

# STUDENT SELECTION BY MEANS OF PSYCHOLOGICAL TESTS—A CRITICAL SURVEY

By H. J. EYSENCK

(Psychological Department, The Maudsley Hospital)

I.—Introduction. II.—Intelligence and success at College. III.—Other factors affecting success at College. IV.—Considerations regarding the use of selection tests. V.—Summary and conclusions. VI.—References.

## I.—INTRODUCTION.

It is universally agreed that university students are not a random sample of the population, but constitute a somewhat homogeneous group. In particular, it is generally conceded that on the average university students, in whatever faculty they may be registered, tend to be of above average intelligence. We may gain an approximate idea of the intelligence of students in American colleges from a paper (Traxler, 1940) bringing together the results of testing 68,899 students in 323 colleges. The mean I.Q. was 109, with a S.D. of approximately 12; in other words some 25 per cent. had I.Qs. above 117, while another 25 per cent. had I.Qs. below 101.

This figure must be interpreted with caution, although other published material tends to confirm it. Different universities differ widely in their requirements; in some American colleges the average intelligence of the students differs hardly at all from that of the non-student population, while in some of the better known universities the standard is so high that the average I.Q. might be in the neighbourhood of 130. The latter figure would seem to bear a closer relation to conditions in this country than the over-all figure of 109, although there is not sufficient evidence to quote any precise figure.

According to Burt's calculations (1943), of 709,580 persons aged eighteen years in 1939-1940, 10,785 or 1.52 per cent. entered one of the English or Welsh universities. If only the most intelligent 1.52 per cent. had been selected, the borderline I.Q. would have been at the  $+2.17$  S.D. level; in terms of I.Qs. this would make the border-line I.Q. = 135 approximately, taking the standard deviation of the I.Q. as 16. However, as Burt points out, "a simple calculation shows that about 40 per cent. of those whose innate abilities are of university standard are failing to reach the university; and presumably an equal number from the fee-paying classes receive a university education to which their innate abilities alone would scarcely entitle them." These figures suggest that the figure of I.Q. 135, rather than marking the lower limit of the intelligence distribution of the student population, is somewhat near the average. It is improbable that we would go very far wrong if we estimated that the average student I.Q. lies somewhere between 125 and 130, with wide differences between different colleges, and possibly also different faculties.<sup>1</sup> In this argument, the term I.Q. has been used although strictly speaking its use with adult groups is fallacious; where the I.Q. is arbitrarily converted from the test S.D. score into an equivalent Binet I.Q. its value is mainly sentimental, and where other methods of conversion are used the resulting I.Q. is often meaningless.

<sup>1</sup> Published and unpublished studies on samples of English and Scottish students by Vernon, White, Dale, Thomson, Leybourne-White, Wright, and others agree in putting the I.Q. within the range indicated. The difference between an I.Q. of 135 marking the lower limit of student intelligence in conditions of optimal selection, and an average I.Q. of 127 under present conditions or selection, points a moral which needs no elaboration.

However much these criticisms may detract from the exactness with which we are able to estimate the intelligence of the university student, there remains little doubt that intelligence is one of the factors which distinguish the student from the non-student. From this observation to the suggestion that a knowledge of the intelligence of prospective students would be of use in selection and in counseling was only a small step, and McKeen Cattell appears to have been the first to make this suggestion upon his return from Wundt's laboratory to America. Wissler (1901) followed up the idea, but with disappointing results, explicable in terms of the very primitive, non-intellectual type of test used by him; with the exception of isolated attempts to apply the Binet scale to small groups of university students, summarized by Caldwell (1919), no progress was made until the time when the Army Alpha and the Thorndike scales became available, in 1916.

In the five years following, a veritable spate of publications emerged from psychological laboratories all over America, all reporting correlations between intelligence test results and students' success; Toops (1926) collected a bibliography of 400 titles covering the period to 1924 and pointed out that less than one in three of experiments carried out had been published. Since then the flood has abated somewhat, but a regular flow of reports has continued to date, summarized at various times by MacPhail (1924), Edgerton (1920), Wagner (1932), Segal (1934), MacPhail and Bernard (1943), Crawford and Burnham (1946), and many others; in addition to these summaries, the Annual Reports of the Commission on Scholastic Aptitude Tests, set up by the College Entrance Examination Board, contain much valuable material.

It may be estimated that the total literature in this field contains well over a thousand titles, and that some three thousand additional studies have been carried out without finding their way into print. These figures are only approximate, and err on the side of under- rather than of over-estimation. Many of the published reports are unobtainable in this country, and no complete survey is possible or even desirable, in view of the uniformity of findings. The present paper is based on findings reported in some 600 articles; those which are available in this country are quoted in the bibliography. In addition, a small number of papers not easily available here are quoted because of their general importance.

## II.—INTELLIGENCE AND SUCCESS AT COLLEGE.

The usual method of estimating the success of intelligence tests in forecasting success at college has been that of giving the test to all entrants, to obtain an estimate of the success of the student through his marks after one, two, three or four years, and of correlating the test scores with the marks. The criterion used for estimating the success or failure of the student will be discussed in a later section; here let us note merely that in nearly every case there has been a tendency to find low to medium positive correlations between the tests given and the final marks obtained by the student. Most publications report coefficients ranging from .3 to .7, with a mean in the neighbourhood of .5.

These values are often reached by straightforward averaging of correlation coefficients, a procedure which lacks statistical validity. The present writer took a random sample of 500 reported coefficients from his files, transformed each coefficient into its corresponding  $z$ -value ( $z = \tan^{-1} r$ ), averaged the resulting  $z$ -values ( $\bar{z} = .586$ ), and transformed this average into  $\bar{r} = .527$ . This value is not sufficiently dissimilar from those found by others to invalidate previous findings; when the distribution of coefficients is plotted in terms of  $z$ ,

however, the marked skew usually found when distributions are plotted in terms of  $r$  now disappears, and the curve of distribution is not significantly different from normal (as tested by the chi square test).

A correlation of the size  $r = .53$  is not very useful for prediction ; it indicates that less than 30 per cent. of the causal factors in college success are measured by the test, and that the percentage reduction of error in prediction over chance is only 15. The coefficient of alienation is too high, and the index of forecasting efficiency too low, to make one place much confidence in prediction. There are several considerations, however, which make the outlook less gloomy.

In the first place, an average correlation based on a random sample of published work does not give a fair idea of what the best tests can do when applied by experienced workers ; it merely shows what happened during the past thirty years when good and bad tests were given to suitable and unsuitable populations by experienced, indifferent, and frankly amateurish workers. When we add the observations that (1) many of the tests used were obviously unsuitable for their purpose, having been constructed for children, or for normal adults, rather than for highly intelligent adults ; (2) that many of the tests used were far too short to give a reliable, let alone a valid, result ; (3) that even the sheer working out of the correlation coefficient often shows signs of incompetence, data showing obvious heterostedasticity being treated by formulæ implying homostedasticity ; (4) that little use appears to have been made of such obvious devices as multiple and partial correlations, of factor analysis, or of differential test scores ; (5) that conditions of homogeneity have seldom been fulfilled adequately ; (6) that the criterion itself against which the tests have been validated is highly unreliable (cf. below) ; when all this is taken into account the figure of  $.53$  may take on a new and rather more hopeful significance.

In order to find out the improvement which could be secured by taking only studies expertly conducted, a subsample of thirty-four coefficients was drawn from our files reporting studies fulfilling the following conditions : Specially designed tests only should have been used ; length of testing should not have been below ninety minutes ; number of cases should have been above 250 ; statistical treatment should have avoided obvious fallacies. When averaged via the  $z$ -function, these thirty-four studies show an average correlation of  $\bar{r} = 0.58$  between intelligence and success. This value would appear to be more meaningful than the previous one ; it is significantly higher than the value of  $.53$  found in the total sample. This finding suggests that such conditions as suitable choice of test, length of testing time, and adequate statistical treatment are necessary if spuriously low predictive values are to be avoided.<sup>1</sup>

<sup>1</sup> In this connection it should also be mentioned that Derfflinger (1943) has shown that correlations between tests and achievements have shown a rise in recent years, such as one might expect on the basis of improvements in test construction, statistical treatment of results, and estimation of achievement. Our own figures, analysed from the point of view of temporal development, are in agreement with Derfflinger's. On the other hand, Feder (1935) and Williamson (1937) have shown a tendency for correlations between tests and achievement to decline in recent years. The explanation for these discordant results lies in the fact that Derfflinger and the present author have summarized work covering great numbers of different universities ; Feder and Williamson were only concerned with the trend in their own universities. Their own explanations of their findings agree with each other, and are probably correct ; they believe that if administrative action is taken in response to the results of tests, coefficients will tend to become lower. Thus by allowing the intelligent to take on extra course work, while preventing the dull from doing so, the two groups of students are more equalized in respect to their degree work, and will tend to show less difference in achievement than they would have done if both groups had been allowed to take on extra work, or if both groups had been forbidden to do so. Services experience in this country indicates that " the more efficient the selection scheme put into operation by psychologists, the lower the validity coefficients of the tests."

