Multiple R.

Sticker and Huber (1967) report similar findings. Although GRE-Q plus GRE-Psychology was found to be the best of the GRE score predictors, the authors conclude that its contribution was slight in the prediction of GGPA and negligible with Orals. They further state, "The wide usage of the GRE for selection must be questioned in the light of these data. Clearly, the data are from too small and limited a sample to suggest that any institution should forego the GRE as part of its selection battery, but they do suggest the necessity for research by each institution into the usefulness of the GRE, rather than its acceptance at face value".(p.467)

Bean (1975) states, "A practice commonly followed in making admissions decisions is to set a minimum GRE total aptitude score instead of using GRE-V and GRE-Q separately. For the data in this sample, such a practice actually reduced the predictive validity from that obtained in using GRE-V alone. Thus, the validity of the GRE total aptitude score should be checked empirically, rather than assumed".(p.966)

Hartogson, Trainer and Chansky (1978) found UGPA to be the best predictor of graduate achievement with a validity coefficient of .40 for UGPA with GGPA. The addition of GRE aptitude scores increased this coefficient to .42. The authors cite these findings as evidence that only undergraduate achievement of the special education majors is important to graduate success for this population. In their opinion the addition of GRE scores adds little to the predictive validity of the UGPA.

In a postdiction study of the GRE and eight semesters of college grades, Humphreys and Tabor (1973) report puzzling results. When GRE-V, GRE-Q and GRE-Advanced scores were correlated with undergraduate grades, it was found that the aptitude test scores correlated most highly with Freshman rather than with Senior grades. The Advanced test scores correlated most highly with Sophomore grades. Senior grades, particularly during the last semester, correlated least highly with GRE Advanced Tests scores. Two explanations are considered: (1) people are changing or (2) the criterion is changing. The evidence does not appear to clearly support either explanation at this point, but discussion of the possibilities lead the authors to take a tentative look at the correlation between UGPA, GRE scores and psychology graduate grades. While they caution that small correlations and large sampling errors make it necessary that these results be viewed cautiously, it appears that Senior grades tend to have their highest correlation with first year graduate grades. The Verbal GRE

tends to have its highest and only significant correlation with Freshman grades, and the Psychology Advanced test tends to have its highest correlation with graduate grades. Its second highest correlation was again with Sophomore grades. Humphreys and Tabor call for additional research but suggest that GRE Aptitude scores may not be appropriate for use in the prediction of graduate school success.

Lannholm (Aug. 1968) reports widely varying results from a study in which one or more graduate departments from ten schools, representing six disciplines, cooperated. Predictor data included GRE Aptitude test scores and/or Advanced test scores and, for some students, undergraduate grade point average. Departmental ratings and level of academic attainment were used as measures of success in the various programs. Neither UGPA nor Advanced Test scores consistently predicted success.

With reference to predicting the success of graduate students in psychology, Lannholm reports that of eight groups of graduate students in psychology, only the data for the groups in one department yielded reasonably high validity coefficients. In two departments modest positive coefficients (.27 and .45) for available test scores were found but there were small negative validities for Undergraduate GPA. In both of these departments the Quantitative Ability score was found to be the best predictor.

In still another department Advanced and Quantative scores appear to show promise but the Verbal score yielded a

negative validity (-.28). And in three other departments all coefficients were low; in one of these all were negative. Lannholm hypothesizes that the specializations within psychology may be more disparate than those within other disciplines.

Obviously the question of the utility of GRE scores remains unanswered. Once again, the literature lacks consensus. Of the eighteen studies reviewed, nine support the use of GRE scores either singly or in combination as useful in predicting graduate school success while nine fail to find these scores consistently valid. Five of the studies involving Ph.D. programs support this variable while three do not. Three studies involving M.A. graduate students find GRE scores useful; four do not. Of three studies using both Ph.D. and M.A. candidate populations, one is in favor of GRE score use, two are against its use.

Other Objective Test Scores

It appears that most attention has been focused upon the use of the UGPA, the GRE scores and, especially in earlier studies, to some extent the MAT as predictors of success in graduate study. Undoubtedly this is because these measures are so widely used in the selection process. However, there have been occasional attempts to investigate the merits of test score predictors not routinely included in the selection formula.

For example, Ayers (1971), reasoning that the ability to use the English language effectively would enhance the performance of graduate students in a masters program in education,

included the New Purdue Placement Test in English (PET) and some portions of the National Teacher's Examination (NET) in a battery of predictor variables to be investigated. The sample (N=241) included all students who had completed the Master of Arts in Education between June, 1963, and August, 1970, at a regional state university. The areas of major emphasis included Administration and Supervision, Curriculum and Instruction and Guidance and Counseling. The correlations between GGPA and UGPA, MAT and PET were highly significant for those students included in the first two areas of emphasis. These correlations were less substantial for those students majoring in Guidance and Counseling. The author suggests that this could be explained, in part, by the curriculum of Guidance and Counseling which places less emphasis on cognitive mastery of academic content and more on performance in the form of counseling techniques, testing, interviewing and intern situations. (It would be interesting to know whether Clinical Psychology students, as opposed to those majoring in other areas of psychology, differ in the same ways on such measures. Lannholm apparently suggests they may.)

The correlations between MAT and GGPA were interpreted by Ayres as justifying the use of the MAT when only a single predictor is used. In addition the use of the UGPA and/or PET as predictors of success in graduate work in education appear justified. The introduction of selected scores from the NET appears to enhance the predictive qualities of these variables.

In the face of criticisms of tests as being inaccurate, biased and irrelevant, Baird (1976) takes a slightly different approach to the assessment of such scores as predictor variables. This study looks at those personal characteristics of students that correlate with various commonly used standardized measures which are thought to predict success in graduate school. These measures include not only the Verbal and Quantitative portions of the GRE but the Law School Admissions Test (LSAT) and the Medical College Admissions Test (MCAT) Science score as well.

Survey questionnaires were collected from a sample of 21,000 college seniors at 94 colleges in 1971. Personal information was obtained in 14 categories. A total of 4,375 of those students had taken the GRE; 1,845 the LSAT; and 959 the MCAT. Personal characteristics were correlated with scores on the various tests.

While the results of this study are somewhat mixed, they appear to indicate that students who receive higher scores on these standardized measures also possess those characteristics valued by selection committees of graduate and professional schools. Several of the tests also appear to be correlated with some background characteristics that appear to be unrelated to the purpose of selection. In general the author seems to believe that these tests do what they are intended to do, but he points to a need for further research to determine whether they may also reflect irrelevant characteristics of students which may contribute to bias in the selection process.

Baird states, ". . .admissions tests have been, and are, merely the technical apparatus for the evaluation of students' academic promise -- an apparatus which is especially effective in predicting grades".(p.415)

Quality of Undergraduate Institution

The selection procedure appears to have traditionally relied heavily upon the use of objective test scores in an effort to choose those applicants best suited to graduate training programs. Mehrabian (1969), for example, found that ratings of a student by the faculty, as well as acceptance decisions by the selection committee, loaded on a GRE-MAT factor defined by GRE-V, MAT, GRE-Advanced and GRE-Quantitative scores. However, there are other pre-selection factors available to the selection committees of graduate departments which may influence their decisions. One such variable is the Quality of the Undergraduate institution (QI). It seems logical that, all things being equal, those students who had done undergraduate work at "better" institutions would be better prepared to succeed in a graduate program than would those from less rigorous institutions. Since differences among "good" applicants are often minimal this would appear to be one factor which might prove useful. Unfortunately, the literature is not supportive of this hypothesis. Dawes (1971), in a study which focused upon the decision making process of the selection committee, found that UGPA and QI alone correlated more highly with later faculty ratings of accepted students than with the ratings of those students by the selection committee.

