CHAPTER IV

TESTING THE TESTS--INTELLIGENCE TESTS GO TO COLLEGE

Introduction

As the psychologists who had been mobilized for war returned to their academic departments in America's universities, they brought with them their new interest in psychometrics. Within months after the armistice, these faculty members were both finding new uses for the original Army Alpha test and introducing new intelligence tests.[1] With no accepted guiding definitions or proscriptions regarding the new examinations, psychologists were free of constraints.[2] They tinkered with both their new instruments and their new tools of descriptive statistics. President Robert Stuart Hyer of Southern Methodist University noted in 1919 that the initial interest in the new intelligence tests among faculty in higher education was to "test the tests, rather than to measure the mentality of the students."[3]

---

[1] J.W. Bridges, "The Value of Intelligence Tests in Universities," School and Society 15 (March 18, 1922):295.

[2] The stage was set for the introduction of these examinations even before the War. In 1916, at an American Psychological Association Meeting, Walter V. Bingham predicted many of the postwar applications: "The demand for psychologists to turn seriously to the task of testing undergraduates arises from the need of supplementary means of selecting from applicants for admissions: of adjusting the curriculum to the particular needs of the individual student; of adjusting the student to the curriculum; of assisting the employment office in placing the seniors in the right positions; and of measuring the results of instruction." published as :Walter V. Bingham, "Mentality Testing of College Students," Psychological Bulletin 14 (February 1917):70.

[3] Quoted in: H. T. Hunter, "Intelligence Tests at Southern Methodist University," School and Society 10 (October 11, 1919):437.

The psychologists' interest in testing the tests--particularly the Alpha--led to large scale experiments across the country in the period from 1918 to 1926. While these psychologists debated the definition of intelligence and the public debated the implications of the tests for the nation, America's colleges became large scale psychological laboratories.

Three universities played leading roles in this experimentation with tests. Stanford University, with Lewis M. Terman as its premier tester, innovatively applied the tests to tracking students and to comparing institutions and departments. Ohio State University, primarily through the work of Herbert C. Toops, led in such applications of the tests as advising students and sectioning classes.[4] Columbia University, the institution of Edward L. Thorndike and Ben D. Wood, led in introducing the "psychological examinations" directly into the admissions process. Additionally, Columbia's Teachers College established the criteria for the test's success--correlations and predictive statistical validity.

Hundreds of other universities tested the tests in their own ways with their own students, while psychologists and admissions officers scurried to inform each other of their innovations and findings. In the early 1920s the pages of America's education periodicals abounded with reports of experimental applications of the intelligence tests. Institutions, ranging from such Ivy League schools as Brown, Smith and Penn to state institutions including the Universities of Missouri and Washington to even relatively unknown, small colleges like

---

[4] Toops is an under-recognized leader in the field of testing. Not only did he directly produce the Ohio State Intelligence Test, he trained many of the future leaders in testing. Henry Chauncey, founding president of Educational Testing Service, and Harold Gulliksen, later to be one of Educational Testing Service's (and psychology in general's) luminaries, both studied under Toops. Chauncey in 1923-24 and Gulliksen in 1927-29.See: Chauncey, Henry, with Gary D. Saretzky. "Oral History with Henry Chauncey," March 28, 1977. ETS Archives Oral History Program. Educational Testing Service, Princeton New Jersey. p. I-11.

Hobart and Sweet Briar, reported their work with the new tests in the pages of such journals as School and Society and Education.

In some cases, the experimental applications of the tests were idiosyncratic and short-lived. One experimental use, however, met so directly the needs of educators who were forced to make difficult admissions decisions that it later blossomed into the Scholastic Aptitude Test.

Testing the Tests: Early Applications of the Intelligence Tests

The first applications of the tests were simply to correlate the examinations with other known data. For example, in January 1919, within two months of the armistice, the Purdue University faculty voted to administer the Army Alpha examination to all students. (Additionally, 90 faculty members voluntarily took the test.) Purdue then correlated scores on the Alpha with specific academic departments.[5] Similarly, at Ohio State University the central administration directed all students to take the Army Alpha Test in October of 1919. Ohio State, too, subjected its test scores--in this case scores for 5,950 students--to assorted statistical analyses. The Purdue and Ohio State results, along with