Even this higher value has to be seen, not in isolation, but against the background of the criterion used to establish validity. The predictive value of a test is probably depressed more by the unreliability of the predictive criterion than by the unreliability of the test itself; certainly the unreliability of the test is measurable and can be remedied, while the unreliability of the criterion is difficult to measure and almost impossible to remedy.

Unfortunately, we know very little about the reliability of examinations at the university level, although work such as that carried out by the International Institute Examinations Enquiry (Hartog and Rhodes, 1935), leads one to suspect that reliabilities of the order of 0.8 will not often be reached or surpassed. Certain data are available to show that correlations are higher between tests of intelligence and objective tests of achievement in courses, than between tests of intelligence and subjective tests of achievement in the same courses (Tharp, 1927, Frazier and Heilman, 1928), but this may be due in part to the similarity in form between intelligence and objective achievement test, rather than to low reliability in the subjective achievement test. Extensive and careful figures are published by the Scottish Council for Research in Education (1934); it was found that correlations between Degree marks and Department marks were as follows: English=0.36; Mathematics=0.30; Latin=0.59; French=0.59. These figures again do not give direct evidence on the reliability of the examination marks, but they suggest that either these marks are not very reliable, or that a special factor of "examination ability" is more powerful in determining final standing of the student than his knowledge and ability as shown by his Department mark. A direct study of these questions is an obvious necessity, and little can be said usefully about what the outcome of such a study might be.

Instead of measuring reliability of the criterion, many investigators have correlated the course marks gained by students in different years; thus the standing of students in the freshman class may be compared with the standing of the same students four years later in their degree examination. These correlations are of interest because they presumably set a limit to our ability to predict success at college by the use of any type of test whatever; if first-year standing and fourth-year standing correlate 0.75, then it is presumably impossible to find any combination of predictive measures which will predict final standing better than to the extent of 0.75. (We are making the assumption here that the marks on which first-year standing is based are perfectly reliable; if they are not perfectly reliable the use of perfectly reliable prediction tests of the highest possible validity might give predictive accuracy equal to the value of 0.75 *corrected for attenuation*, and therefore rather higher.)

Evidence presented by Toops (1926), Guiler (1927), Read (1940), Wood (1923), Cleeton (1931), and others shows that correlations between students' standing in different years vary between 0.45 and 0.85, with a mean in the neighbourhood of 0.75. Differences between different universities are too great to be due to sampling errors, and presumably have much to do with such obvious causes as carefulness with which papers are marked, relative homogeneity of student population, and amount of student counselling available. Whatever may be the true figure in this country, it seems safe to say that we have not reached the stage where we can say that our tests predict final standing as well as does the sum-total of the work done during the first year at university. (This is the enviable result of McClelland's work (1942) in selection for secondary education, where apparently optimum forecasting accuracy has been achieved. Quite generally it may be said that McClelland's discussion of selection procedures is of interest no less to universities than to those responsible for

secondary education, and that everyone concerned with selection ought to be familiar with the methods, experimental and statistical, described in his book.)

The data quoted so far establish the fact that tests can predict reliably and reasonably validly the future performance of university students. The knowledge thus gained can be utilized in a great variety of ways, and has been so used in the past; selection of students on the basis of their test scores is only one of the many possible uses. Jones (1928) gave intelligence and achievement tests to 112 freshmen and reports that counselling service given to those working below expectancy led to marked improvement in their work. Starrack (1937) reports follow-up studies at the Iowa State College of those falling below their ability level, including an investigation of reasons for failure, a systematic programme of personality development, and the use of "therapeutic methods of suggestion"; he also reports on the granting of special privileges for students doing better than expected on the basis of their ability. Compton (1941), in a study of 331 institutions, found that thirty-eight informed all students of their test results, eighty-one informed none, the others gave information to varying proportions along with interpretation and attempted guidance. Toops (1926) points out that "college intelligence tests are not entrance tests. They are primarily educational administrative devices for dealing with administrative and pedagogical problems of students, rather than the criterion of intelligence of applicants for admission."

Toops bases his opinion on the results of a twenty-five query questionnaire sent to 110 colleges and universities, all of whom replied. Sixty per cent. of these institutions used tests administratively, while another 12 per cent. used them experimentally. There was a tendency for smaller colleges not to give tests at all, and for the larger ones to give them only experimentally; the medium-sized institutions appeared to make the most use of the results. Toops tabulated the number of colleges making certain uses of test data and found that of sixty-six institutions giving tests, none used the result as the sole basis for admission, nineteen used it as a partial basis for admission, forty-nine in determining dismissal for low scholarship, thirty-four in determining probation for low scholarship, thirty in dealing with disciplinary cases, thirty-six in determining amount of school work to carry, eighteen in determining amount of work for self-supporting, twenty-five in encouraging bright students to undertake graduate work, forty-two in encouraging extra effort in case of unmotivated bright students, thirteen in the selection of assistants, twenty-three in making recommendations for scholarships, fifteen in making recommendations for fellowships, thirteen in hiring student clerical help, six in determining membership in honorary scholastic societies, and eighteen in "sectioning" students according to capacity for progress. Thus 341 uses are reported from sixty-six institutions, making an average of five uses per college or university, apart from the possible use for selection.

Even where no use is made of intelligence test results in determining who shall go to university, most administrators and psychologists who have had experience of using the results in the running of educational institutions agree that they are of considerable help on such points as those mentioned in the last paragraph. However, a situation has now arisen in this country in which there are two or three applicants for every vacant place in the universities. Thus the necessity of selection in one form or another has been forced on the authorities, and it would appear that a knowledge of the intelligence of the applicants, together with such other information as may be likely to help the selection committee, would be of great value in ensuring a supply of suitable candidates to the universities. Before discussing this question in detail, a

number of problems must be at least mentioned, as they supply the basis on which some of our recommendations are based. Thus no argument regarding the question of the use of tests for selection can be very helpful which does not deal with such problems as the function of the selection ratio, the measurement of primary abilities, the predictability of success in special subjects (teaching, medicine, etc.), or the relative usefulness of non-cognitive tests. These and other points will now be taken up in order.

One argument against the prevalent practice of giving only one type of intelligence test (nearly always a verbal one) is that the isolation of a number of group factors by Burt, Thurstone, Holzinger, and others makes it likely that more analytic studies involving the correlation of success in college with each of several different abilities might be more diagnostic than the mere measurement of 'g,' adulterated by an admixture of 'v.' The main studies in this connection are those by Ball (1942), Ellison (1941), Goodman (1944) and Tredick (1942), using the Thurstone battery of Primary Abilities tests, and the Ohio studies of non-verbal tests by Edgerton (1940), Heston (1942), and Troyer (1942). These studies, as well as the related work of Butch (1939) on differential weighting, show that the analytic approach to the problem bids fair to increase the predictive validity of tests. A more accurate assessment is not possible at the present time in view of the small number of the studies carried out.

The same might be said of the related question of predictive validity of tests, not for success at university as such, but for success in specified courses. Thus prediction, to be useful, would have to be specific; instead of saying: "This person will or will not succeed at college," it would have to say: "This person would succeed in becoming a teacher, but would fail in the medical course," or "This person might do well on the Arts side, particularly in the linguistic field, but he would be hopelessly out of place on the Science side." This form of analytic prediction and test interpretation is, of course, closely bound up with the question of primary abilities, discussed in the preceding paragraph, and with the question of non-cognitive selection, discussed below. As long as we remain at the stage where selection is on the basis of a 'g' score alone selective prediction is impossible, except in so far as different faculties show different mean ability scores. While the existence of such differences seems to be well established, the differences are usually much too small, and the overlap much too great, for any specific counseling to be based on them.