He states,

> "The weights used to predict the faculty ratings are
> presented in Equation 1, while those used to predict
> the admissions committee ratings are presented in
> Equation 2.

$$.0006 \text{ GRE} + .76 \text{ GPA} + .2518 \text{ QI} \quad (1)$$

$$.0032 \text{ GRE} + 1.02 \text{ GPA} + .0791 \text{ QI} \quad (2)$$

> It is of passing interest to note that the admis-
> sions committee does not place sufficient weight on
> the quality of the undergraduate academic institu-
> tion. In fact, this quality index is the best single
> predictor of later faculty ratings. The interpreta-
> tion made here of this finding, however, is not that
> all admissions committees everywhere should place
> more weight on the variable of undergraduate insti-
> tutional quality, but rather that such quality be-
> came a good predictor among the selected group by
> virtue of the fact that the admissions committee
> tended to ignore it." (p.185)

Other authors are less supportive of the QI as a measure
of future success. Heritage (1978) found undergraduate col-
lege quality not discriminative between his graduated and not-
graduated groups of former masters students in the Rutgers
University Reading Program. Hirschberg and Itkin (1978)
found that QI failed to predict either first year grades or
completion of degree. They state that its further use as a
single predictor is not warranted.

Goldberg (1977) argues against the inclusion of an esti-
mate of the QI in a formula for use by the selection committee
at the University of Oregon. According to this author the
use of this variable only increases opportunities for clerical
error and time required for computation of the linear compo-
site score. He states that it adds nothing else.

Merenda and Riely (1971) report that QI assumed the smallest, but significant, positive weight in a valid predictor set for selecting graduate students in psychology. Mehrabian (1969) found QI not to be significantly related to any of the performance indices used in a study of admissions criteria at UCLA.

Personal Characteristics

In addition to UGPA, objective scores, and quality of undergraduate institution, the admissions committee normally has access to certain personal and individual characteristics of applicants such as sex and age. Letters of recommendation are usually required and, in some cases, an evaluation of the applicant's research orientation, promise and committment to psychology.

The sex of the applicant has received passing comment in the literature. Hirschberg and Itkin (1978) found sex and time to completion of various program requirements to predict Ph.D. attainment. They state, "It is safe to say, based on these data, that one of the best predictors of who would obtain the degree was sex: Men did and women didn't."(p.1090) They further state that women did not differ greatly from men on those variables related to obtaining a degree. "Women were not rated by their peers as being less able or less conscientious; they were rated (accurately) as not finishing." (p.1091) According to these authors, only 35 percent of the women in their sample had obtained a degree by 1975. Sixty-eight percent of the men had received a degree by that time.

Mehrabrian (1969) found sex unrelated to any of the performance indices used in his study of relationships among criteria which could be used in the selection of students for graduate training. However, these indices did not include attainment of degree. His focus was upon the performance of a student during the first year of graduate school. This appears to fit the findings of Hirschberg and Itkin that women are no less capable than men in graduate performance even though they finish the degree less often. This phenomenon appears not to have been systematically explored.

A second bit of demographic data is the age of the applicant. Heritage (1978) found UGPA and age to yield significant correlations with time to degree completion. Swanson and others (1969) state that on the basis of the data analyzed, students who continue on to graduate work immediately after receiving the baccularate degree are significantly more likely to complete the masters degree than are those for whom a period of time elapses between undergraduate and graduate work. Lafferty (1969) found no significant correlation between age and the predictability of GRE Aptitude Test scores for successful graduate students.

Research orientation has been thought to hold promise as a predictor of success in graduate training but generally has not been found fruitful. Mehrabian (1969), for example, found a Research Orientation factor composed of research experience, research orientation and the letter of recommendation rating a candidate's research potential to be unrelated to

performance during the first year of graduate training. However, this author and others (Hirschberg and Itkin, 1978; Goldberg, 1977; Lannholm, Aug. 1968) have suggested that ratings, obtained from letters of recommendation, of research versus service orientation may contribute to increased accuracy of prediction in this area.

Although it is generally agreed that letters of reference as currently used are of little value in the selection process (Hirschberg and Itkin, 1978; Goldberg, 1977; Lannholm, Aug. 1968; and Kelsey and Dobson, 1977), many of these authors call for revision of the traditional letter of recommendation. Ratings of promise, of research orientation and of the rater's familiarity with the ratee are suggested. Goldberg (1977) also suggests that applicants be allowed to choose the number of letters of recommendation they provide, based on how well they are known by the faculty members, or that they be allowed to send a paper written within the last four years, and representative of their best work, in lieu of letters of recommendation. These revisions and innovations are, of course, intended to provide more objective and less "creative" information upon which to base acceptance decisions. Willingham (1974), however, notes:

> "One might suppose that motivation to undertake graduate work would be one important quality reflected in letters of recommendation, but the validity of such references is disappointingly low. In extensive studies of NSF fellowship applicants, the reliability of single references was reported to be in the low .30's. This may be the main reason why recommendations are poor predictors, but careful efforts to improve that reliability with

multiple ratings did not result in good validity
for the NSF fellowship recommendations. Such re-
sults do not suggest that improved letters of
reference would increase accuracy of prediction."
(p.276)

Combined Variables and Weighted Composites

As stated earlier, few studies actually recommend the
use of a single predictor variable in selecting graduate stu-
dents. Although many report single predictors to be signi-
ficantly correlated with various criterion measures, the
majority appear to find that combinations of variables and
composite formulations increase the validity of prediction.

Robertson and Nielson (1961) found the combination of
Mean GRE and UGPA in math and science courses to yield the
highest correlations with faculty ratings. They conclude
that combining the two predictors would be a definite improve-
ment over selection based on either one alone.

Lannholm (Mar. 1968) in a review of studies which used
GRE scores as predictors of success in graduate school during
1972 and 1976, makes four broad generalizations. The first
three are concerned with the validity of the GRE scores as
predictor variables and are supportive of that measure. The
fourth generalization is that the best predictions of success
were obtained when UGPA and GRE scores were used in combina-
tion.

Willingham (1974) states that a weighted composite of
undergraduate grade point average and GRE scores provides
substantially more accurate predictions than does undergradu-
ate grade point average alone. According to Willingham, this

composite provides a validity coefficient in the .40 to .45 range for various criteria of success across academic fields. This coefficient is reported to be somewhat higher than the validity of the GRE scores alone.

Lannholm (Aug. 1968), following a report on studies in which ten graduate schools cooperated, states that, when undergraduate GPA was available, the judgmentally weighted total obtained by applying reasonable weights to each predictor was found to yield better predictions than Undergraduate GPA used alone.

Robertson and Hall (1964) report that when a selection index based upon a weighted combination of GRE, MAT and GPA was developed, it was found to correlate significantly with faculty ratings and comprehensive examination scores and to do a better job of prediction than did any of the three predictors used alone.

Merenda and Reily (1971) conclude that total undergraduate GPA, GPA in psychology courses, GRE-V, GRE-Q, GRE-Advanced and quality of the college in which the baccalaureate degree was earned constitute a valid predictor set for selecting graduate students in psychology. On the average, the most successful students had the highest mean scores on all six predictor variables. The failure group had, on the average, the lowest mean scores.

Mehrabian (1969) investigated the relationships among a variety of criteria which could be used in the selection of students for graduate training. In addition, the validity of

these criteria as predictors of success in graduate school
was assessed. In the first part of the study, unselected
applicants for the graduate psychology program at UCLA were
rated on each of thirteen admissions criteria. These criteria
were then factor analyzed. Next, the admissions criteria
scores were related to three indices of performance in the
psychology program. These indices of performance consisted
of an average evaluation of a student's competence and pro-
mise in the area of research, the average grades of the stu-
dent during the first year of graduate work in content courses
and the average grades received by the student in statistical
courses during the first year of graduate school.

When the original thirteen admissions variables were
factor analyzed and a principal component solution found, six
factors were found which had eigenvalues greater than unity.
These factors accounted for 75 per cent of the total variance.
Varimax rotation of these factors yielded criteria groupings
as follows:

1) GRE-MAT factor defined by GRE-V, the MAT, GRE-Ad-
   vanced and GRE-Q scores.

2) Grade Point Average factor, defined by overall and
   last two years undergraduate grade point average.

3) Research Orientation factor including research expe-
   rience, research orientation and the letter of recom-
   mendation rating of a candidate's potential as gradu-
   ate student and research worker.

4) Grade Point Average Improvement factor consisting of the increase in undergraduate grade point average during the last two years of undergraduate work as opposed to the first two years.

5) Sex factor, determined by the sex of the candidate.