---

[5] G.L. Roberts and G.C. Brandenburg. "Army Intelligence Tests at Purdue University," School and Society 10 (September, 1919):777. Students in Chemical Engineering obtained the highest average; students in science, the lowest. In explaining the low score for science students, the researchers noted, "The science school was composed largely of women at this time, there being a total of only seventeen men in the school, so that this school may be taken as representative of the women of the university." Continuing with their assessment of women's ability, the researchers introduce the University of Illinois and estimate other national results, "At the University of Illinois the men proved superior in the tests by almost exactly this amount, and in every report on the test, so far as the writer know, men have made slightly better scores than women."

reports of studies at dozens other colleges were then reported nationwide in various education and educational psychology journals as the nation's psychology faculties rushed to publish their findings about the tests.[6]

Three important studies in 1922 and 1923 revealed the extent of the use of intelligence tests in American colleges and universities. Donald A. Laird and Annageaen Andrews sent survey materials to forty-three state schools, forty-three denominational schools and twenty one private schools "of collegiate grade;"[7] The Association of American College's Commission on Psychological Tests and Rating Scales surveyed four hundred schools;[8] and Herbert Toops of Ohio State surveyed 102 schools. In his case, Toops was not seeking to obtain precise percentages of schools using the tests, but "rather to determine certain

---

[6] George F. Arps and Ellis L. Noble, "University Students Intelligence Ratings According to the Army Alpha Test," School and Society 11 (February 21, 1920):233-237. The reports became countless; between 1918 and 1926 the single journal School and Society, published, in addition to aggregate surveys, individual articles discussing the uses and findings regarding intelligence tests on at least the following institutions: Oberlin, Brown, University of Illinois, Bryn Mawr, Yale, Syracuse, Dartmouth, Stanford, Hobart, William Smith College, Southern Methodist, Purdue, Illinois State Normal University, Sweet Briar, University of Washington, Northwestern, University of Pennsylvania, Columbia, Chicago, Ohio State, Smith College.

[7] Donald A. Laird and Annageaen Andrews, "The Status of Mental Testing in colleges and Universities in the United States," School and Society 18 (November 17, 1923):597. Laird found that by 1923 35% of the state universities he polled, 78 percent of the private institutions and 21% of the denominational schools were using intelligence tests as a routine or required part of their admissions process.
Furthermore, many other institutions were administering the tests to large portions of their entering students but were not requiring that all students take them.

[8] Richard Wellington Husband, "Psychological Tests and Rating Scales," School and Society, (April 19, 1924):443-450. Husband, who chaired the commission, reported that "about 80 percent of the institution reporting tests are required of all freshmen, usually a short time after admission to college. "In a few instances, they are also given to seniors a short time prior to graduation." p. 446. Brigham would later cite this study in explaining why the College Board was introducing its Scholastic Aptitude Test. See: Carl Campbell Brigham et. al. Scholastic Aptitude Tests: A Manual for the Use of Schools, Prepared by the College Entrance Examination Board. p. 8. [Document is housed in Educational Testing Services Archives.]

working principles which have been developed by universities throughout the country in regard to the use of tests."[9]

The results of the surveys were consistent; in Laird's words, the "spreading contagion of mental tests has not been sluggish in attaching itself to the administration of our higher educational institution."[10] The Association of American Colleges study concluded that there was no subject "of more interest to American Colleges than the subject of psychological tests."[11] The results of all of the surveys indicated that the tests were well distributed both geographically and by institution type throughout American Higher education.

Having recently embraced descriptive statistics as the foundation of their emerging profession, psychologists were quick to promote their new tools to educators in general. The Toops study indicated, however, that the impetus for the examinations in 1923-24 was still coming almost exclusively from the psychology departments.[12] In fact, other academic faculties were expressing no

---

[9] Herbert A. Toops, "The Status of University Intelligence Tests," Journal of Educational Psychology 17 (January and February, 1926):23-36, 110-124.

[10] Laird's survey consisted of 43 state schools (he did not include and normal schools, 43 denominational schools and 21 private schools. Only one state school was selected from any single state. The use of intelligence tests for admissions was geographically widespread. Laird specifically mentions The University of Vermont, University of Washington, Ohio State University, University of North Carolina, State University of New Mexico, Miami University, University of Kansas, University of Delaware, Rhode Island College of Education, University of Idaho, University of Kentucky, University of Minnesota and University of Wyoming among his state institutions. The private schools cited were similarly spread across the nation; they included for example, Pomona College, Swarthmore, Stanford, Johns Hopkins and Reed College. Some respondents to the Laird survey indicated that while they did not require tests, these new instruments were often prescribed. For example, the University of Nevada administered the test to 195 students out of a class of 230.

