ABSTRACT

A review of current procedures for selecting students for graduate programs in psychology reveals a number of problems due to predictor criterion, and methodological variables. The idea is advanced that with the multiplicity of procedures used, plus the number of students applying to graduate programs in psychology, the effort on the part of individual selection committees to differentiate students on a number of relevant variables is both time consuming and not sufficiently powerful to be able to make meaningful statements about each applicant's relative potential. A tentative proposal to establish a centralized registry of psychology graduate students is made. A centralized registry might be more efficient and economical and also would establish a research data pool which could provide a more reliable means for identifying and matriculating graduate students who will make significant contributions to psychology. (Author/EVH)
On Selecting Graduate Psychology Students:
A Tentative Proposal
John W. Kelsey and William R. Dobson
Utah State University
Abstract

A review concerning the selection of students for graduate programs in psychology notes a number of problems in the present procedures due to predictor, criterion, and methodological variables. The idea is advanced that with the multiplicity of procedures used, plus the number of students applying to graduate schools in psychology, the effort on the part of individual selection committees to differentiate students on a number of relevant variables is both time-consuming and not sufficiently powerful to be able to make meaningful statements about each applicant's relative potential. A tentative proposal to establish a centralized registry of psychology graduate students is made which hopefully would be more efficient and economical, but also would establish a research data pool which could well provide a more reliable means for identifying and matriculating graduate students who will make significant contributions to psychology.
On Selecting Graduate Psychology Students:
A Tentative Proposal

The selection process concerning candidates for graduate degrees in psychology is only the beginning of an enormous expenditure of time and money on the part of the applicants, selection committees, and for those candidates who matriculate, the graduate faculty. It would seem imperative to select those students who have an optimal chance of succeeding within a program, and of making subsequent contributions to the field. The literature to date has been suggestive of numerous variables or combinations of variables which may be useful for predicting future success, but which have consistently proven inadequate and/or less than optimally reliable in selecting the "best" candidates. Consequently, selection committees each year choose the required number of applicants, hoping that they have offered admittance to the most qualified candidates by the best available method, but never really being certain that they have done so. Historically, certain predictors are used for a period of time but are supplanted when new predictors are purported to be more effective. Comparatively little concern is given to the criterion variables, or methodology used to determine the efficacy of any one selection method.

Successful selection is related to a number of factors which revolve around the issue of validity and methodology. Therefore, predictor and criterion variables must be related in such a way that useful information is obtained with relative ease if results are to be duplicated for candidates in succeeding years. In general, studies of selection procedures often attempt to demonstrate the efficacy of single predictor variables.
(i.e. GPA, GRE, MAT, personality test scores, etc.) by inter-comparison with other single variable research. Also, criterion variables which are easy to collect (i.e. practicum instructor ratings, or first year graduate grades) are often used. Longitudinal criterion variable research in the area of graduate selection is exceedingly rare, or is conducted by \textit{ex post facto} designs. A final factor which also must be considered is the consequences of research methodologies which have constricted samples, and which therefore bias the findings and conclusions that the research purports to show.

This article will first review and critique previous variables and methodologies used in the study of the selection process in graduate psychology, and to suggest ways to increase the predictive validity of further research in this area. Finally, a tentative proposal will be offered, which if implemented would hopefully supply a system whereby the selection process would be considerably less time consuming for all concerned, and which would also begin a data pool for further research and development. The ultimate goal projected for the proposed procedure would be to increase the probability of selecting those candidates who have the greatest chance of making significant contributions to both the service and research components of psychology.

\textbf{Predictor Variables}

Numerous predictor variables have been used to select applicants believed most capable for matriculation into graduate psychology programs. Variables upon which selection has typically been based have included some form of undergraduate grades, standardized tests, letters of recommendation,
biographical information, essays, and interviews. As will be mentioned in a later section, new tests have been suggested, along with an assessment of various skills a potential professional psychologist might need to possess. For the most part no single variable approaches a degree of prediction where the variable alone could be usefully applied. Where some of the newly suggested factors are significantly related to success, generally they are cumbersome or time consuming than one could realistically expect either applicant or selection committee to use. For the most part selection committees continue to use the aforementioned variables singly or in various and often unique combinations.