6) Mathematical Training factor, defined by the total number of mathematics and logic courses taken as an undergraduate and a relatively low rating of the psychology program attended by the student as an undergraduate.

When the three criteria of graduate school performance and the admissions criteria were correlated, it was found that neither student sex, increase in undergraduate grade point average, the rating of the program in which the student did his undergraduate work nor research experience were significantly related to any of the performance indices.

Through regression analysis, Mehrabian developed the following formula for the selection of students by the psychology program at UCLA:

Graduate school performance = 1.34 (GRE-MAT index) + 105.7 (letter of recommendation rating) + 22.5 (research orientation rating) + 18 (number of math and logic courses) + 91 (last two years' GPA).

The author concludes that the GRE-MAT scores bear the strongest relationship to performance in graduate school. However, ratings obtained from letters of recommendation and a rating of research versus service orientation may also contribute to increased accuracy of prediction of success in graduate school.

Goldberg (1977) describes the graduate admissions process at the University of Oregon Psychology Department. When a prospective applicant applies to the department, a packet of materials is sent which includes a linear composite that potential applicants can compute for themselves. It is identical to the formula used by the selection committee as a preliminary screening index for majority students except that it does not include a quality index rating of the applicant's undergraduate institution. Goldberg suggests that a new formula for the use of the admissions committee not include the QI because, as mentioned earlier, he believes this to be an ineffective measure. He adds that virtually nothing is gained from its use. His suggested composite score formula is $GLD = GPA + \frac{GRE\text{-}V + GRE\text{-}Q}{200}$. Majority students whose composite scores fall below 9.5 have no possibility of admission, are rejected and immediately notified to this effect. Those who score above 9.5 and all minority students then become potential candidates for selection. In this way, the number of applications to be reviewed by the selection committee is effectively reduced by the elimination of applicants who, in all likelihood, would neither be accepted nor successful had they been accepted.

Others suggest the use of composites but do not focus upon the identification of specific variables. These include Dawes (1971), Hirschberg and Itkin (1978) and Wiggins and Kohen (1971).

Notes on Methodology

Although many authors discuss the difficulties of predicting success in graduate school, Chansky (1964) specifically addresses the problem of the GPA as a dependent variable in studies of academic achievement. He points out that grades have no inherent stable meaning because the several sources of variance which would explain a given grade are not known. Since an underlying assumption of interval scales used in product moment correlations is normal distribution of the characteristic, and grades are not normally distributed, but skewed, usually negatively, Chansky states that computations involving the assumption of normality are not permissible. He suggests the use of ordinal scales rather than nominal scales since the categories are not equal. In other words, A is greater than B, B is greater than C and so on. Grades would be rank ordered, A being first, B being second and so on. The GPA would be the median grade and correlations would be of the rank type. Chansky cautions that even when non-parametric techniques are used the findings would apply only to the sample surveyed because schools differ so widely in curricular goals and marking practices.

A study concerned with the investigation of differences in the predictive efficiency of the Graduate Records Examination Aptitude Tests for various departments in the graduate school of a New England University is reported by Madaus and Walsh (1965). The sample of 569 first year graduate students was enrolled in a number of departments within the university.

The criterion variable was graduate grade point average at the
end of the first semester's work. Zero order product moment
correlation coefficients for GRE-V with GGPA, GRE-Q with GGPA and
GRE-V with GRE-Q were computed to be .19, .18 and .45 respectively.
These correlations are all significant beyond the .01 level. The
Multiple R between GGPA and GRE-V and GRE-Q was found to be .22.
This correlation is significantly different from zero at the .01
level, with 4.84 percent of the variance in GGPA accounted for
by GRE-V and GRE-Q combined.

Six departments, taken separately, yielded significant
correlations between GRE scores and GGPA. The median sample size
for these departments was 43. Six other departments, taken separately,
failed to yield significant correlations. However, the median sample
size for these departments was only 18. The authors suggest that
the size of the sample is a definite factor relative to whether
significant correlations are found between predictor and criterion.
They caution that grouping of departments for predictive purposes
should not be done. They further conclude that GRE scores are, from
a practical standpoint, inefficient predictors of success in gradu-
ate school. They also state that using GRE scores with a Multiple
Regression model does not provide administrators with helpful
information regarding graduate school admissions.

Attenuation and grade inflation further complicate the
prediction process (Willingham, 1974; Madaus and Walsh, 1965;
Chansky, 1964).

Willingham identifies two important weaknesses in the UGPA when used as a predictor of performance in graduate school: The very narrow range of scores and the variation of grading systems among various institutions. For example, the meaning of a B average may vary from one institution to another. Restriction of range is also seen by this author as the major weakness of grades in graduate school when used as a criteria measure of success.

Chansky (1964) points out that teachers assign grades for many reasons some of which are unrelated to academic achievement. In addition, teachers frequently disagree as to the appropriate grade which should be assigned to a given test paper and have been known to change their own grades from time to time. He summarizes that the GPA bases its existence upon capricious judgments and volatile criteria.

Dawes (1871) states that the validity of the usual selection criteria considered alone can be expected to be low because the restricted range of talent among applicants selected attenuates correlations. He also feels that selection committees tend to use compensatory methods of selection. By this he means that students low in one measure must be high on another in order to be selected which, of course, means that people who are low on one value should be expected to succeed on the basis of the other variables used in selection.

Robertson and Hall (1964) explain the low but statistically significant correlations typically found in studies of the prediction of success in graduate school as due to the

relative homogeneity of the samples. They state that the predictive measures being evaluated in relation to certain criteria of success in graduate school have already been employed to select only those students who had shown 'promise'.

Hartogson, Trainer and Chansky (1978) also speak of low validity coefficients. They state that the range is restricted in both predictors and criterion since the students in their sample had to meet certain admissions criterion prior to selection and could receive no more than six credits of C to remain in the program. They further state that 95 percent of the graduate grade averages lie between A and B (SD = 0.26) and that the lack of differentiation among graduate grades may be viewed as grade inflation. According to these authors, such conditions argue against obtaining high coefficients.

Suggestions for Improvement

Finally, the literature is replete with suggestions for improving the predictive validity of the selection process. Willingham (1974) takes a pessimistic view of the possibilities of improving prediction of graduate success but proposes alternate prediction strategies which would take into account multiple criteria of success related to different training objectives. He concludes,

> "The best way to improve selection of graduate students will be to develop improved criteria of success. This is no small job for graduate faculties, but it carries the promise of more effective utilization of talent and greater assurance of equity in admitting students to advanced training and the privilege associated with such training". (p.183)

Lannholm (Aug. 1968) makes four suggestions for improving the prediction of success in graduate school: (1) clarification of the "nature of success in graduate school," (2) exploration of issues involved in the assessment of each student's performance, (3) the identification of additional predictors and (4) the formation of an overall level of "promise" for each student using all the available data at the time of selection.

Hirschberg and Itkin (1978) propose peer ratings taken before admissions as one possibility; rating scales of motivation and committment and familiarity of the rater with the ratee either in lieu of or in addition to the usual letter of recommendation; and the use of "multiple hurdles" model in the graduate student selection and deselection process. The multiple hurdles model involves the student's satisfying various requirements at various time periods from application to completion of degree. In addition, these authors would include non-intellectual measures and, like Willingham, raise the question of how different graduate school treatments, and expected outcomes, relate to constellations of student characteristics and, in turn, to student success in graduate training.

Humphreys and Taber (1973) conclude,

"The prediction of graduate school success may have to be restructured along radically different lines. If change as indicated by intercorrelations and validities continues smoothly from the senior year to the first graduate year, the continued use of the aptitude tests of the GRE becomes highly suspect. Perhaps it would be more useful at present

for an institution to require two or more
advanced tests, say in the major and in one
or two minors, than to require the present
combination of two aptitude and one advanced
test. The data also suggest that the advanced
tests may need to be revised to make them
more responsive to advanced undergraduate
achievement. As a first step, it might be
profitable to look for items in present advanced
tests that are more highly correlated with
senior than with sophomore grades". (p.184)

Bean (1975), Willingham (1974), Lannholm (1968),
Robertson and Hall (1964) and others clearly support the
necessity of validation of predictor variables by individual
graduate programs. Robertson and Hall, for example, suggest
that the predictors, as well as their individual weightings,
should be determined by finding how well each predictor corre-
lates with the various criteria of success in a particular
department.