[11] Husband, "Psychological Tests and Rating Scales," p. 450.

[12] Toops, "The Status of University Intelligence Tests," pp. 23-36 and 110-124) Toops' information indicated that out of 102 colleges that reported use of intelligence tests, the psychology department was the responsible agency in 31 institutions; the registrar was

interest in implementing the examinations, and only a few administrative units within a very small number of colleges were seeking the exams. Writing in 1923, Lewis Terman lamented that "worthwhile evaluations in higher education will continue to be as rare as they now unhappily are until the rank and file of college and university teachers become able to think in more exact quantitative terms than they are yet accustomed to."[13]

If the impulse to apply the tests was not campus-wide, the potential applications nonetheless ranged across institutions.[14] Some of these applications were ephemeral; others, such as the use of intelligence tests for admissions purposes, have endured. The initial predominant use of the tests was not, however, for admissions purposes. Toops found no schools using the tests as the sole admissions criterion, and found only 19 among the 102 schools surveyed using the tests for admissions purposes at all. Toops also discovered that only 26 percent of the colleges using testing had the tests administered

---

responsible in 15; the Dean of Men in 15; the president in 8; the Dean of Women in 7; and the entrance board in only four. At other institutions the tests were administered by personnel ranging from physicians to student advisors. That the impulse to test was still from the psychology department coincides with consistent phenomena throughout the various levels of education. Hugo Munsterberg noted in 1900, "I do not think that if teachers go through psychological and pedagogical studies they really will suffer very much; they will do what they do with most studies--they will forget them." Hugo Munsterberg, "School Reform," Atlantic Monthly 85 (May 1900):666.

[13] Terman, Lewis, Editors Introduction to Ben D. Wood, Measurement and Adjustment series, Measurement in Higher Education (Yonkers on Hudson, New York: World Book Company, 1923), p. 4.

[14] The types of institutions employing or considering employing "intelligence tests" were quite varied. For example, The Hebrew Technical Institute in New York considered using both group tests then following up with individual tests for those candidates who performed poorly. Henry Goldman, "Memorandum on A Method of Testing Applicants for Admission to The Hebrew Technical Institute," Twenty Page Manuscript in possession of Professor Paul Goldman, Department of Educational Policy and Management, University of Oregon.

before the opening day of school.[15] Toops concluded that the tests were "primarily educational administrative devices for dealing with administrative and pedagogical problems of students, rather than the criteria of intelligence of applicants for admissions."[16]

In 1924 the four most frequent uses of intelligence tests in American higher education were (a) in determining dismissal for low scholarship--49 schools out of 110; (b) in encouraging extra effort in the case of unmotivated bright students--42 out of 110; and (c) in determining amount of school work a student shall be allowed to carry--34 out of 110; [and] (d) in "encouraging bright students to undertake graduate work"--25 out of 110.[17]

Intelligence Tests for Disciplinary Action

In contrast to the use of intelligence tests for admission to higher education--a procedure which would become and remain commonplace-- other applications tried by the colleges in the 1920s were less long-lived. One such use was for disciplinary action. The Ohio State University reported in 1923 that it was using the Army Alpha Test to recognize ability so that it could reprimand students who were not trying.[18] Ohio State used tests "particularly in cases of

---

[15] Toops, "The Status of University Intelligence Tests," p. 297.

[16] Ibid., p. 110. The pattern of extensive use of the tests after admissions for advising purposes continued into the 1930s. For example, The Ohio College Association introduced a state-wide intelligence testing program in 1928 in which high school seniors could take the Ohio State University Intelligence Examination. Administering the test prior to admission, in this case, was not to allow consideration of the results for admissions purposes; it was simply to make the testing procedure for advising more efficient. See: Charles Wells Reeder, "Forecasting Academic Success in the College of Commerce and Administration," in Ralph W. Tyler, Service Studies in Higher Education (Columbus: The Ohio State University, 1932), pp. 206-220.

[17] Toops, "The Status of University Intelligence Tests," p.121.