Grades and Standardized Tests as Predictors

The most widely used factors in the selection of psychology graduate students are some combination of undergraduate grade point average (GPA), and standardized test scores. The underlying assumption for their usage seems to be that selection committees need both a measure of past achievement and of current scholastic aptitude or ability to differentiate candidates. These seem to be logical criteria.

The most popular method of using undergraduate grades as a predictor utilizes total undergraduate GPA (Hurst, 1974; Dawes, 1971; Merenda & Reilly, 1971; Gertler, 1970; Hoyt, 1966), or a combination of undergraduate GPA and undergraduate psychology GPA (Ewen, 1969). Siegel, Klein, and Ritigstein (1968), in a poll of 86 school psychology programs, found total undergraduate GPA and undergraduate psychology GPA ranked 1 and 2 in terms of use by selection committees. Some studies have used other combinations, however, such as GPA for the last two years of undergraduate study (Mehrabian, 1969; Allen, 1967; and Robertson & Nielsen, 1961), as well as GPA in
Various specific course areas other than psychology (Rawls, Rawls & Harrison, 1969). In most cases, depending on the criterion variable used, correlations have been low and randomly significant (r's of -0.2 to .34).

Standardized tests, basically the Graduate Record Examination (GRE) or Miller Analogies Test (MAT), constitute another major means of predicting graduate success. Siegel, Klein, and Ritigstein (1968) found that 73% of the schools polled used GRE scores, while only 45% listed the MAT. GRE scores have been correlated in various combinations of using, as the predictor, Total, Verbal, Quantitative, and/or mean of Verbal, Quantitative, and Advanced (psychology) scores (Mehrabian, 1969; Rawls, Rawls & Harrison, 1969; and Borg, 1963). Reported correlations with the criterion variable have ranged from .08 to .64. Lannholm (1968) noted that generally the Advanced Test in psychology has been somewhat the best predictor, and Willingham (1973) in a review of 43 studies in nine different disciplines indicated that the GRE Advanced Tests were the most valid predictors of success. At least for the Advanced (psychology) Test this was explained by Ewen (1969) in terms of the test being an unobtrusive measure of the motivation of the applicant, since a good score can generally be obtained by studying the jargon of the discipline. Thus, it is assumed that a motivated student will spend time prior to the examination possibly reading one of the available dictionaries of psychology.

Although used less frequently at present, the MAT was widely required by selection committees during the 1950's. The MAT has fared poorly, however, in that studies have generally reported non-significant correlations with the criterion variables chosen (Rawls, Rawls & Harrison, 1969; Robertson & Hall, 1964; and Watters & Patterson, 1953). A unique use of
the MAT was noted by Ewe'n (1969) in which Advanced (psychology) Test GRE scores minus MAT scores were found to be significantly correlated ($r = .49$) with graduate percentage of A grades, but no theoretical explanation of the phenomenon was presented.

Letters of Recommendation, Biographical Information, and Essays as Predictors

Another broad category of predictors encompasses what may be termed a description of the candidate. Letters of recommendation are widely used by selection committees, ranking third in use (Siegel, Klein & Ritigstein, 1969), but were found to be non-significantly correlated, in any combination, to any success criteria (Rawls, Rawls, & Harrison, 1969). One might, therefore, conjecture that most committees only require letters of recommendation on the off-chance that a bad letter (assuming that people only ask for letters from those they perceive would give them a good letter) would drastically aid the committee in eliminating some candidates from consideration. And one can only assume that recent legal mandates regarding rights of individuals to review their school records and letters of recommendation written about them will result in even less credence being given to letters of reference as a useable predictor.

Biographical data (a seemingly important variable considering the role it plays in psychological diagnosis) and essays have been examined by Rawls, Rawls, and Harrison (1969) in research on forty separate predictor variables. They concluded that a significant predictor of Ph.D. attainment was the age of the applicant (older), whether or not applicant had already obtained a Master's degree (favoring a yes answer), and if applicant was married (favoring a yes answer). The essay "Why I want to go to graduate
school in psychology" had no predictive validity, although the argument can be raised that it be used only when spelling or language usage are particularly bad. Of course this assumes that meticulousness in checking one's spelling and grammatical usage in personal essays is a desirable trait in prospective graduate students.