In an effort to look beyond predictor and criterion vari-
ables, and the problems inherent in them, Dawes (1971) has
focused upon the decision-making process of the selection
committee. He argues, quite convincingly, for the develop-
ment of a simulation of the selection committee's judgmental
process, termed a "paramorphic representation," based upon a
linear combination of the criteria used by the committee mem-
bers in making their selections. This paramorphic represen-
tation could then be used in place of the admissions committee
to make initial screening decisions such as the rejection of
students who would clearly not be accepted by the committee.
The use of this procedure is quite similar to that described
by Goldberg (1977).

Dawes argues that, not only is this a possibility but that, in the final analysis, a linear model based upon the behavior of the committee can be more accurate, less costly, and more human.

In support of his arguments, he report that a model, developed at the University of Oregon to simulate the behavior of the Department of Psychology selection committee, was able to screen out fifty-five per cent of the applicants without a single error being committed. In addition, the predictions of performance made by the paramorphic representation of the selection committee correlated more highly with actual faculty ratings of performance than did the ratings of the committee at the time of selection. According to Dawes, the representation accounted for approximately 25 times as much variance as did the judgment.

Additional support for the use of the model in selecting graduate students is found in a study by Wiggins and Kohen (1971). The purpose of the study was to test, in a situation other than clinical diagnosis of the MMPI, the hypothesis advanced by Goldberg (1970) favoring model over man in clinical judgment situations. When graduate students were asked to predict first year grade point averages of profiles representing other graduate students, they were fairly accurate but in every case the model of the judge was more valid than the judgments themselves. The authors conclude that members of an admissions committee might profitably use a model of their own judgments in order to free time which they feel may be better employed in the search of new admissions indexes.

Finally, there are a number of authors who see the solution to the Graduate Student Selection Dilemma in a radical departure from present procedures. Citing such obvious difficulties with the present system as the expenditure of time, effort and money on the part of both the applicants and the institutions; the emotional frustrations; the inadequacy of predictors and criterion variables; the difficulties of methodology in research when samples are improperly assembled; and the lack of needed longitudinal research, Kelsey and Dobson (1977) call for a centralized registery of psychology graduate students. They feel that this registery might not only be more efficient and economical but might also provide the research data pool necessary in order to establish a more reliable means of identifying and matriculating graduate students who would make significant contributions to psychology. Goldberg (1977) also supports the development of such a centralized system.

This proposal appears to merit consideration on the basis of its research potential alone. However, it would appear that individual selection committees would, for the time being, given the state of the selection art, still be placed in an unenviable position. They would still be required to select small numbers of potentially successful graduate students from large populations of applicants whose qualifications were often very similiar, on the bases of predictor variables whose validity is questionable and criterion measures that may or may not tap the essence of success in graduate school.

Statement of the Problem

The selection of graduate students who will be successful in
a particular graduate program is a difficult and time-consuming task.
It is made more difficult by a lack of agreement as to valid predictor
measures and clearly defined criterion variables. Those students who
would, or could, be successful in one graduate program may, or may
not, be equally suitable to another. Yet, many graduate departments
base these crucial decisions upon "traditional" rather than empiri-
cally validated measures. In this study information will be gathered
regarding the sex, undergraduate grade point average and Graduate
Records Examination scores of those students enrolled in the clinical,
school, industrial and general psychology programs, at Western Kentucky
University, for the years 1977, 1978 and 1979. These data will then
be analyzed to determine whether they constitute a valid predictor
set for selecting future potentially successful graduate students
for these programs and the extent to which each of several predictor
variables contributes to the effectiveness of the linear prediction.
Four Hypotheses will be tested.

Hypothesis I: Cumulative UGPA will predict GGPA during the
first two semesters of graduate work in the Master of Arts degree
program of the Psychology Department at Western Kentucky as efficiently
as will the GPA received during the last two years of undergraduate work.

39

Hypothesis II: The Graduate Records Examination Quantitative
score will predict grades in Experimental Design and Psychometric
Theory.

Hypothesis III: A linear combination of UGPA and scores on the
Verbal and Quantitative portions of the Graduate Records Examination
will predict Graduate Grade Point Average during the first two semesters
of graduate work in the Master of Arts program of the Psychology
Department at Western Kentucky University.

Hypothesis IV: The addition of either the Advanced Test score
or the Analytical Test score on the Graduate Records Examination to
the linear combination of UGPA and GRE-V plus GRE-Q will improve
the prediction of GGPA during the first two semesters of graduate
work at the M. A. level in psychology at Western Kentucky University.

The Null Hypothesis then may be stated as follows: None of
the selected predictor variables bears a significant relationship to
success in the first two semesters of graduate work in the Master
of Arts degree program in the Psychology Department at Western Kentucky
University.

Method

Subjects

There were fifty-five subjects in the sample of graduate students who had enrolled in the four M.A. level psychology programs at Western Kentucky University from the Spring Semester, 1977, through the Spring Semester, 1979. This sample includes all students accepted into the Clinical, Industrial, School or General programs.

As one focus of the study was to be the validity of the Graduate Records Examinations Analytical Test (GRE Analytical) as a predictor of success in these programs, only those subjects were chosen to whom this test was available before application to the programs. A survey of the application files of past entering classes showed that no student had taken this measure prior to the Spring Semester, 1977.

The fifty-five subjects had transcripts from their undergraduate academic institutions on file with the University. Over-all undergraduate grade point average, as well as last two years' grade point average, was computed for each subject from these transcripts. It was not possible to compute last two years GPA from the data available on two of these students.

All subjects had Graduate Records Examination Aptitude scores on file. Seventeen students had taken the Graduate Records Examination Advanced Psychology Test while 28 had

41

taken the Graduate Records Examination Analytical Test. No student had taken both the Advanced and the Analytical Exams. Ten students had taken neither the Advanced nor the Analytical Examinations.

The sex of each subject was noted. There are 24 female students and 31 male students in the sample.

Measures

Predictor Variables

The following measures were obtained for each subject based on information available from the student's application file: overall undergraduate grade point average, last two years' undergraduate grade point average, scores on the Verbal and Quantitative portions of the Graduate Records Examination and either the Advanced Psychology or Analytical Graduate Records Examination scores when available.

Overall Undergraduate grade point average and last two years' undergraduate grade point average were recorded on a five point scale ranging from zero to four. A was considered as 4, B as 3, C as 2, D as 1 and F as zero.

Criterion Measures

The following measures were requested from the office of the registrar for each subject: first semester graduate grade point average, second semester graduate grade point average and the grades of each student in the two required statistical courses (Experimental Design and Psychometric Theory). Again, grade point average was recorded on a five-point scale ranging from zero to four with A equal to 4 and F equal to zero.

Analysis

One-way frequency distributions with descriptive statistics were generated for all variables using the Frequencies procedure found in the Statistical Package for the Social Sciences, 2nd Edition (Nie, Hull, Jenkins, Steinbrenner and Brent, 1975). Descriptive statistics were generated for categorical variables using the Condescriptive procedure found in the same package.

A test of the validity of the Graduate Records Examination Quantitative (GREQ) score in relation to the grades received in the two research methods courses (GPAPT and GPAEXD) used the Discriminant Analysis procedure of the same statistical package.

The Pearson Correlation procedure of the above cited package was used to generate Pearson Product-Moment correlation coefficients for all measures. This same procedure was used to obtain the correlations of undergraduate grade point average with graduate grade point average for two subgroups of the sample: (1) those with GREM scores less than the mean of the GREM scores of the total group and (2) GREM scores equal to or greater than the mean of the GREM scores of the total group.

Finally, the procedure, Regression, of this statistical package was used to analyze the relationship between the criterion variable, GGPA, and a set of predictor variables, UGPA plus GREM, UGPA plus GREM plus GREAN and UGPA plus GREM plus GREAD.