Most studies in this field have dealt with the selection of teachers, and the result has uniformly been that forecasts based on an assessment of 'g' are relatively useless; while predicting with moderate accuracy the success of the prospective teachers in their examination, correlations with their actual teaching ability as assessed by subjective estimates are rather low. The fault here may lie in the unreliability of the criterion, rather than in the inaccuracy of the test; but commonsense suggests that in teaching, personality traits may be of more importance than particularly high intelligence, and findings so far have not shown commonsense to be mistaken, although it cannot be said that they have added much to its pronouncement.<sup>1</sup>

Prediction in the faculty of medicine has been rather more successful on the whole. Apart from work such as that of Edholm and Gibson (1944) in which unsuitable tests were employed, the studies of Cowdery (1931) and Moss (1931,

<sup>1</sup> Relevant to this point may be the work of Cattell (1931), Pinsent (1933), Saer (1941), Turnbull (1934) and others, which has appeared in this *Journal*. Also of interest here is an unpublished memorandum by N. Wright on "The Use of Psychological Tests in the Selection of Candidates for Training as Teachers." (Jordanhill Training College, Glasgow; I am indebted to Dr. P. E. Vernon for drawing my attention to this work.)

1933) have demonstrated that failures can be predicted with a good deal of accuracy. The figures given by Moss may perhaps be quoted as representative, as they are based on a careful, large-scale study using 6,000 students. He showed that the Moss Medical Aptitude Test predicted 69 per cent. of the failures during the first and second years of the course, while Grades in Pre-Medical subjects predicted only 43 per cent., and Interview Ratings only 33 per cent. Test and Pre-Medical Grades together predicted 73 per cent. of the failures. Correlation of the Moss test with first-year grades was  $+0.59$ , and while in the best test decile there were no failures at all, there were 56 per cent. in the tenth decile.<sup>1</sup> Similarly, Cowdery and Ewell found a correlation of  $.64$  between test and success at medical school, and a multiple correlation of  $.76$  with test and pre-medical grades.

The objection has been made to studies of the kind mentioned that success in the test does not guarantee good performance as a doctor; that it only means that the candidate is likely to succeed in passing his examination. This objection is perfectly true. Success in the test means that the requisite ability for success in the examination is present; it does not mean that the ability will be employed in such a way as to ensure such success. Consequently many students possessing the ability will succeed, but a certain number will also fail, through lack of ambition, through neurotic and emotional upsets which interfere with their work, because of having to earn their living at the same time as studying for their degree, and for a large number of similar causes. Consequently the "examination variance" of the bright students will be considerable. On the other hand, the dull student who has not got the requisite ability to succeed will fail regardless of outside circumstances; consequently the "examination variance" of the dull students will be very small. This point suggests the answer to the objection quoted at the beginning of the paragraph; the test does not guarantee that those passing will be successful, but it does guarantee that they have the ability to pass if outside circumstances and personality qualities are favourable, while those who fail the test are almost certain to fail the examination however favourable outside circumstances and personality qualities may be. Going one step further, we may here say that while even examination success is not final, and does not guarantee success as a doctor, or a teacher, failure at the examination makes success impossible and failure certain.

The whole question of the relation between examination success and quality of work done in one's chosen profession is an extremely difficult one, and one which demands much more investigation than it has received. One cannot help feeling that if a certain proportion of the thousand or so investigations published in the field of prediction of examination success had been devoted instead to the investigation of what examinations predict in terms of quality of later work, and to the elaboration of some reliable and valid criterion against which to check both test results and examinations, science would have profited to a considerable extent. After all, if test and examination fail to show a high correlation, this may be due not only to the unreliability of the examination, or the inadequacy of the test, but also to the specific nature of the examination which makes it a bad and invalid criterion of professional quality. A last possibility which should be borne in mind is the fact that the type of training which precedes the examination may have much to do with the type of person who succeeds in the examination; if training consists largely of the inadequate "cramming" method it is quite possible that the highly intelligent, conscientious type of person is penalized, while the born "crammer" is

<sup>1</sup> N.B.—These failures occurred during the first two years. Presumably many more students from this decile failed during the next two years.

rewarded. Thus a mere investigation of the correlation between test and examination tells us little about those points concerning which we are most ignorant; quite generally it may be said that attempts to predict without knowledge of or control over training will fall short of the full measure of responsibility that the psychologist, in close collaboration with the expert teacher, doctor, architect, or whatever it may be, ought to be ready to take.

We may sum up the discussion so far by saying that in considering only the requirements of intelligence, work reported hitherto has shown quite definitely that moderately high correlations between tests and examination success can be expected; that prediction is becoming more analytic, both in the choice of tests and in the specific nature of the prediction; that examination success itself is not a perfect criterion, but is deficient from the points of view of reliability and validity; and that many outside factors, such as personality traits, types of training, and various others must be taken into account. We may next turn to a consideration of such work as has been done on factors other than intelligence and other ability tests. While the results in this field are not as clear-cut as in the field of intelligence, there is not the slightest doubt that progress is more likely to be made here than in the further development of tests of intelligence.

### III.—OTHER FACTORS AFFECTING SUCCESS AT COLLEGE.

Commonsense indicates that success depends not only on ability, but also on character, temperament, and interest, as well as on other imponderable factors. That this view is broadly speaking true is shown by a number of factorial studies in which correlations were calculated between a variety of tests of intelligence, and marks in a number of school or university subjects. The work of Alexander (1935), Holzinger (1939), Vernon (1939), Braddock (1946) and others shows with surprising unanimity that there is a factor common to all the scholastic subjects which cannot be identified with 'g,' 'v,' 'k,' or any other primary ability factor. This factor, called 'X' by Alexander, would appear to be the resultant of all the non-cognitive factors in so far as these influence success at school or in college. The amount contributed to the total variance by this factor is often surprisingly large, and if we could measure the influences which go to determine it we would have here a means of improving considerably our predictive accuracy.<sup>1</sup> Upwards of five hundred studies are recorded in this field; excellent summaries are given in two papers by Harris. The present section will give only the present author's conclusion resulting from a survey of the evidence; no attempt is made to duplicate Harris' (1931, 1940), reliable reviews and bibliographies which are easily accessible. We have, however, added a few references to the literature since 1940, and not contained in the summaries mentioned.

An indirect measure of the "X" factor is given by the prospective student's standing at school; presumably his success in school involves his persistence, studiousness, interest, application and so forth to a similar degree as later work in college, and can, therefore, be used to predict the latter. Unfortunately, it is difficult to compare different studies of this question because of the wide differences in quality of teaching, standards of examination, and composition of the student body in the high schools studied; in general it may be said that school standing does not predict college success less well than do intelligence tests, and many reports indicate that it may predict success at college even

<sup>1</sup> Vernon explains this factor as one of "studiousness, or willingness to work," an interpretation similar to our own.



better. There is, however, general agreement that school standing and intelligence combined give a better prediction than either of them singly; multiple correlations are often surprisingly high. Much the same is true of an indirect test of school success, the entrance examination based upon acquired knowledge rather than upon innate ability. Such entrance examinations often predict as well or better than intelligence tests; best prediction is usually achieved by a combination of the two types of test.

Other data dealing with certain objective items relating to personal background are of less value in prediction. There are many studies of the influence of age, sex, and family data; these are difficult to evaluate because of technical faults which invalidate many findings; often there is a failure to partial out intelligence before studying the influence of other factors! However, certain questionnaires dealing with background data have shown reasonable correlations with success at college (Scott, 1938, Gray, 1938, Asher, 1940, Bittner, 1945), generally in the neighbourhood of  $+0.4$ , and it is usual to find that when these results are added to intelligence test scores they raise the multiple predictive  $R$ . It must remain doubtful, however, in how far correlations between personal data and success obtained in, say, the State of Kentucky, are likely to be duplicated in other states, or in this country; no data are available to answer this question.

Physical data show uniformly low correlations; the literature suggests that smokers are inferior to non-smokers, colour-blind students to non-colour-blind students, students of euryomorphic body build to students of leptomorphic body build, the ugly to the good-looking. The first two findings appear valid because intelligence was partialled out; the other two are probably due to the correlation of intelligence with leptomorphic body build (Eysenck, 1947) and with beauty. For all practical purposes these data can be disregarded.