The Interview as a Predictor

Siegel, Klein, and Ritigstein (1968) reported that 51% of school psychology programs utilize an interview in the selection procedures. Munday (1968) found that 90% of some 42 businesses polled rated the interview as the most effective technique. Ulrich and Trumbo (1965), however, in a comprehensive review of the selection interview for the preceding 20 years, concluded earlier that the interview is generally unreliable. They did report that "...greatest gains in validity over other predictors involved interviews described as systematic, designed, structured, or guided" and "limited in purpose" (p. 100). Two areas which they identified as relevant and with greatest validity for the selection interview were measures of the applicant's motivation and personal relations ("Will applicant fit into the social context of the job?"). Personal relations seems to have a drawback in that a selection committee could have legal problems if this were the only basis for rejecting a candidate. Ulrich and Trumbo also concluded that ancillary data can best be collected outside the interview with a considerable saving of time. Further, Springbett (1958) found that an interviewer is usually prepared to make a decision after only 4 minutes, although the average interviewer makes his decision on only certain cues, and that the same cues could be rated in some other fashion. A recent study by Broadhurst (1974), investigating the reliability of the
Interview for selecting students for post-graduate work in clinical psychology, found no significant relationship between interviewer ratings and selection committee ratings using ancillary information.

Although the reliability of the interview is in question, except for possible assessment of the applicant's motivation and personal relations, two methods have been applied to obtain ratings of the interviewee. Asher (1970) has determined that question-by-question ratings which disregard previous answers are more reliable than global ratings. This method, however, would be exceedingly cumbersome to apply to large numbers of interviewees, as one could expect the time frame of the interview to expand greatly as the interviewer ponders what score to give a previous answer.

Directly relevant to the question of selection of psychologists in professional service settings is the development of the Psychology Intern Rating Scale (Plutchik, Klein & Contes, 1960), which evaluates, in a systematic manner, the relevant factors suggested by Ulrich and Trumbo (1965). Although promising, the presented reliability data is, at best, open to interpretation. Some of the factors which they found could be significantly ascertained in a short interview included: "Tendency to complain", "Ability to work with staff", "Sensitivity to others", "Intellectual grasp of psychology", and seven other factors. Ten non-significant discriminators were found, including rated items such as "Mood", "Creativity", "Openness", "Personal appearance", "Anxiety", etc.

Personality, Skills and Abilities as Predictors

Traditionally, graduate students in psychology have been selected, as have students in other disciplines, by using past academic performance factors as a logical indicator of future academic success. More recently,
with the increased awareness that many graduate students in psychology become service oriented rather than research oriented, the possibility of determining what characteristics need to be possessed by a professional psychologist has been researched. For example, Coombs, Avila, and Purkey (1971) and Carkhuff (1969) have suggested that the more effective counselors possess personal traits or skills related to empathy, genuineness, openness, respect, and other personality factors necessary for client growth, and that procedures for selection of potential counselors should take into account the potential value of utilizing these factors.

Hurst and Shatkin, (1974), and Anthony and Wain (1971) have used the Counseling Simulation Inventory devised by Carkhuff with some success for selection of potential counselors and "helpers." Another potentially useful instrument, which can be mailed to prospective counselors and which measures empathy through a Trainability Index, was devised by Anthony, Gormally, and Miller (1974). The usefulness of a measure of "openness" on the part of a counselor has been demonstrated in two studies by Allen (1972; 1967) using the Rorschach Index of Repressive Style, the Group Supervisor Report Scale, the Truax Scales to measure "openness," and the Supervisor Rating of General Competence to measure "effectiveness."