Results

Table I contains mean scores, standard deviations and maximum and minimum scores of ten of the eleven variables used. This table excludes the variable sex which was not found to relate significantly to any of the variables. An inspection of skewness of the data indicates approximation of the normal curve. As can be seen in Table I, UGPA has a mean of 3.22 with a standard deviation of 0.43. The maximum score was 4.0, while the minimum score was 2.09. The mean score of LTYR was 3.39 with a standard deviation of 0.45, a maximum score of 4.0 and a minimum score of 2.0.

GREV and GREQ were quite similar for this sample. GREV was found to have a mean score of 554.55, a standard deviation of 86.88, a maximum score of 750 and a minimum score of 390. GREQ's mean score was found to be 543.09 with a standard deviation of 80.07, a maximum score of 710 and a minimum score of 330.

GREM (that is, GREV plus GREQ) was found to have a standard deviation of 128.90, a mean of 1100.55, a maximum score of 1430 and a minimum score of 830.

The mean score of GREAN (568.21) was found to be somewhat higher than either GREV, GREQ or GREAD but its standard deviation was somewhat higher (92.90). Its range is slightly less than the range of GREV and GREQ (720-370). The number

of students who had taken the Analytical Exam was only 28 while all 55 had taken both the GREV and the GREQ.

With a sample of only 16 GREAD had a mean score of 536.88, a standard deviation of 50.03, a maximum score of 630 and a minimum score of 450.

The GGPA mean score was found to be 3.60, the standard deviation 0.32, the maximum score 4.0 and the minimum score 2.82. While the mean scores of GPAEXD and GPAPT were somewhat lower than the mean of GGPA (3.13 and 3.26 respectively) the standard deviations were considerably higher (0.80 and 0.77 respectively). These variables also have wider ranges than does GGPA. GPAEXD has a maximum score of 4.0 and a minimum score of 2.0. GPAPT has a maximum score of 4.0 but a minimum score of 1.0. All grade point averages are presented on a five point score: A=4, B=3, C=2, D=1 and F=0.

Table II contains the correlation matrix of the variables. Of 55 non-redundant Pearson Product-Moment correlation coefficients, 19 are significant at or above the .05 alpha level. Of these, seven are at the .001 level, eight at the .01 and four at the .05 level of significance.

As would be expected, Graduate Records Examination scores frequently yield significant correlations with one another. GREM scores correlate significantly with the GREV scores ($r=.77$, $p=.001$), with GREQ scores ($r=.72$, $p=.001$) and with GREAN scores ($r=.36$, $p=.01$). GREAD scores failed to correlate significantly with any other GRE scores.

TABLE I

Mean Scores, Standard Deviations

and Maximum and Minimum Scores

	N	MEAN	SD	MAX	MIN
UGPA	55	3.22	0.43	4.00	2.09
LTYR	53	3.39	0.45	4.00	2.00
GREV	55	554.55	86.88	750.00	390.00
GREQ	55	543.09	80.07	710.00	330.00
GREM	55	1100.55	128.90	1430.00	830.00
GREAN	28	568.21	92.90	720.00	370.00
GREAD	16	536.88	50.03	630.00	450.00
GGPA	55	3.60	0.32	4.00	2.82
GPAEXD	55	3.13	0.80	4.00	2.00
GPAPT	47	3.26	0.77	4.00	1.00

TABLE II

NON-REDUNDANT PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS
(rounded to the nearest hundredth)

	UGPA	LTYR	GREV	GREQ	GREM	GREAN	GREAD	GGPA	GPAEXD	GPAPT	SEX
UGPA	1.00	0.77***	0.03	0.05	0.05	0.30	-0.09	0.36**	0.14	0.45**	-0.08
LTYR		1.00	-0.14	-0.02	-0.09	0.23	0.10	0.24	0.20	0.41**	-0.17
GREV			1.00	0.14	0.77***	0.48+	0.38	0.30*	0.03	0.12	0.05
GREQ				1.00	0.72***	0.36	0.05	0.30*	0.19	0.31*	0.27
GREM					1.00	0.54**	0.22	0.41**	0.17	0.31*	0.18
GREAN						1.00	99.00	0.46***	0.22	0.59***	-0.19
GREAD							1.00	0.36**	0.68***	0.14	-0.07
GGPA								1.00	0.44**	0.80***	-0.17
GPAEXD									1.00	0.43**	-0.23
GPAPT										1.00	0.01
SEX											1.00

*** significant at or beyond the .001 alpha level

** significant at or beyond the .01 alpha level

+ significant at or beyond the .02 alpha level

* significant at or beyond the .05 alpha level

Undergraduate Grade Point Average yielded significant correlations with LTYR (r=.77, p=.001), Graduate Grade Point Average (r=.36, p=.01) and with the Graduate Grade Point Average of Psychometric theory (r=.45, p=.01).

LTYR was found to correlate significantly with one variable other than UGPA. Its correlation with GPAPT was found to be r=.41, p=.01.

The criterion variable GGPA yielded the greatest number of significant correlations. Correlation with GREAN (r=.46) and with GPAPT (r=.80) are significant at the .001 level, while correlations of this variable with UGPA (r=.36) and with the GPAEXD (r=.44) are all significant at the .01 alpha level. At the .05 level of significance, GGPA correlated with GREV (r=.30) and with GREQ (r=.30).

GPAEXD was found to yield significant correlations with GREAD (r=.68, p=.001), GPAPT (r=.43, p=.01) and with GGPA (r=.44, p=.01), while GPAPT was found to correlate significantly with GREAN (r=.59, p=.001), GGPA (r=.44, p=.01), UGPA (r=.45, p=.01), LTYR (r=.41, p=.01), GPAEXD (r=.43, p=.01) and with GREQ and GREM (r=.31, p=.05).

It should be noted that these correlations are not always based upon the entire sample of fifty-five subjects. In correlations involving GREAN, 28 cases were present and in GREAD, only 16; LTYR has a sample of 53 and GPAPT of 47. All other correlations are based upon the entire sample of 55 cases.

Table III contains the classification results of the Discriminant Analysis using GPAEXD (grades received in the

Experimental Design course) as the categorical dependent variable
and GREQ as the continuous independent variable. Table IV contains
the same data generated for the variable GPAPT (Psychometric Theory)
used as a categorical dependent variable and GREQ as the continuous
independent variable. This analysis was done in an effort to estab-
lishthe effectiveness of the GREQ score as a predictor of success in
the two required research methods courses, Experimental Design and
Psychometric Theory. For the purposes of these analyses these grades
were recoded from the original five point scale (A=4, B=3, C=2, D=1,
F=0) into three categories (A=3; B=2; C, D, F=1). This was done
because of the very small number of C, D, and F grades.

As can be seen in Table III the percentage of "grouped" cases
correctly classified in the Experimental Design course was 31.48.
Of the 18 cases actually receiving an A grade, 8 or 44.4 percent were
correctly assigned to Group 3 but 10 (five in each group) were in-
correctly assigned to groups predicted to receive grades of B or below.
Fifty-three percent of those who received B grades in this course
were predicted to receive grades of A while 28.6 percent were assigned
to the group receiving grades of C or below. Only 17.9 percent or
five students were correctly identified here. Of those who
actually received grades of C and below, 50 percent were correctly
identified, 25 percent were assigned to the B group and 25 percent to
the A group. Grades were not efficiently predicted in either of the
analyses. GREQ proved not to be a good measure of performance in these

TABLE III

CLASSIFICATION RESULTS OF DISCRIMINANT
ANALYSIS -- GROUP GPAEXD

ACTUAL GROUP *	NO. OF CASES	PREDICTED GROUP MEMBERSHIP* 1	2	3
Group 1	8	4 50.0%	2 25.0%	2 25.0%
Group 2	29	8 28.6%	5 17.9%	15 53.6%
Group 3	18	5 27.8%	5 27.8%	8 44.4%
Ungrouped Cases	1	0 0.0%	1 100.0%	0 0.0%

* Group Membership (A=3, B=2, C, D, F=1)   Ungrouped cases
have no assigned grade.

Percent of "grouped" cases correctly classified: 31.48%

courses for this sample. In neither case was Wilks' Lambda found to be statistically significant at the .05 level.