Direct tests of personality qualities of the questionnaire type often show slight correlations with success at college, but results are too divergent, and correlations too low, for this type of work to be of practical use.<sup>1</sup> Ratings by school teachers and others who know the prospective student well are more promising, but not routinely available. Of more importance are studies of interests, reviewed extensively by Strong (1943) in his book on *Vocational Interests of Men and Women*. Work along these lines suggests very strongly that here we have a promising indicator of success at college which is relatively independent of intelligence, and which would be certain to add to the multiple correlation of a selection of other tests. A limited amount of work done in this country with the Strong test (unpublished) suggests that it might be as useful here as it has been found to be in America.

Many of the studies reviewed so far fail to satisfy because they are not based on any general theory, but appear merely as accidental successes among large numbers of failures. If a hundred "personal data" are correlated with success at college, it is reasonable to expect five of them to show a significant correlation with the criterion, on a purely chance basis; combining these five correlations by means of multiple correlation formulæ would result in quite high  $R$ , which would, however, drop to zero if another sample were to be tested. In some cases the amount of "shrinkage" likely to occur has been estimated by suitable formulæ, but this has not been the uniform practice, and even when

<sup>1</sup> It is usually found that 'neuroticism' correlates negatively and 'introversion' positively, with success at college. But as both variables are known to be correlated to some extent with intelligence reported correlations may be spuriously high. But the existence of such correlations appears reasonable enough in view of the writer's work on these two personality variables (Eysenck, 1947).

the shrinkage formula is used, it is still doubtful how the resulting  $R$  would compare with the correlation found in an altogether different sample.

An exception to these strictures is the excellent work of Munroe (1945). Building on the findings of many investigators that students showing neurotic traits usually succeed less well at college than students relatively free from such traits, she evolved a simplified and more objective scoring method for the Rorschach test, involving the counting up of all neurotic indications on the record, and scored each of the students tested by her on the basis of this general quality of "neuroticism." She then showed that prediction of college success on the basis of neuroticism score was slightly more successful than prediction on the basis of intelligence tests, and that in combination these two tests showed a much better predictive power than either test by itself.

Further work of this kind, based on sound theoretical background, may succeed in transforming the rather drab picture of non-cognitive prediction. It may be that the factorial demonstration of the existence of two major personality factors, together with the provision of batteries of tests for their measurement, will aid in the development of such a background (Eysenck, 1947); at present, little use has been made of such objective tests in the prediction of college success.

In summing up briefly the conclusions derived from a rapid survey of work on non-cognitive factors, it may be said that promising correlations are reported in the field of pre-college achievement, university entrance (achievement) examinations, personal data sheets, vocational interests, and the measurement of "neuroticism," while hope is held out that the application of more objective measures of personality variables and traits may lead to even higher predictivity. There is little doubt that no single test by itself can give sufficient information to guide selection policy wisely; whenever possible, all the tests and measures found useful should be employed, and multiple regression and factorial techniques used to ensure the best available prediction.

#### IV.—CONSIDERATIONS REGARDING THE USE OF SELECTION TESTS.

The studies quoted in this paper have been almost exclusively American. The question arises, can conclusions appropriate to one country be applied immediately to another country? And are there special considerations applying to this country which would make it more or less likely that selection methods could be used with effect at the university level? Discussion up to this point has been entirely factual, although inevitably selective; this section is the writer's attempt to draw conclusions from the evidence cited, and to bring various other considerations to bear on the problem. Consequently it will be appreciated that a certain amount of subjectivity is not to be avoided, and that while the writer may at times appear to sound dogmatic because of the necessity to put his points as briefly as possible, he wishes to encourage discussion of the proposals, rather than to make definitive statements.

The small volume of published and unpublished work done in this country shows that in a very general way the relation between intelligence tests and achievement is positive, and that we can expect similar types of relation to obtain here as are found to obtain in America. (Vernon, Dale, White.) Differences at the moment centre around two concepts; relative homogeneity and the respective selection ratios.

It is, of course, well-known that in relatively homogeneous groups, i.e., groups differing little with respect to intelligence, correlations of that variable with any outside variable will be considerably lower than they would be in a

relatively heterogeneous population. Prospective students are quite clearly a more homogeneous group with respect to intelligence than would be a random sample of the population ; for America, the figures quoted in the first paragraph of this paper suggest a 30 per cent. reduction of heterogeneity for university students as compared with the whole population. If our argument be correct that English students are a relatively more highly selected group than are American students, it would follow automatically that they would also be a more homogeneous group. In other words one would have to conclude that correlations between intelligence tests and university achievement would tend to be lower in this country than in America.

Other considerations, however, must be taken into account. Homogeneity with respect to a variable correlated with the predictive variable tends to lower correlations of that variable with the criterion ; on the other hand, homogeneity with respect to a variable not correlated with the predictive variable tends to increase correlations of that variable with the criterion (Wagner and Strabel, 1935 ; Dwyer, 1938). This is clear if we consider the case of a variable which is correlated with the criterion, but not with intelligence ; the more heterogeneous the population studied with respect to that variable, the less valid will the prediction of the intelligence test be. Now American students are probably on the whole far more heterogeneous with respect to sex, age, racial and national background, upbringing, and so forth than are English students ; consequently we might expect correlations between intelligence tests and achievement to be rather higher here.

It may be suggested that these two factors tend to balance out on the whole ; comparison of the results reported from the average American State University with results from some of the more selective universities, which in respect of homogeneity rather resemble English universities, does not indicate very considerable differences in the size of the predictive coefficients.

The question of the selection ratio is based on the simple consideration that the selective value of a test depends not only on its correlation with the criterion ; it also depends on the ratio of applicants to number of positions vacant (Tiffin, 1943 ; Taylor-Russell, 1939). Clearly, if there are roughly as many vacancies at a university as there are prospective students wishing to enrol, tests might still be given for the sake of counseling, or experimentation, or for other reasons, but from the point of view of selection they would be quite useless as every applicant would be taken anyway. If, however, the selection ratio is 5 (five applicants for every vacancy), even a test showing a relatively low correlation with achievement would serve quite well to pick out students almost certain to do well, and to reject those likely to fail. In general the higher the selection ratio, the more successfully does a test perform its function.

Now at present, and for several years to come, the selection ratio at our universities is likely to be relatively high ; possibly in the neighbourhood of two or three with yet higher figures in certain faculties, such as the medical. Under those conditions, even if tests should show lower correlations with achievement in this country than they do in the United States, they would nonetheless be of the greatest value because of the high selection ratios expected. This important point is often disregarded in discussions of the value of selection tests.

These considerations suggest that selection tests would be of definite value in showing clearly which of the prospective students would have only a very slight chance of completing the course successfully, and in designating those who could reasonably be expected to succeed. Perhaps the best method of utilizing test results would be to have available for each student data regarding his

pre-college scholastic achievement, his university entrance (achievement test), exhaustive intelligence tests covering not only 'g' but also other primary abilities, a personal data sheet covering items shown by experience to correlate with success and achievement, a measure of his vocational interests, such as the Strong blank, and an indication of his neurotic predisposition, such as the Munroe test. (Ideally, of course, it might be suggested that university selection should be only the last stage in educational guidance, a process which should be more or less continuous from the time the child enters the school. There should be various stages of selection, in the basis of tests, records and interviews during the child's school career; the accumulated mass of knowledge about the candidate would be invaluable to the university authorities.)

It would be easy to determine the level of intelligence below which achievement at the university level is impossible, or at least highly unlikely; all candidates falling below this level ought to be disqualified straight away. The remainder of the students should be selected on the basis of an interview with a board of several members of the university, including a psychologist; this board should have before them the scores of the student being interviewed on all the tests mentioned above. Their conclusion and recommendation would then be based not only on the purely subjective elements of the interview, but would proceed from a factual background integrated with personal knowledge of the candidate, such as can be gained only in an interview.

Certain points should be emphasised in this connection. In the first place, records should be kept of the correlations between achievement and detailed test scores, and between achievement and opinions expressed on the basis of the interview; only in this way can we learn something about the actual contribution made by the interview to the objective data. It is quite possible that multiple regressions predicting achievement could be written (without containing the interview) as high, or even higher, than similar equations containing the interview; little is known about this aspect of the problem. Another point is that test construction, administration and interpretation is a highly skilled task, and that nothing is more likely to bring predictive procedures into disrepute than the all-too-frequent practice of having this work done by the amiable amateur not fully versed in the literature, in test construction, and in the various statistical techniques of factorial analysis, multiple regression, analysis of variance, and so forth. If the literature contains one warning more obvious than any other, it is this—that no success can be expected from the haphazard throwing together of items and tests of unknown difficulty or factorial composition, a practice unfortunately too widespread. Only the most rigorous testing of the tests themselves, by the use of item analysis and the various other techniques elaborated for this purpose, can ensure that technical competence in test construction so absolutely essential for fruitful work, and so deplorably missing in many published tests.