The major drawback of the above, as in other promising predictors, is the need to look at a number of instruments at the same time, thus magnifying the time frame of the selection process. Additionally, Allen found that GRE/MAT scores and GPA (freshman or senior) predictors were not significantly related to "openness." Other studies have suggested as possible predictors the Test of Social Intelligence (Osipow and Walsh, 1973), the concept of cognitive-flexibility predicted from TAT and Rorschach scores.
On Selecting

(Whitely, Sprinthall, Masher & Donagly, 1967), and the Personal Orientation Inventory (Melchers, 1972). Although low correlations have been observed, significant relationships using an alpha = .10 have been found between some scales of the California Psychological Inventory (Sociability, Achievement via Independence, Achievement via Conformity, and Psychological Mindedness) and practicum instructor ratings of counseling success (Puranajoti, 1972). However the high alpha of .10 does call these particular findings into question.

Criterion Variables

The above predictor variables can only be useful in terms of the way in which success is defined through the criterion variable. Since success is often complex to define, it would seem important to evaluate predictor variables against pertinent criteria such as effectiveness and attainment within the profession of psychology. Also, criteria which evaluate success over long periods would logically be better measures. Although some attempts have been made to collect longitudinal data, most research in this area has tended to use short-term criteria.

Short-term criteria typically fall into the categories of either grades or practicum evaluations. Graduate grade point average is most frequently used (Ewen, 1969; Borg, 1963; Hyman, 1957), but since graduate grades are generally restricted to letter grades of A or B, this criterion is extremely homogeneous across subjects (Angoff, 1969; Schagrin, 1969). Practicum evaluation (Melchers, 1972; Whitely, et al., 1967) and faculty ratings (Robertson & Hall, 1964; and Watters & Patterson, 1953) are also used, but these have the drawback that they evaluate the subject at a fixed
point of time, often close to the beginning of graduate training, and therefore do not take into consideration later learning and experience which would significantly influence the trainee's ultimate effectiveness and success.

Although more difficult to collect, longitudinal criteria assess the subject's success in more general terms. Degree attainment (Altman, 1971; Merenda & Reilly, 1974) evaluates whether or not a candidate successfully completes the necessary requirements of an academic program. Research productivity (Mehrabian, 1969; Hoyt, 1966) relates one aspect of professional success which is, by definition, part of the graduate degree. Hoyt's study (1966) found no significant correlation between quality or quantity of an individual's research and his/her undergraduate GPA.

Although faculty ratings or practicum evaluations may be logically related to the selection of potential counselors, Sagedahl, Lesar, and Markwardt (1969), in polling 69 school district superintendents, found that work experience variables and perceived compatibility with the staff ranked higher in hiring practices than practicum supervisors' recommendations (4th), practicum grades (10th), or grades in counseling courses (14th). It should be noted that the top-rated variable concerned with hiring later in one's career is directly related to the personal relations predictor suggested by Ulrich and Trumbo (1965).

Methodological Variables

In general, research of this nature is difficult to conduct even if the proper predictor and criterion variables have been selected. Restriction of results applicability often occurs because of the nature of the
available subjects, and the significance of the calculated correlations. Most studies have used subjects pooled from candidates already accepted into a graduate program. Thus the possible range of the predictor variables are typically compressed, due to the homogeneous nature of selected candidates. In one study where all applicants were allowed to enter a graduate program (Merenda & Reilly, 1971), it was found that undergraduate GPA, GPA in undergraduate psychology, GRE Verbal, Quantitative, and Advanced (psychology) scores, and a rating of the undergraduate institution were highest for graduate degree attainers. However, this study used as degree attainer those students who also left the program and subsequently finished at another school. While helpful in proving the efficacy of using GPAs and GREs in the selection process, the study did fail to identify any variables which would predict degree success in the specific program studied.

Simple correlation studies generally suffer from the fact that low but significant values of a Pearson r only account for a small portion of the variance between predictor and criterion even if the question of validity can be ignored. Proposals have been made (Dawes, 1971; Tully, 1962; Hoffman, 1960) to use prediction models which take into account the use of multiple prediction techniques. Dawes (1971), using a paramorphic representation of the admissions committee at one university, found that a simple mathematical equation could model admissions committee behavior so that 55% of the applicants could be eliminated without a single error of omission of an actually accepted candidate. Although the criteria used included GRE scores, undergraduate GPA and a quality rating of the undergraduate institution attended by the applicant, Dawes suggested that
more applicable factors might be substituted. As was pointed out in pre-
ceeding sections though, some promising measures of potential success are
very time consuming for both the applicant and evaluator. It may be best
to assume, until further research indicates otherwise, that the more tra-
ditional variables for prediction should be used.