As can be seen in Table IV the percentage of "grouped" cases correctly classified in the Psychometric Theory class by the GREQ score was 42.55 percent. Of twenty cases who received grades of A in this course, 12 or 60 percent were predicted correctly on the basis of the GREQ scores but 8 or 35 percent of this group were incorrectly predicted to receive grades of B or below. Of Group 2 (those who actually received grades of B, (N=20), 9 were predicted to receive a grade of A, 5 to receive a grade of B and six to receive a grade of C or below. Thus, of the 20 students who actually received grades of B, only five were correctly identified on the basis of the GREQ score while 15 were incorrectly predicted to receive grades other than B. Of those students who in fact received grades of C or below (N=7) only 3 were predicted on the basis of their GREQ scores to receive these grades. An equal number were predicted to receive B's and one case was placed in Group 3 (ie, to receive a grade of A).

Next, the sample of fifty-five subjects was divided into two subgroups, those with GREM scores equal to or greater than the mean GREM score for the total group and those whose GREM scores fell below the mean GREM score for the total group. A Pearson Correlation was used to generate Person Product-Moment correlation coefficients for UGPA with GGPA for each of the two subgroups. This was done to test the relationship of UGPA and relatively higher and lower GREM scores to GGPA.

TABLE IV

CLASSIFICATION RESULTS OF DISCRIMINANT
ANALYSIS -- GROUP GPAPT

ACTUAL GROUP *	NO. OF CASES	PREDICTED GROUP MEMBERSHIP*		
		1	2	3
Group 1	7	3 42.9%	3 42.9%	1 14.3%
Group 2	20	6 30.0%	5 25.0%	9 45.0%
Group 3	20	5 25.0%	3 15.0%	12 60.0%
Ungrouped Cases	8	3 37.5%	2 25.0%	3 37.5%

* Group Membership (A=3, B=2, C, D, F=1) Ungrouped cases
have no assigned grade.

Percent of "grouped" cases correctly classified: 42.55%

There is a strong relationship between UGPA and GGPA for those students whose GREM scores are 1100.545 or above ($r = .53, N = 26$, $p = .002$), but there also exists a less strong but statistically significant relationship between these two variables for those students whose GREM scores fall below the group mean ($r = .30$, $N = 29$, $p = .05$), which appears to indicate that the better the GREM score the stronger the predictive relationship of UGPA is to GGPA.

When a Stepwise Multiple Regression procedure was used to investigate the relative contributions of UGPA and GREM to the prediction of GGPA, a multiple R of .41, R^2 of .16 and adjusted R^2 of .15 were generated by Step 1, GREM on GGPA. UGPA was entered as Step number 2 and yielded a multiple R of .53, R^2 of .28 and Adjusted R^2 of .25. $F = 10.82$ and 10.28 respectively for Steps 1 and 2. These are significant beyond the .001 level. Both GREM and UGPA appear to be strongly and significantly related to GGPA. Beta weights of .39 and .33 respectively suggest that their contributions are approximately equal.

Finally, UGPA, GREM, GREAN and GREAD were transformed into standard scores. This was done in an effort to minimize the danger of overfitting and shrinkage which typically occur when multiple regression and certain other statistical procedures are used with samples having small N's. Dunnette and Borman (1979) address this issue.

> "When we use more than one predictor and wish to combine this information optimally to gain maximum accuracy for predicting scores on a single criterion,

> any weights we select based on sample data (e.g.
> regression weights) will take advantage of sample-
> specific configuations of the data, that is, will
> 'overfit' the data; thus the validity obtained in
> the sample provides an overestimate of the validity
> to be expected in the long run."(p.491)

Transformation to standard scores gave both UGPA and GRE
scores equal weighting. These transformed scores were desig-
nated ZUGPA, ZGREM, ZGREAN, and ZGREAD. A Pearson Correlation
procedure was then used to generate Product-Moment correlation
coefficients of various combinations of these scores with GGPA.
Combinations used were ZGPA plus ZGREM, ZGPA plus ZGREM plus
ZGREAN, and ZUGPA plus ZGREM plus ZGREAD. These combinations
were renamed as follows:

FIRST   ZUGPA plus ZGREM

SECOND   ZUGPA plus ZGREM plus ZGREAN

THIRD   ZUGPA plus ZGREM plus ZGREAD

The correlation of GGPA with FIRST yielded a coefficient
of $r=.53$ (N=55, p=.001); with SECOND, $r=.65$ (N=28, $p<.001$)
and with THIRD, $r=.46$ (N=16, p=.036). Thus it appears that
the relationship of UGPA plus GREM may be slightly improved
by the addition of the GREAN score but weakened by the addi-
tion of the GREAD score. It must be noted that the small
numbers of subjects having GREAN (N=28) and GREAD (N=16) make
interpretation of this finding a cautious one.

Discussion

This study was undertaken in an effort to establish the validity of those predictor variables presently in use at Western Kentucky University in selecting students for its Master of Arts degree programs in Psychology and to identify the most valid variable or combination of variables for future use in student selection. Those predictor variables presently in use include the UGPA and GRE scores in addition to letters of reference and some intuitive "feel" on the part of the committee members which probably takes into account such subjective factors as quality of the undergraduate institution, dedication to the field, potential and motivation on the part of the applicant.

A review of the literature showed that UGPA (cumulative or some portion thereof) and GRE scores are most commonly used by graduate programs in the selection of graduate students and that they probably represent a useful selection battery, though consensus is far from unanimous. Letters of reference and the quality of the undergraduate institution appear, for the most part, to be useless in their present forms. Therefore, they were eliminated from this study. As far as motivation, dedication and potential are concerned, it is commonly agreed that while these are important factors in "success" at any level, present measures of such personal characteristics

55

are not sufficiently sophisticated to make them useful as valid predictor measures. For this reason, only UGPA, LTYR and GRE scores as predictors of success in this graduate program were considered. "Success" is defined, for the purposes of this study, as GGPA for the first two semesters of the graduate program and by grades received in the two required statistical courses, Psychometric Theory and Experimental Design.

Hypothesis I of this study states that UGPA (cumulative grade point average in undergraduate work) will predict GGPA as efficiently as will LTYR (grade point average during the last two years of undergraduate work). This hypothesis appears to have been supported by the results of the study.

The correlation of UGPA with GGPA (r=.36) is statistically significant at the .01 alpha level but the correlation of LTYR with GGPA is .24 (not significant at the .05 level). It is important to note that these variables bear no significant relationship to any of the predictor variables other than to one another (r=.77, p=.001). It would be expected that UGPA and LTYR would be highly correlated. It appears from the results of this study that UGPA, LTYR and GRE scores are relatively independent of one another for this sample.

Both UGPA and LTYR are significantly related to grades received in the Psychometric Theory course (r=.45, p=.01, r=.41, p=.01 respectively) but neither are significantly related to grades received in Experimental Design.

In view of these findings it appears that there would be
no advantage in using LTYR rather than UGPA as a pre-selection
variable for this sample. Thus, the first hypothesis was
accepted and UGPA rather than LTYR was included in the linear
combination correlations and regression analyses.

The low but statistically significant correlation of
UGPA with GGPA is consistent with much of the literature. No
doubt the failure of these variables to correlate more highly
is explained at least in part by the unexplained variance of
which Chansky (1964) speaks. Differences in grading practices
and curricular expectations from one undergraduate institution
to another, and even from one course to another within the
same school, undoubtedly renders the grade point average less
valid than one would wish.

Hypothesis II concerns the predictive validity of the
Graduate Records Examination Quantitative score in relation
to grades received in the two required research methods courses,
Psychometric Theory (GPAPT) and Experimental Design (GPAEXD).
The correlation of GREQ with GPAPT is significant at the .05
alpha level (r=.31). Interestingly, the correlation of GREQ
with GPAEXD is non-significant (r=.19). This is a puzzling
result in view of the literature which appears to suggest that
GREQ is frequently found to be a good predictor of grades in
statistics and research methods courses (Willingham, 1974;
Bean, 1975; Lannholm, 1968; Hirshberg and Itkin, 1978).