[18] Laird and Andrews, "The Status of Mental Testing," p. 597.

students delinquent in their work or as corroborating the need for dismissal or pointing out that the student is not trying or has too many outside activities."[19]

Dartmouth College, too, used the Army Alpha test to ferret out sloth among students. Charles L. Stone of Dartmouth's psychology department expressed the view that "one conspicuous objective at present is the elimination or explanation of cases of extreme disparity between intelligence and scholarship."[20] Stone contended that, "One very satisfactory way to detect idlers, men too heavily loaded with extra-curricular activities, and men with unusual capacity to develop their potential, is, if we may at least tentatively trust intelligence tests, to compare intelligence with scholarship percentiles of the individual men."[21] Yale contacted students whose performances were lower than that predicted. Using language that is less draconic than Stone's, Yale's Paul Burnham recalls, "There was quite a bit of effort to understand what the problem was and to give some help if that was possible."[22] This former director of admissions also candidly notes that when a student who had an average prediction performed superbly, the admissions office would typically say, 'Well

---

[19] Ibid.

[20] C.L. Stone, "The Significant of Alpha in College," Journal of Educational Psychology 13 (May 1922):301.

[21] Ibid. According to William Wilson of the University of Washington, other educators discussed the possibility of a "grade based on a ratio between what the student has accomplished and what it is believed that he should have accomplished." The intent of these educators was similar to Stone's, to assure that students worked to their potential. William R. Wilson, "Mental Tests and College Teaching," School and Society 15:634.

[22] Burnham, Paul S. with Gary D. Saretzky and David R. Hubin. Oral History with Paul S. Burnham. ETS Oral History Series, April 1985. p. 20.

that's great. Yale has motivated that fellow and he really has developed an interest.'"[23]

Stanford tried two apparently conflicting applications of the test. If grades disqualified a student, he would be given another chance if he could perform at the seventy-fifth percentile on the Thorndike Intelligence Test.[24] Stanford also reported, however, that "all students who test above the 80[th percentile] but do poor work are called in for conferences with the scholarship committee.[25] Terman noted in 1921 that "a number of universities that do not give mental tests to all their students do test all who are placed on probation."[26]

Intelligence Tests for Advising

Psychologists very quickly applied the tests to the process of advising students into specific courses, particular majors and even into vocations.[27] Brown University psychologists Stephen S. Colvin and Andrew H. MacPail reported that their institution was making extensive use of intelligence tests in advising. Between 1918 and 1925, Brown used the Army Alpha, the Brown University, the Thorndike Intelligence Examination and combinations of the Brown and Thorndike. Intelligence tests were but a part of Brown's approach to

---

[23] Ibid.

[24] Laird and Andrews, "The Status of Mental Testing" p. 597.

[25] Ibid., p. 597.

[26] Lewis M. Terman, "Intelligence Tests in colleges and universities," School and Society, (April 28,1921):482.

[27] Laird and Andrews, "The Status of Mental Testing," p. 597.

academic advising, but, in the words of Colvin and MacPhail, "these tests constitute a very vital part of that scheme . . ."[28]

Consistent with a general impulse to test the tests, Smith College psychologist David Camp Rogers advocated the use of intelligence tests for advising, but not simply because of the potential benefits for the students. Rogers contended that the tests could have pedagogical value in helping students choose a major suited to their abilities while simultaneously helping educators determine the epistemological and pedagogical underpinnings of disciplines. He presented a hypothetical example of a student who tests very well in sections of intelligence tests that depend mainly on comprehension and reasoning, but ranks very low in those sections requiring memory. If this student then performs poorly in chemistry classes, "It is then no great strain on the astuteness of the adviser to suggest that she has difficulty in memorizing the symbols and formulas of chemistry. To this she assents, and chemistry goes down in the list as a field in which good memory is valuable."[29]

Several schools introduced intelligence tests into vocational guidance. The University of Missouri reported that it was using the tests to provide "sound advice in the choice of a vocation."[30] Stanford's Terman discussed the role of aptitude testing in higher education in terms that reflected his "tracking" work in

---

[28] Stephen S. Colvin and Andrew H. MacPhail, "The Value of Psychological Tests at Brown University." School and Society 16 (July 29, 1922):117.

[29] David Camp Rogers, "The Intelligence Examination and College Problems," The Smith Alumnae Quarterly 14 (May 1923):270.