Summary of the Present Selection System

The above review of predictor, criterion, and methodological variables
used in research of candidate selection into graduate psychology programs
has noted that the significance of current data is often tenuous. Predic-
tor variables must be proven relevant while also being realistic in terms
of the applicant's ability to supply the required information. Anyone who
has applied to graduate programs within the last five years is probably
aware that no two schools request the exact same information and that the
composite requirements for transcripts, standardized tests, personality
tests, letters of recommendation, writing samples, interviews, etc. is
often taxing and expensive. Criterion variables also need to be defined
in a relevant manner. Graduate grades are not useful because they are
generally homogeneous, and single class ratings such as in a practicum
course do not take into account the fact that ultimate skills are also
related to subsequent practice, experience, and learning. A poor evalua-
tion today in a practicum class is relevant only when compared with fur-
ther evaluations obtained two, three, or five years later. Relevance for
criterion variables are directly related to longitudinal success defini-
tions. Finally, methodological problems of past research include studies
which utilized incorrectly constituted samples (although realistically it
would be tremendously difficult to initiate a system at most institutions where all applicants were allowed to matriculate), and also, research which has attempted to predict success from single variables.

Why then does the system work so well, as attested by the fact that schools produce possibly more psychologists than we need (Atbee, 1976)? Although no recent data can be found on the national attrition rate from graduate programs, there probably are not many long-term losses. This could be due in part to selection factors which are at work prior to one's application to graduate school. The student whose academic record is poor, who may not be able to obtain three strong letters of recommendation, who is not motivated for the line of work of a professional psychologist, or who monetarily cannot afford to stay out of the job market any longer, probably does not apply in any significant number. The pool of aspiring applicants who are left tend to vie for candidacy in more than a few programs, and are eventually accepted by some and rejected by others. Strongly motivated candidates who do not receive acceptance to the schools of their choice have the option of working for awhile then re-applying to the same institution, or alternately to seek admission to and accept training in a less selective institution. The point is that at the doctoral level almost any candidate could be selected, given the present system, with few selection errors. Although not as refined in terms of pre-selection factors, the same is also probably true of Master's degree programs. Occasionally some inappropriate students do matriculate, but the demands of course work, research, economics, and life being what it is, programs do have some mechanisms for terminating incompetent students.
Heretofore there has been little concern for what may properly be termed the Type II errors of the selection system, i.e. the exclusion of those students who could make a significant contribution to psychology but never matriculate into graduate school. One could expect that given the large number of applicants into the typical graduate program, any application which presents reason for rejection is easily discarded as a time-saving measure for the selection committee. Some examples of rejection criteria may include a questionable or confusing letter of recommendation, GRE or GPA below the present year's cutoff point, or a late application.

Viewing the system globally, another important factor for many applicants is their own choice of schools to which they seek admission and the limited information available to them to aid them in making the most appropriate choices. Although there seems no easy way to absolutely determine the potential number of good students lost at selection time, this question should be of vital concern to applicants and training institutions alike.

A Tentative Proposal

If the present system in fact operates reasonably well, it would also seem reasonable to sit back and let the natural filtering of applicants proceed on course. Unfortunately (or fortunately, depending on one's view) there are a number of disquieting or disturbing aspects in allowing the present procedures to go unquestioned. One of the first of these is simply the time and energy expended by those applying for graduate study, those recommending their acceptance, the members of selection committees, and all others involved in the total process. There ought to be a more
efficient way. A second concern immediately presents itself also; namely, how many individuals who may really have extremely high potential for significant contributions to the field are dropped by the wayside unrecognized because predictor and criterion variables are still not identified or adequately defined. This leads to yet a third and possibly the most important concern—the need to initiate long-range coordinative research to systematically explore, identify, define and refine the variables involved, and thus increase our ability to more accurately predict the kinds of scientific behaviors and personal skills associated with productive professionals. The authors obviously endorse the above research notion. The following proposal, admittedly brief and in skeletal form, is therefore presented.