This point clearly implies the next, viz., the impossibility of carrying out work of this kind successfully without some form of central organisation. The psychologist employed as lecturer, demonstrator, or in some other capacity in the university has neither the time, nor the facilities, for designing and scoring large numbers of tests, even where he is familiar with the principles according to which they ought to be constructed; as Vernon points out, American tests cannot be taken over directly as they are often quite unsuitable for students lacking an American background. Even in America, where psychological departments are vastly larger and better supplied with staff and money than here, it has been found necessary to have large central organisations which undertake yearly revisions of standard selection tests, and which undertake to

have them scored ; such a course would appear even more necessary in this country.

In addition to such a central organisation, however, it would appear desirable that full- or part-time psychologists specially trained for work in this field should be attached to universities to supervise the testing, carry out follow-up studies, and give advice to faculty and students when requested to do so, on the basis of the test results. American experience indicates that such use of psychological tests is perhaps even more urgent and fruitful than use for simple selection only, and if it is worth while to prevent the certain failure from entering college, surely it is worth while to try to prevent a possible success from failing through causes which can often be removed by counseling, advice, changes in courses undertaken, and remedial teaching. What can be done in this way is indicated by reports such as those quoted above showing that counseling service given to those working below expectancy may lead to marked improvement in their work.

Objections are certain to be raised to these suggestions on the score of cost. There is no doubt that to provide a service of the kind envisaged would be moderately expensive. However, the cost involved should be set against the tremendous social as well as financial gain resulting from improved selection and from successful guidance ; these gains, because difficult to measure in monetary terms are easily overlooked, but are nevertheless very real. Looking back over the past twenty-five years of American experience, we find much evidence to encourage the belief that a psychological service of the kind suggested would result in gains, social, financial and psychological, which far outweigh the limited outlay required.

This conclusion holds true even for the traditional type of forecasting and counseling ; it is the writer's contention that the total gain would be incomparably higher if a much greater imbrication of academic policy and psychological method could be achieved along the lines suggested in a previous part of the paper, i.e., by having psychological investigation of training courses, examinations, and so forth going on concomitantly with a selection and counseling programme. Even without such developments, however, the national need for trained personnel in practically all the professions is such that laissez-faire methods of selection will have to give way to a more scientific approach which can, up to a point, be provided by the use of psychological tests. It is the contention of this paper that a recognition of this need at the moment is a national priority.

#### V.—SUMMARY AND CONCLUSIONS.

An examination was made of the literature dealing with the use of psychological tests in the selection of university students. The following conclusions appeared well established :

1.—Tests of intelligence show correlations with university achievement of between +0.5 and +0.6 when suitable tests are chosen and conditions of testing are favourable.

2.—Standing at school shows equally high correlations with university achievement, and in combination with intelligence tests predicts such achievement with considerable accuracy.

3.—Special achievement tests may serve instead of school standing when external circumstances make it difficult to obtain suitable and comparable records for candidates.

4.—Questionnaires covering personal background and details of past history show relatively low correlations with success, but may be helpful when added to other types of tests.

5.—Personal inventories are unlikely to be of much use in practice, even when used in combination with other tests ; their use should be discouraged.

6.—Interest questionnaires show correlations with general and specific achievement which make it likely that such questionnaires would form a useful part of a battery of selection tests.

7.—Tests of temperamental qualities are of promise, but only in isolated cases can it be said that they are already of practical use ; Munroe's use of the Rorschach test is perhaps the outstanding example.

8.—A battery of tests covering various aspects of personality is more likely to give high predictive accuracy than any one test ; thus a single test of 'g' is less likely to be helpful in prediction, than a battery containing tests of several of the " primary abilities."

9.—Selection is only one of many possible uses to which results from a psychological examination can be put ; of equal or even greater importance may be sectioning, counseling, and advice on administrative problems.

10.—The detailed uses of tests, and the procedures recommended, cannot be summarized briefly ; it should be noted, however, that definite organisation, both within the university and also on a national basis, is required for the full benefits to be reaped.

In addition, psychological assistance will be useful in proportion as such matters as methods of training and construction of examinations become imbued with the psychological spirit.

#### REFERENCES

In order to make the list of references more useful, each citation has been prefixed by a letter denoting the relevance of the material contained in the article in question. "G" denotes papers dealing with the general problem of the relation of intelligence to college achievement ; "T" denotes prediction of teaching success ; "M" denotes prediction of success at medical school ; "S" denotes summaries ; "C" denotes studies of the criterion (success) itself ; "F" denotes studies dealing with factors other than intelligence. Certain references have not been marked as they are not directly relevant to the question of selection, but have been quoted in the course of the paper. An asterisk is prefixed to papers which the writer was not in a position to consult personally, but had to quote from abstracts.

G—ALEXANDER, N., and WOODRUFF, R. J. : Determinants of college success. *J. Higher Educ.*, 1940, 11, 479-485.

F—ALEXANDER, W. P. : Intelligence, concrete and abstract. *Writ. J. Psychol. Monogr. Suppl.*, No. 19, 1935.

G—ANDERSON, E., et al. : Wilson college studies in psychology : I. A comparison of Wechsler-Bellevue, Revised Stanford-Binet, and American Council of Education tests at college level. *J. Psychol.*, 1942, 14, 317-326.

G—ANDERSON, E., and SPENCER, L. T. : The predictive value of the Yale Classification tests. *Sch. and Soc.*, 1926, 24, 305-312.

G—ARLITT, A. H., and HALL, M. : Intelligence tests versus entrance examinations as a means of predicting success in college. *J. Appl. Psychol.*, 1923, 7, 330-338.

P—ASHER, E. J., and GRAY, F. E. : The relation of personal history data to college success. *J. Educ. Psychol.*, 1940, 31, 517-526.

G—AVERILL, L. A. : A mental survey of fifty-nine normal school students ; some correlations and criticisms. *J. Educ. Res.*, 1923, 7, 331-337.

\*G—BALL, F. J. : A study of the predictive values of the Thurstone primary mental abilities tests as applied to lower division freshmen. *Penn. St. Coll. Stud. Educ.*, 1942, No. 24, 51.

F—BEAR, R. M. : Factors in the achievement of college freshmen. *Sch. and Soc.*, 1926, 24, 802-804.