A Discriminant Analysis using two groups (those defined
by GPAPT and those defined by GPAEXD) and the variable GREQ

was performed in an effort to clarify the relationship of GREQ to grades received in these courses. For the group defined by GPAPT (N=47) the percent of grouped cases correctly classified according to grades actually received was 42.55. This function was reported to be non-significant. For the group defined by GPAEXD (N=54) the precent of grouped cases correctly classified was 31.48. This function also was determined to be non-significant. The total number of cases in which grades were correctly predicted on the basis of the GREQ scores for GPAPT was 20 while 27 were incorrectly predicted. With reference to the GPAEXD group, only 17 cases were predicted to receive the grades actually received in the course and 37 were predicted to receive some grade other than the one actually received.

These results indicate that, had GREQ scores been used as a measure of an applicant's ability to perform adequately in the two required statistics and research methods courses (that is to receive grades above C), 14 applicants would habe been rejected because of their inability to receive an A or B in Psychometric Theory. In fact, only 7 were actually unsuccessful. If this variable had been used to select students based upon their ability to receive grades of A or B in the Experimental Design course, 17 would have been rejected while only 8 actually failed to perform adequately. On the basis of this analysis, Hypothesis II was rejected. The use of the GREQ score as a predictor of grades in the two research methods and statistics courses does not appear justified.

While this finding is somewhat atypical, it is not with-
out precedent in the literature. Borg (1963) concluded that
using GREV scores with a cut-off of one-half standard devia-
tion below the mean as a selection criteria for M.A. level
students would have eliminated 41 students who were actually
successful and 21 who were not successful. His criterion was
GGPA rather than grades in specific courses, but a similar
procedure was used in an effort to further support the rejec-
tion of Hypothesis II of this study.

As can be seen in Table V, if a cut-off score of one
standard deviation below the GREQ mean score for this sample
were used to predict failure in the two statistics courses,
the cut-off point would be 463.03. This score would have
eliminated a total of 14 applicants; 9 of those were actually
successful while only 5 of those eliminated were unsuccessful
in the required courses GPAPT and GPAEXD. If the cut-off
point had been established at one-half standard deviation
below the GREQ mean score for the total sample, it would then
be 503.07. Nine students would then be eliminated, 3
correctly but 6 incorrectly. It can only be concluded that
for this sample the GREQ score is not a valid predictor of
grades in the two required statistics and research methods
courses. One possible explanation may lie in the fact that
during the three-year period of this sample's participation
in the graduate program, each of the two courses had more
than one instructor. At no time were the two courses taught
by the same instructor, and the Psychometric Theory course was

TABLE V

Number of Successful Students Predicted to be Unsuccessful
in Statistics Courses With GREQ Cut-Off Intervals of One
Standard Deviation and One-Half Standard Deviation Below
the Mean GREQ Score of the Total Sample.

GREQ cut-off point	No. of students predicted to be unsuccessful in statistics courses	No. of students actually unsuccessful in statistics courses	No. of successful students incorrectly identified
1 SD (463.03)	14	5	9
½ SD (503.07)	9	3	6

never taught by the instructors who taught Experimental Design or vise versa. Once again, grading practices and curricular expectations may have been quite different and this uncontrolled variance may have resulted in the failure of GREQ to adequately predict grades received in the two courses. The fact that GPAPT is correlated significantly with GREAN (r=.59, p=.001) but not with GREAD while GPAEXD correlates significantly with GREAD (r=.68, p=.001) but not with GREAN appears to indicate that the two courses are quite different either in content or in instructor expectation. However both are correlated with GREQ at the .05 alpha level (r=.31 and .30), which would appear to indicate that they do have 9 percent shared variance. (GREAN does not correlate significantly with GREAD).

Hypothesis III states that a linear combination of UGPA and GRE Verbal plus GRE Quantitative (GREM) will prove to be a valid predictor of GGPA over the first two semesters of graduate work in the M.A. level psychology programs at Western Kentucky University. The findings do appear to support this hypothesis. Both UGPA and GREM are significantly correlated with GGPA (r=.36 and r=.41, p=.01). Each variable is also correlated at a significant level with the grades received in GPAPT. UGPA is significantly correlated with GPAPT at the .01 alpha level (r=.45) while GREM yields a significant correlation with GPAPT at the .05 level of alpha (r=.31).

UGPA fails to show a statistically significant correlation with GREM (r=.05). These variables appear to be relatively

independent of one another. This finding lends support to
the use of the two variables combined as predictors of success.
They appear to relate to GGPA in different ways and without an
unacceptable degree of overlap.

In an effort to determine the relationship of high GREM
scores (those at or above the mean GREM score for the total
group) and UGPA to GGPA as well as that of lower GREM scores
(those below the sample GREM mean score) the sample was divided
into two subgroups and the correlations of UGPA with GGPA for
each group were computed. The results of this procedure indi-
cate that the relationship of UGPA to GGPA for those students
with high GREM scores (N=26) is significant beyond the .001
level (r=.53). For students whose GREM scores fall below the
GREM mean score for the total group (N=29), there also exists
a significant but less strong relationship (r=.30, p=.05).
These findings appear to suggest that UGPA maintains a rela-
tionship with GGPA both with high GREM scores and with low
GREM scores but that this relationship becomes increasingly
strong as the level of the GRE approaches and surpasses the
mean of the sample. In other words, it appears that for a
student with a high GREM score, UGPA would be more strongly
predictive of GGPA than would the UGPA of a student with a
lower GREM score. In any event, it seems that undergraduate
grades do predict graduate grades for this sample of subjects
and that the addition of the GRE Verbal plus Quantitative
score may supplement this prediction by indicating the
strength of the UGPA to GGPA relationship.

A multiple regression procedure was used to regress UGPA and GREM to GGPA. This indicated that both GREM and UGPA are significantly related to GGPA ($R=.41$ and $.53$, $p=.001$). The Beta weights of $.39$ for GREM and $.33$ for UGPA further indicate that these variables are relatively equal in their contributions to the prediction of GGPA.

However, when multiple regression procedures are used with samples containing small numbers of subjects, there is a danger of overfitting the data (Dunnette and Borman, 1979). In an effort to minimize this error, UGPA, GREM, GREAN and GREAD scores were transformed to standard scores so that each score would have equal weighting. The correlation of UGPA plus GREM with GGPA using these transformed scores was found to be significant beyond the $.001$ level ($r=.53$), the same value obtained with the multiple regression analysis. Thus it appears that for this sample of graduate students in psychology at Western Kentucky University, the use of UGPA plus GREM as a predictor set in the prediction of GGPA is highly justified.

One additional factor in this justification is the fact that the reliability of the criterion measure (GGPA) was found to be extremely high. When first semester grades were correlated with second semester grades there was found to be a correlation of $r=.85$, $p<.001$).

Hypotheses IV concerns the merits of the GRE Advanced Psychology (GREAD) and GRE Analytical (GREAN) scores as additional variables in the prediction of GGPA. The correlation

of UGPA plus GREM with GGPA, using transformed scores, was increased from r=.53 to r=.66 when the GREAN score was added. This correlation is significant beyond the .001 level of alpha. It indicates that the prediction of GGPA could be improved by the use of a predictor set which included not only UGPA and GREM but GREAN as well.

When the GREAD score was added to the UGPA, GREM predictor set, the correlation of the set with GGPA fell from the original r=.53 to r=.46 (p=.03). This is a puzzling development particularily in view of the findings of previous studies which indicate that the Advanced Psychology portion of the GRE is frequently found to be the best predictor of GGPA (Mehrabian, 1969; Lannholm, 1968; Hirschberg and Itkin, 1978). A careful inspection of the data reveals that of the total sample of 55 used in the present study, only 16 had taken the Advanced Psychology GRE. In addition, the GREM mean score for this group of subjects when transformed to standard scores is such that it becomes slightly less than zero and therefore produces a negative correlation, reducing the predictive validity of the predictor set. It appears that these results may be an artifact of the small sample and its individual configuration.