First a relook at how the present system seems to typically operate. Assuming that a potential applicant has reached the decision to formally apply for admission to graduate school, he/she would probably peruse available university catalogues, read descriptions of graduate programs, assess relative costs, etc. Then 5 to 15 universities would be selected, the applicant would request the necessary instructions and application forms and proceed to supply all the required information, including transcripts of previous credits, the usual letters of recommendation, etc. This information is duly received and processed by each of the universities, forwarded to their respective psychology departments, who then in turn appoint a selection committee to review and begin to screen the many applicants. This process of eliminating candidates continues, each step being more painful than the last, until the final selections are made.
the applicants are notified, the chosen few happily accept, and the selection committees then collapse until the next February - May rerun. Assuming, also, that a given department receives only 125 to 150 applications for 6-10 possible openings (the situation is apparently even worse at many institutions), each selection committee is faced with the awesome responsibility of making somewhat arbitrary and terribly difficult decisions, which have crucial impact on the applicants concerned, and selection committee members become more and more frustrated because the decisions have to be made on inadequate or insufficient data with unknown predictability.

As one possible alternative to the present procedure, let us now reconsider the potential applicant under a different system. The applicant proceeds pretty much as before, selecting the universities of his/her choice, fills out an application form and collects the other necessary information. But at this point the applicant is required to mail only one packet, to a central registry, whose initial task would be to classify, record and do the initial screening—the typical process that would ordinarily be taking place at perhaps 10 universities for a given applicant. On the application form, the applicant would indicate either the specific universities to which he/she would like the application sent, (including any required fee), or the applicant may simply request that his/her application be placed in the central registry's file. (In all cases, of course, applicants could also apply directly to the universities of their choice, although this might hopefully become less essential if the proposed centralization proved more efficient and acceptable). The central registry's staff could then proceed further to roughly rank-order and
otherwise categorize applications, using the best available criteria.

By this time the universities may have considered all of the applications sent directly to them, as well as those forwarded by the registry, but may still not be satisfied with the selection available to them—or for other reasons a given institution may still have additional openings. The institution could then request from the registry, for a nominal fee, the applications from the top 10 rated individuals from any combination of categories (e.g. women, minorities, etc.) who had not as yet been accepted into some other graduate program. In this manner each institution would have the potential to fill all training slots from the widest possible population base, and all applicants would enjoy increased potential for consideration and acceptance.

In essence, the proposed procedure would provide a central clearing house and data bank which would certainly appear to be less time consuming for all concerned, would cut down the attrition rate for high potential applicants and perhaps, most importantly, would develop skilled staff and researchers who could begin to assemble a data pool for needed longitudinal studies of the selection process, as well as for other appropriate research. There no doubt would be some complications and even disadvantages to such a procedure, e.g. some universities might simply prefer to handle their own applications directly, but one can also see some unique advantages as well, which would seem to make the proposal worthwhile.

It would be presumptuous of the authors at this point to try to specify in any great detail how such a central registry might be established and organized, but it would be seen as being under the aegis of the APA, possibly in conjunction with a particular university, agency or private corporation.
On balance the concept appears sound and would be seen as warranting serious discussion in some broad APA forum. Responses to this proposed concept would also be welcomed.
References


Hurst, J., & Shatkin, S. Relationship between standardized admissions variables and certain interpersonal skills. *Counselor Education and Supervision*, 1974, **14**(1), 22-33.


Merenda, P., & Reilly, R. Validity of selection criteria in determining success of graduate students in psychology. Psychological Reports, 1971, 28(1), 259-266.


Puranajoti, V. Relationship of personality characteristics to counseling success of graduate students in psychology in counseling practicum courses. Dissertation Abstracts International, 1972, 32(7-A), 3700-3701.


Tully, Emerson. Screening applicants for graduate study with the aptitude test of the Graduate Record Examination. *College and University*, 1962, 38, 51-60.