- \*G—BELL, W. J. : Prognostication of college marks by tests of intelligence and of high school content. *Proc. Iowa Acad. Sci.*, 1935, 42, 174-175.
- G—BELL, C. J. : Mental tests and college freshmen. *J. Educ. Psychol.*, 1916, 7, 381-387.
- \*F—BERG, I. A., et al. : Scholastic achievement of students entering college from the lowest quarter of their high school graduating class. *J. Amer. Ass. College. Registr.*, 1944, 20, 53-59.
- F—BILLHART, W. H., and HUTSON, P. W. : Determining college ability during junior high school years. *Sch. and Soc.*, 1941, 53, 547-552.
- F—BITTNER, R. H. : Quantitative predictions from qualitative data : predicting college entrance from biographical information. *J. Psychol.*, 1945, 19, 97-108.
- G—BLAIR, J. L. : Significant factors in the prediction of the success of college freshmen. *Thesis, Univ. of Chicago*, 1931. Quoted in Freeman, F. M., *Mental Tests*. Houghton Mifflin, 1939.
- F—BRADFORD, E. J. G. : Selection for technical education. *Brit. J. Educ. Psychol.*, 1946, 16, 20-31.
- \*F—BRECKINRIDGE, E. : A study of the relation of preparatory school records and intelligence test scores to teaching success. *Educ. Admin. and Superv.*, 1931, 17, 649-660.
- S—BRIGHAM, C. C. : Annual reports of commission on scholastic aptitude tests. *Coll. Entrance Examination Board, N.Y.*, 1926, 1929.
- G—BROOM, M. E. : Predicting success in first semester college social science. *Sch. and Soc.*, 1928, 28, 689-690.
- G—BRUCE, L. : A study of the C.A.V.D. intelligence scale and its results in the doctor's level at Colorado State College of Education. *J. Exper. Educ.*, 1941, 9, 216-218.
- BURT, C. : Ability and income. *Brit. J. Educ. Psychol.*, 1943, 13, 83-98.
- \*S—BURTT, H. E. : The university intelligence tests, 1919-1922. *Ohio State Univ. Bull.*, 1922, No. 5.
- G—*et al.* : Efficiency of instruction in unselected sections in elementary psychology compared with that in sections selected on the basis of intelligence tests. *J. Educ. Psychol.*, 1923, 14, 154-161.
- G—BURWELL, W. B., and MACPHAIL, A. H. : Some practical results of psychological testing at Brown University. *Sch. and Soc.*, 1925, 22, 48-56.
- G—BUTCH, R. L. C. : Improving the prediction of academic success through differential weighting. *J. Educ. Psychol.*, 1939, 30, 401-420.
- F—BYRNS, R. K. : Scholastic aptitude and freshmen achievement. *Sch. and Soc.*, 1932, 35, 713-718.
- F—*and* HENMON, A. : Long range prediction of college achievement. *Sch. and Soc.*, 1935, 41, 877-880.
- S—CALDWELL, H. W. : Adult tests of the Stanford Revision applied to college students. *J. Educ. Psychol.*, 1919, 10, 477-485.
- G—CALFEE, M. : College freshmen and four general intelligence tests. *J. Educ. Psychol.*, 1913, 4, 223-229.
- G—CARTER, T. W. : Standardized mental test vs. an unstandardized English test as a measure of predicting success in college. *Sch. and Soc.*, 1927, 26, 151-152.
- T—CATTELL, R. B. : The assessment of teaching ability. *Brit. J. Educ. Psychol.*, 1931, 1, 48-72.
- C—CLEETON, G. V. : The predictive value of certain measures of ability in college freshman. *J. Educ. Res.*, 1917, 15, 357-370.
- S—College Entrance Examination Board, Comm. in Scholast. Aptitude Tests, 8th Ann. Rep., 1933.
- G—COLVIN, S. S. : Psychological tests at Brown University. *Sch. and Soc.*, 1919, 10, 27-34.
- F—COMPTON, R. K. : Informing college freshmen of their scholastic aptitude test scores. *Sch. and Soc.*, 1941, 53, 62-64.
- G—CONDIT, P. M. : The prediction of scholastic success by means of classification examinations. *J. Educ. Res.*, 1929, 19, 321-335.
- T—Predicting success of graduate students in a college of education. *Sch. and Soc.*, 1942, 56, 192-195.
- \*M—COWDERY, K. M., and EWELL, B.C. : Aptitude for medical school success. *Fac. Bull, Stanford Univ.*, 1931, No. 19, 3-4.
- G—CRANE, E. : An investigation of three plans for selecting students to be admitted to college. *J. Educ. Psychol.*, 1926, 17, 322-330.
- G—*et al.* : The results of some psychological tests at Bryn Mawr College. *Sch. and Soc.*, 1927, 25, 640-644.
- F—CRAWFORD, A. B. : Incentives to study. *Yale Univ. Press*, 1929.
- S—*et al.* : Forecasting freshman achievement. *Sch. and Soc.*, 1930, 31, 125-132.
- S—*et al.*, and BURNHAM, P. S. : Forecasting college achievement. *Yale Univ. Press*, 1946.

- T—CUFF, N. B., and DONOVAN, H. : Intelligence testing in teachers' colleges.  
 G——— : Prognosis and diagnosis of success in college. *J. Appl. Psychol.*, 1930, 14, 612-619.
- G—DALE, B. : The use of mental tests with university women students. *Brit. J. Educ. Psychol.*, 1935, 5, 59-75.
- G—DARLEY, J. E. : A study of clinical predictions of student success or failure in professional training. *J. Educ. Psych.*, 1938, 29, 335-354.
- G—DAVIDSON, M. R., and MACPHAIL, A. H. : Psychological testing in a women's college. *Person. J.*, 1927, 6, 266-275.
- \*F—DEARBORN, W. F. : Relative standing of pupils in the high school and the university. *Univ. of Wisconsin Bull.*, 1909, No. 312.
- \*S—DERFFLINGER, G. W. : The prediction of college success—a summary of recent findings. *J. Amer. Ass. College Registr.*, 1943, 19, 68-78.
- F—DILLEY, F. B. : Scholastic records of students failing a general intelligence test. *Sch. and Soc.*, 1931, 33, 311-314.
- G—DWYER, P. S. : The use of subcorrelations in determining the predictive power of high-school grades. *J. Educ. Psychol.*, 1938, 28, 673-680.
- F——— : The correlation between age at entrance and success in college. *Ibid.*, 1939, 30, 251-264.
- S—EDGERTON, H. A. : Academic prognosis in the university. Baltimore : Warwick and York, 1930.
- G—EDGERTON, H. A., et. al. : Statistical studies of a college freshman program. *J. Exp. Educ.*, 1940, 8, 247-252.
- \*S———, and TOOPS, H. A. : Academic progress. *Ohio State Univ. Stud., Contrib. to Admin.*, No. 1, 1929.
- G—EDHOLM, O. G., and GIBSON, O. H. : Examination results and an intell. test. *Lancet*, 1944, 247, 294-296.
- G—ELLISON, M. L., and EDGERTON, H. A. : The Thurstone Primary Abilities Tests and College marks. *Educ. Psychol. Measurement*, 1941, 1, 399-406.
- G—ERFFMEYER, C. E. : Intelligence tests as an aid in the diagnosis of academic maladies. *Sch. and Soc.*, 1924, 20, 317-320.
- EYSENCK, H. J. : Dimensions of Personality. London : Keegan Paul, 1947.
- F—FEDER, D. P. : An evaluation of some problems in the prediction of achievement at the college level. *J. Educ. Psychol.*, 1935, 26, 597-603.
- G—FICKER, C. E. : Predicting achievement in the liberal arts college. *Sch. and Soc.*, 1935, 42, 158-520.
- F—FINCH, F. H., and NEMZEK, C. L. : Prediction of college achievement from data collected during the secondary school period. *J. Appl. Psychol.*, 1934, 18, 454-460.
- G—FINDLEY, W. G. : The validity of a comprehensive examination for scholastic awards in New York State. *J. Exp. Educ.*, 1943, 11, 250-256.
- \*T—FRASIER, C. M. : Intelligence as a factor in determining student teaching success. *Educ. Adm. and Sup.*, 1929, 15, 623-629.
- \*T———, and HEILMAN, J. D. : Experiments in teacher college administration, III. Intelligence tests. *Educ. Adm. and Sup.*, 1928, 14, 268-278.
- F—FREEMAN, F. S. : Elusive factors tending to reduce correlation between intelligence test results and college grades. *Sch. and Soc.*, 1929, 29, 284-286.
- G—FREEMAN, F. S. : Predicting academic survival. *J. Educ. Res.*, 1931, 23, 113-123.
- F—GARNETT, R. L. : Some factors in college success. (No publisher, no date). Quoted by Pintner, R. Intelligence tests. *Psychol. Bull.*, 1935, 32, p. 459.
- G—GARRETT, H. E. : A study of the C.A.V.D. intelligence examination. *J. Educ. Res.*, 1930, 21, 103-108.
- G—GARRETT, S. W. : The Ohio State Psychological Examination : an instrument for predicting success in college. *Occupations*, 1944, 22, 489-495.
- G—GOODMAN, C. H. : Prediction of college success by means of Thurstone's primary abilities tests. *Educ. Psychol. Measmt.*, 1944, 4, 25-140.
- G—GRAVER, D., and ROOT, W. T. : The Thorndike intelligence test and academic grades. *J. Appl. Psychol.*, 1927, 11, 297-318.
- \*F—GRAY, E. F. : Validation of the University of Kentucky personal history blank. *Kentucky Person. Bull.*, 1938, No. 20.
- \*G—GUILER, W. S. : The selective value of mental test. *Elem. Sch. J.*, 1925, 25, 112-115.
- C——— : The predictive value of group intelligence tests. *J. Educ. Res.*, 1927, 16, 365-374.
- G—HAMMOND, H. P., and STODDARD, G. D. : A study of placement examinations. *Stud. in Educ.*, 4, No. 7.
- F—HANNA, J. V. : A comparison of Co-operative Test scores and high school grades as measures for predicting achievement in college. *J. Appl. Psychol.*, 1939, 23, 289-297.