For example, the GREV, GREQ and GREAN scores are mandatory but, for this sample, the GREAD was never required. Therefore, it is difficult to know what motivated those students who took the Graduate Record Examination Advanced Psychology Test when they were not required to. After all, thirty-nine of the fifty-five students in the sample did not

take this test. Perhaps those students who did elect to take the
test, even though it was not required, did so in an effort to
compensate for some real, or perhaps imagined, weakness in their
required scores and other credentials. This is, of course, only
speculation but it does point out the very real possibility that
the sample of 16 students who did take the Advanced test may, in
fact, differ in some way from the remaining 39 students and cannot
be considered completely representative. Further research is necessary
in order to rule out this possibility but, on the basis of the
present data, the addition of the GREAD score would not appear to
add to the validity of the predictor set, UGPA plus GREM, for this
sample of students.

The variable sex was entered into the analysis but was found
to be nonsignificant in its relationship to the other variables.
This seems to be consistent with the findings of Hirschberg and Itkin
(1978) that males and females do not differ in ability to succeed.
If, as they suggest, females do not finish graduate programs as often
as do males, other factors appear to be at work.

In interpreting these findings, it must be remembered that the
usual problems of attenuation and grade inflation both in pre-selection
predictor variables and in criterion variables are also present in this
study. Both UGPA and GRE scores are attenuated because this sample
of students was selected upon the basis of these scores.

No effort was made to look at applicants who were not accepted
and how they may have differed from those who were accepted. The
subjects of this study were selected because the members of the selection

committee of this program had chosen those applicants with
the higher scores on the predictor measures and rejected
those who fell below some unspecified, consenually agreed
upon minimum score. Thus, the range is fairly restricted.
In addition, it must be recognized that those students with
extremely high scores are most often accepted by Ph.D. pro-
grams and, therefore, are not available to this program.
This factor also contributes to the restriction of range of the
GRE scores. Further, the GGPA mean score of 3.60 indicates
that the majority of grades given in the program range from
A to B but the range, 4.00 to 2.82 also indicates that A-B
grades are not assured. Given the expectation on the part of
the selection committee members and possibly on the part of
other faculty members as well, that accepted applicants are
those who "should" succeed, it is likely that factors other
than demonstrated ability and past achievement most often
account for failure to succeed. Of course, these factors
have not been the focus of this study but it would appear
that if more valid selection procedures are to be developed,
some investigation of them must be undertaken.

However, given the strength of the UGPA/GREM predictor
set and the lack of significant correlation between these
variables it seems likely that UGPA may be viewed not only
as a measure of academic ability and past achievement but
of personal characteristics such as motivation, tenacity,
dedication and determination as well. Perhaps there has been
too little thought given in the past to what those predictor

variables commonly used but frequently criticized are actually measuring. It may well be that the search for new predictor variables should begin with a thorough investigation of the old ones.

Given the available data, it does appear that UGPA and GREM plus GREAN scores do a reasonable job of selecting students who are capable of success in the Master of Arts degree programs in the Psychology Department at Western Kentucky University and their use is highly justified in the absence of more valid and inclusive predictor variables.

References

Ayers, Jerry B. Predicting quality point averages in master's
degree programs in education. Educational and Psychological Measurement, 1971, 31, 491-495.

Baird, Leonard L. Biographical and educational correlates
of graduate and professional school admissions test
scores. Educational and Psychological Measurement, 1976, 36, (2), 415-420.

Bean, Andrew The prediction of performance in an educational
psychology master's degree program. Educational and
Psychological Measurement, 1975, 35, 963-967.

Borg, W. R. GRE aptitude scores as predictors of GPA for
graduate students in education. Educational and Psychological Measurement, 1963, 23, (2) 379-382.

Chansky, Norman A note on the grade point average in research.
Educational and Psychological Measurement, 1964, 24, (1), 95-98.

Dawes, Robyn M. A case study of graduate admissions: Application of three principles of human decision making.
American Psychologist, 1971, 26, 180-188.

Dunnette, M. D., Borman, W. C. Personnel selection and classification systems. Annual Review of Psychology, 1979, 30, 477-525.

Eckhoff, Constance M. Predicting graduate success at Winona
State College. Educational and Psychological Measurement, 1966, 26, 483-485.

Ewen, Robert B. The GRE psychology test as an unobtrusive
     measure of motivation. Journal of Applied Psychology,
     1969, 53, (5), 383-387.

Forest, Robert F. The undergraduate grade point average as
     a predictor of the graduate index of students in the
     leadership-administration emphasis in the master of arts
     in education program at Castleton State College, In
     Education Research Information Center (ERIC), Castleton
     State College, 1979, File No. Ed 160 007.

Goldberg, L. R. Man versus model of man: A rationale, plus
     some evidence for a method of improving on clinical infer-
     ences. Psychological Bulletin, 1970, 73, 422-434.

Goldberg, Lewis R. Admissions to the Ph.D. program in the
     department of psychology at the University of Oregon.
     American Psychologist, 1977, 32, 663-668.

Hartogson, Y., Trainer, J., Chansky, N.M. Predicting grades
     of special education majors. Psychological Reports,
     1978, 42, (1), 290.

Heritage, John O. A validity study of admissions criteria
     for masters students in a reading program. In Educational
     Research Information Center (ERIC), Rutgers The State
     University of New Jersey, 1977, File No. 144 035.

Hirschberg, Nancy and Itkin, Stuart Graduate student success
     in psychology. American Psychologist, 1978, 33, (12),
     1083-1093.

Humphreys, L. G., Taber, Thomas Postdiction study of the
     graduate record examination and eight semesters of col-
     lege grades. Journal of Educational Measurement, 1973,
     10, (3), 179-184.

Kelsey, J. W., Dobson, W. R.   On selecting graduate psychology students: A tentative proposal. In Educational Research Information Center (ERIC), 1977, File No. ED 141 414.

Lafferty, Gladys E.   A study of the influence of age on predictability of Graduate Record Examinations aptitude tests for successful graduate students. In Educational Research Information Center (ERIC), South Carolina School of Education, 1969, File No. ED 029 606.

Lannholm, Gerald V. Cooperative studies of predicting graduate school success. Graduate Record Examinations Special Report Number 68-3, In Educational Research Information Center (ERIC), Educational Testing Service, Princeton, New Jersey, 1968, File No. ED 163 083.

Lannholm, Gerald V. Review of studies employing GRE scores in predicting success in graduate study 1952-1967. In Educational Research Information Center (ERIC), Educational Testing Service, Princeton, New Jersey, 1968, File No. ED 163 084.

Madaus, G., Walsh, John J. Departmental differentials in the predictive validity of the graduate record examination aptitude tests. Educational and Psychological Measurement, 1965 , 25, (4), 1105-1110.

Mehrabian, Albert Undergraduate ability factors in relationship to graduate performance. Educational and Psychological Measurement, 1969, 29, 409-419.

Merenda, P. F., Reilly, R. Validity of selection criteria in determining success of graduate students in psychology. Psychological Reports, 1971, 28, 259-266.

Newman, Richard I. GRE scores as predictors of GPA for psychology graduate students. Educational and Psychological Measurement, 1968, 28, 433-436.

Robertson, M., Hall, E. Predicting success in graduate study. The Journal of General Psychology, 1964, 71, 359-365.

Robertson, M., Nielsen, W. The graduate record examination and selection of graduate students. American Psychologist, 1961, 16, (10), 648-650.

Schmidt, F. L., Hunter, J. E. Development of a general solution to the problem of validity generalization. Journal of Applied Psychology, 1977, 62, 529-540.

Shannon, R. S. and others Factors related to success in completing a master's degree at Stout State University 1960-1968. In Educational Research Information Center (ERIC), Stout State University, 1969, File No. ED 0506 78.

Stordahl, Kalmer E. Predictive validity of the graduate record aptitude test. In Educational Research Information Center (ERIC), Northern Michigan University, 1970, File No. ED 044 071.

Sticker, George and Huber, J. T. The graduate record examination and undergraduate grades as predictors of success in graduate school. The Journal of Educational Research, 1967, 60, (10), 466-468.

Wiggins, Nancy, Blackburn, Margaret, Hackman, J. R. Prediction of first year graduate success in psychology: Peer ratings. The Journal of Educational Research, 1969, 63, 81-85.

Wiggins, Nancy, Kohen, Eileen S. Man versus model of man revisited: The forecasting of graduate school success. Journal of Personality and Social Psychology, 1971, 19, (1), 100-106.