F—HARRIS, D. : The relation to college grades of some factors other than intelligence. *Arch. Psychol.*, 1931, 20, No. 131.

F——— : Factors affecting college grades ; a review of the literature, 1930-1937. *Psychol. Bull.*, 1940, 37, 1250-166.

HARTOG, P., and RHODES, E. C. : An examination of examinations. London : Macmillan, 1935.

F—HARTSON, L. D., and SPROW, A. J. : The value of intelligence quotients obtained in secondary schools for predicting college scholarship. *Educ. Psychol. Measmt.*, 1941, 1, 387-398.

G—HELD, O. C. : An attempt to predict the success of University freshmen in their adjustment to scholastic work. Pittsburgh : Univ. of Pittsb. Press, 1933.

G—HESTON, J. C. : The use of non-verbal tests in the prediction of academic success. *J. Educ. Psychol.*, 1942, 33, 608-614.

F—HOLZINGER, K. J., and SWINEFORD, F. : The stability of a bi-factor solution. *Suppl. Educ. Monogr.*, No. 48. Chicago : Univ. of Chicago Press, 1939.

F—JOHNSON, D. M. : Confidence and achievement in eight branches of knowledge. *J. Educ. Psychol.*, 1941, 32, 23-36.

G—JOHNSON, J. B. : Predicting success or failure in college at the time of entrance. *Sch. and Soc.*, 1924, 19, 772-776.

G—JOHNSTON, J. B., and WILLIAMSON, E. G. : A follow-up study of early scholastic prediction in the University of Minnesota. *Sch. and Soc.*, 1934, 40, 730-738.

G—JONES, A. H. : The prognostic value of the low range Army Alpha scores. *J. Educ. Psychol.*, 1929, 20, 539-541.

F—JONES, E. S. : The grade-test correlation as an index of motivation. *Sch. and Soc.*, 1932, 36, 478-480.

F—JONES, L. : A project in student personnel service at the college level. *Sch. and Soc.*, 1928, 28, 765-768.

G—JORDAN, A. M. : Student mortality. *Sch. and Soc.*, 1925, 22, 821-824.

F—KELLOG, C. E. : Relative values of intelligence tests and matriculation examinations as means of estimating probable success in college. *Sch. and Soc.*, 1929, 30, 893-896.

G—KING, J., and McCORRY, M. : Freshmen tests at the State University of Iowa. *J. Educ. Psychol.*, 1918, 9, 32-39.

G—KIRKPATRICK, F. H. : Four mental tests and freshmen marks. *J. Higher Educ.*, 1940, 11, 38-39.

G—KITSON, H. D. : The scientific study of the college student. *Psychol. Rev. Mon.*, 1917, 23, I.

T—KRIEGER, L. B. : Prediction of success in professional courses for teachers. *Teach. Coll. Contrib. Educ.*, No. 420, 1930.

F—LAIRD, D. A. : A study of some factors causing a disparity between intelligence and scholarship in college students. *Sch. and Soc.*, 1924, 19, 290-292.

S———, and ANDREWS, A. : The status of mental testing in colleges and universities in the United States. *Sch. and Soc.*, 1923, 18, 594-600.

F—LANDRY, H. A. : The relative predictive value of certain college entrance criteria. *J. Exp. Educ.*, 1937, 5, 256-260.

G—LANGLIE, T. A. : Predicting scholarship. *J. Higher Educ.*, 1938, 7, 390-391.

\*G—LEFEVER, D. W. : The prognostic value of certain groupings of the test elements of the Thorndike Intelligence Examination. Univ. of Calif. Press, 1930.

\*G—MANNING, F. L. : How accurately can we predict success in college? *J. Amer. Ass. Coll. Registr.*, 1938, 14, 35-38.

G—MANUEL, H. T., et. al. : The new Stanford-Binet at the college level. *J. Educ. Psychol.*, 1940, 31, 705-709.

G—MATSHALL, M. V. : What intelligence quotient is necessary to success? *J. Higher Educ.*, 1943, 14, 99-100.

G—McCLELLAND, W. : Selection for secondary education. London : University of London Press, 1942.

G—McGEHEE, W. : The prediction of differential achievement in a technological college. *J. Appl. Psychol.*, 1943, 27, 88-92.

S—MACPHAIL, A. H. : The intelligence of College Students. Baltimore : Warwick and York, 1924.

S———, and BERNARD, W. : Ten years of intelligence testing. *Educ. Psychol. Measmt.*, 1943, 3, 157-165.

\*F—McQUITTY, J. W. : Student mortality in relation to scores on the Kentucky classification test. *Kentucky Person. Bull.*, 1937, No. 19.

G—MINER, J. B. : Scholarship and intelligence. *Person. J.*, 1927, 6, 113-118.

G—MITCHELL, W. B. : The revised Stanford-Binet for university students. *J. Educ. Res.*, 1943, 36, 507-511.

- F—MORRIS, E. M. : Some results secured in personal work in a teachers' college. *Sch. and Soc.*, 1934, 574-576.
- M—MOSS, F. A. : Scholastic aptitude tests for medical students. *J. of Ass. of Amer. Med. Coll.*, 1933, quoted by Pintner, R., Intelligence tests. *Psychol. Bull.*, 1934, 31, 453-475.
- M—MOSS, F. A. : Aptitude tests in selecting medical students. *Person. J.*, 1931, 10, 79-94.
- F—MUNROE, R. L. : Prediction of the adjustment and academic performance of college students by a modification of the Rorschach method. *Appl. Psychol. Monogr.*, 1945, No. 7.
- G—MURRAY, E. : Freshman tests in the small college : *J. Appl. Psychol.*, 1923, 7, 258-276.
- G—NELSON, M. J., and DENNY, E. C. : The Terman and Thurstone group tests as criteria for predicting college success. *Sch. and Soc.*, 1927, 26, 501-502.
- G—NEMZEK, C. L. : Intelligence tests at the college level. *J. Educ. Res.*, 1932, 26, 617-618.
- \*G—ODELL, C. W. : Predicting the scholastic success of college students. *Univ. Illinois Bull.*, 1930, 28, No. 5.
- \*G—OWENS, W. A. : On a certain value of the Thurstone tests in predicting scholarship. *Educ. Admin. and Super.*, 1927, 13, 495-499.
- G—PAYNE, A. E., and PERRY, J. D. : The intelligence ranking of 250 Cith College honor students. *Sch. and Soc.*, 1935, 42, 387-384.
- \*G—PERRIN, F. A. C. : The freshman psychological examination in the College of Arts and Sciences of the University of Texas. *Univ. of Texas Bull.*, No. 2445, 1924.
- G—PETERSON, B. H. : Can you predict who'll be able to go to college ? *Sch. and Soc.*, 1941, 61, 26-27.
- F—PIERSON, C. D., and NETTLES, C. H. : A study of high school seniors to determine who shall be recommended to college. *Sch. and Soc.*, 1928, 215-216.
- T—PINSENT, A. : Pre-college teaching experience and other factors in the teaching success of university students. *Brit. J. Educ. Psychol.*, 1933, 3, 99-126, 201-221.
- \*T—PYLE, W. H. : The relation, between intelligence and teaching success ; a supplementary study. *Educ. Admin. and Super.*, 1928, 14, 257-267.
- G—QUAID, T. : A study of the prediction of college freshmen marks. *J. Exp. Educ.*, 1938, 8, 350 to 375.
- F—REED, H. R. : The place of the Bernreuter Personality, Stenquist Mechanical Aptitude, and Thurstone Vocational Interest tests in college entrance tests. *J. Appl. Psychol.*, 1941, 25, 528-534.
- C—READ, C. B. : The relationship of scholastic averages in various semesters. *Sch. and Soc.*, 1940, 52, 468-469.
- G—REITZ, W. : Forecasting marks of New Plan students at the University of Chicago. *School Rev.*, 1935, 43, 34-48.
- F—RIGG, M. G. : The relation of college achievement test to grades and to intelligence. *J. Educ. Psychol.*, 1939, 30, 397-400.
- T—ROGERS, A. L. : The intelligence of women graduates of colleges of Liberal Arts entering the teaching profession. *Sch. and Soc.*, 1924, 19, 201-202.
- S— : Mental tests for the selection of university students. *Brit. J. Psychol.*, 1924, 15, 405-415.
- G— : Mental tests as a means of selecting and classifying students. *J. Educ. Psychol.*, 1920, 11, 181-187.
- G—ROWLAND, E., and LOWDEN, G. : Report of psychological tests at Reed College. *J. Exp. Psychol.*, 1916, 1, 211-217.
- G—ROOT, W. T. : The freshman Thorndike College Entrance Tests, first Semester grades, Binet tests. *J. Appl. Psychol.*, 1923, 7, 77-92.
- G—RUEDSLIL, C. H. : Differential prediction of college achievement in the College of Letters and Science of the University of Wisconsin. *Psychol. Bull.*, 1941, 38, 553.
- T—SAER, H. : A further investigation of pre-college teaching experience and other factors in the teaching success of university students. *Brit. J. Educ. Psychol.*, 1941, 11, 183-196.
- F—SAPPENFIELD, B. R. : Prediction of college scholarship for groups having effort endices of restricted range. *J. Appl. Psychol.*, 1943, 27, 448-451.
- C—SARBIN, T. R., and BORDIN, E. S. : New criteria for old. *Educ. Psychol. Meas.*, 1941, 1, 173-186.
- G—SARTAIN, S. Q. : A comparison of the new revised Stanford-Binet, the Bellevue Scale, and central group tests of Intelligence. *J. Soc. Psychol.*, 1946, 23, 237-239.
- C—SCHMITZ, S. B. : Predicting success in college ; a study of various criteria. *J. Educ. Psychol.*, 1937, 28, 465-473.

- F—SCOTT, C. M. : Background and personal data as factors in the prediction of scholastic success in college. *J. Appl. Psychol.*, 1938, 22, 42-49.
- F—SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION : The Prognostic Value of University Entrance Examinations in Scotland. Univ. of London Press, 1934.
- \*S—SEGEL, O. : Prediction of success in college. *U.S. Off. Educ. Bull.*, 1934, No. 15.
- G——, and GERBERICH, J. R. : Differential college achievement predicted by the American Council Psychological Examination. *J. Appl. Psychol.*, 1933, 17, 637-645.
- F—SELOVER, R. B., and PORTER, J. B. : Prediction of the scholarship of freshmen by tests of listening and learning ability. *J. Appl. Psychol.*, 1937, 21, 583-588.
- \*F—SEYLER, G. C. : The value of rank in high school graduating class for predicting scholastic achievement in college. *J. Amer. Ass. Col. Registr.*, 1941, 16, 123-133.
- F—SMITH, F. : The use of previous record in estimating college success. *J. Educ. Psychol.*, 1945, 36, 167-176.
- \*T—SORENSEN, H. : Why teaching success does not correlate highly with measured intelligence. *Educ. Adm. and Superv.*, 1929, 15, 602-606.
- G—STALNAKER, J. M. : American Council Psychological Examinations for 1926 at Purdue University. *Sch. and Soc.*, 1926, 22, 86-88.
- G—— : A four-year study of the freshmen class of 1935 at the West Virginia University. *J. Educ. Res.*, 1942, 36, 100-118.
- F—STARRAK, J. A. : Matching ability and achievement. *J. Higher Educ.*, 1937, 8, 315-320.
- F—STRONG, G. K. : Vocational Interests of Men and Women. Stanford ; Stanford University Press, 1943.
- G—SYMONDS, P. M. : The significance of intelligence tests in the University of Hawai. *Sch. and Soc.*, 1924, 20, 601-606.
- TAYLOR, H. C., and RUSSELL, J. T. : The relationship of validity coefficients to the practical effectiveness of tests of selection. *J. Appl. Psychol.*, 1939, 23, 565-578.
- F—TEMPLEMAN, W. D. : Vocabulary and success in college. *Sch. and Soc.*, 1940, 51, 221-224.
- \*S—TERMAN, L. M., et. al. : Report of sub-committee of committee on scholarship and student ability. Stanford Univ., 1923.
- C—THARP, J. B. : Sectioning classes in Romance languages. *Mod. Lang. J.*, 1927, 12, 95-114.
- G—THOMSON, G. H. : The distribution of intelligence among university and college students. *Brit. J. Educ. Psychol.*, 1945, 15, 76-79.
- G—THORNDIKE, E. L. : An empirical study of college entrance examinations. *Science*, 1906, 23, 893-845.
- TIFFIN, J. : Industrial Psychology. New York. Prentice Hall, 1943.
- G—TOLL, C. A. : Scholastic aptitude tests in Amherst College. *Sch. and Soc.*, 1928, 28, 524-528.
- S—TOOPS, H. A. : The status of University Intelligence Tests in 1923-1924. *J. Educ. Psychol.*, 1926, 17, 23-36 ; 110-124.
- S—TRAXLER, A. G. : What is a satisfactory I.Q. for admission to college? *Sch. and Soc.*, 1940, 51, 462-463.
- \*G—TREDICK, V. P. : The Thurstone Primary Mental Abilities tests and a battery of vocational guidance tests as predictors of academic success. *Penn. St. Coll. Stud. Educ.*, 1942, No. 24, 58.
- G—TROYER, M. E. : Unpublished, quoted by Heston, C., 1942.
- T—TURNBULL, C. H. : The influence of previous teaching experience on results obtained by students in a university department of education. *Brit. J. Educ. Psychol.*, 1934, 4, 1-10.
- \*F—UPSHALL, C. C., and MASTERS, H. V. : An evaluation of certain tests and information for predicting success in normal school. *U.S. Dept. of the Int., Off. of Education Bull.*, No. 12, 1931, 21-32.
- G—VERNON, P. E. : Educational abilities of training college students. *Brit. J. Educ. Psychol.*, 1939, 9, 233-250.
- G—VITELES, M. S. : A comparison of three tests of "general intelligence." *J. Appl. Psychol.*, 1922, 6, 391-402.
- \*T—WADDELL, C. W. : The prognostic value of Army Alpha scores for success in practice teaching. *Educ. Admin., and Superv.*, 1927, 13, 577-592.
- \*T—WAGENHORST, H. : The relation between ratings of student teachers in college and success in first year of teaching. *Educ. Admin. and Superv.*, 1930, 16, 249-253.
- \*S—WAGNER, M. E. : A survey of the literature on college performance prediction. In : Jones, E. S. (Ed.) : Studies in articulation of high school and college. *Univ. of Buffalo Stud.*, 1932, 9.
- G——, and STRABEL, E. : Predicting success and failure in college ancient and modern foreign languages. *Mod Lang. J.*, 1935, 19, 285-293.

G——— : Homogeneous grouping as a technique for improving prediction coefficients. *Sch. and Soc.*, 1934, 40, 887-888.

\*G—WAUGH, K. A new diagnosis of the college student. *New York Times Mag. Supp.*, Jan. 2, 1916.

G—WEBER, J., et. al. : Success in the graduate school. *J. Higher Educ.*, 1942, 13, 19-34.

G—WHITE, H. D. J. : An application of mental tests to university students. *Brit. J. Educ. Psychol.*, 1931, 1, 297-295.

G—WHITNEY, F. L. : Intelligence Levels among state normal-school graduates. *J. Educ. Res.*, 1923, 7, 229-235.

\*T———, and FRAZIER, C. M. : The relation of intelligence to student teaching success. *Peabody J. Educ.*, 1936, 8, 3-6.

T———, and LENEGER, H. W. : The college success and mortality of state teacher college freshmen as related to intelligence and high school achievement. *Educ. Admin. and Superv.*, 1930, 16, 668-674.

G—WILLIAMSON, E. G. : An analysis of scholastic aptitude of freshmen in some private colleges in Minnesota. *Sch. and Soc.*, 1931, 34, 674-680.

S——— : The decreasing accuracy of scholastic prediction. *J. Educ. Psychol.*, 1937, 28, 1-16.

F—WISSLER, C. : The correlation of mental and physical tests. *Psychol. Rev. Mon. Supp.*, 1901, III.

G—WOLF, R. R. : Differential forecasts of achievement and their use in educational counselling. *Psychol. Mon.*, 1939, 51, No. 1.

S, C—WOOD, B. D. : Measurement in higher education. Yonkers : World Book Co., 1923.

G—YUNN, K. S. : Primary mental abilities and scholastic success in the divisional studies of the University of Chicago. *Psychol. Bull.*, 1941, 38, 552.