[30] Laird and Andrews, "The Status of Mental Testing," p. 597. By 1925, this use was widespread enough to prompt the Vocational Guidance Association to recommend guidelines. See: Morris S. Viteles, (Chair) "Psychological Tests in Guidance: Their Use and Abuse," Reprinted in School and Society, 12 (September 19, 1925):350-356.

the Palo Alto public schools and saw their potential impact as comprehensively involving the student from admission to career placement. Terman suggested that the tests could be useful to colleges in ways ranging from selecting students to later placing those students after graduation. He proposed using intelligence tests to compare students in social groups, to compare students in different academic departments, and to make provide a basis for recommendation of certain students to prospective employers. He argued that recommendations based on "intelligence scores and indications of special aptitudes . . . would soon come to have far greater weight . . . than recommendations based upon off-hand impressions."[31] The practice of using the tests--as opposed to "off-hand impressions"--for advising was popular despite a recognition among some that the additional information the tests provided was minimal. Paul Burnham of Yale recalls that the tests were not all that useful in advising. "It didn't take very long to find out that the student who came out with a spectacularly high profile in, say, mathematics, sciences and spatial relations, and a low verbal, already knew it. The profile was no surprise to him. And he already had made his plans in that way. He had elected to go into the engineering school."[32]

Intelligence Tests for Sectioning of Classes

In the interest of efficiency, many psychologists called for the use of intelligence tests in "sectioning classes." Rogers asserted to Smith College alumni that just as a college should appropriately be seeking homogeneity at admissions, it should also be seeking homogeneity within sections of courses. Rogers supposed that those parts of the intelligence examinations that were

---

[31] Terman in Wood, Measurement, p. 11.

[32] Burnham, "Oral History," p. 27.

influenced most by speed would be the best criteria for sectioning.[33] Using language compatible with the Boring's image of the mind as the motor of thought--a high horsepower motor would not require downshifting to climb hills--Rogers argued strongly that a student's speed with items indicated that student's intelligence.[34] By the end of the decade, the calls for careful sectioning came from varied quarters of campuses. For example, the Ohio State University's Romance Language Departments used the institutional intelligence examination as a factor in sectioning their classes to assure that the "bright student [is in] the company of his intellectual equals."[35]

The practice of using intelligence tests for sectioning was one of the applications that, with refinements, has continued.[36] In 1930 at the urging of University President George W. Rightmire, Ohio State University's College of Arts and Sciences began "a consistent program of discovering, encouraging, and stimulating the superior Freshmen," in order to prescribe a special "suggested reading and purchase list for thoughtful students in the College of Arts and Sciences." Based primarily on intelligence test scores, and only partially on academic achievement, the college chose a group of "superior" students, then

---

[33] Psychologists refer to such sections as "highly speeded."

[34] Rogers, "The Intelligence Examination and College Problems," p. 272.

[35] Robert E. Monroe, "Adapting Instruction to the Ability of the Student in The Romance Languages." in Ralph Tyler, Service Studies in Higher Education: Ohio State University Studies 15, (Columbus, Ohio: The Ohio State University, 1932), p.126. Monroe was Professor of Romance Languages and Supervisor of Elementary Courses in Romance Languages.

[36] Colleges quite commonly use both the SAT Verbal Section and Math Section for placement in writing and math sections respectively. However, using tests with such little diagnostic value in this manner is not even recommended by the test's authors.

sent them and their parents a letter of congratulations along with the reading list.[37]

Intelligence Tests for Comparisons of Institutions

Perhaps one of the most novel experimental uses of intelligence tests by colleges was to enhance their standings and reputations vis-a-vis their counterparts. Beginning in mid-1919, when Brown University published the Alpha test scores of its freshmen, large numbers of schools began to boast about high scores.[38]

Lewis Terman, never one to shy away from a possible comparison no matter how invidious it might be, stated most clearly the reason for such comparisons: "the value of a university's product is determined as much by the original quality of the raw material with which it works as it is by the salary, budget, or curriculum."[39] He argued in 1923 that "two universities may be regarded by the general public as in all respects equal when an amount of ability below the minimum requisite for bare survival in one would considerably exceed

---

[37] Huntley Dupre, "Encouraging the Superior Freshman Students in the College of Arts and Sciences." in Ralph Tyler, Service Studies in Higher Education, Ohio State University Studies 15, p.228. Huntley Dupre was then "Junior Dean of the College of Arts and Sciences.

[38] Colvin and MacPhail, "The Value of Psychological Tests at Brown University," p. 113. From 1920 through 1923 the pages of School and Society, a journal that drew non-technical publications from psychologists as well as educators and academic administrators, were replete with articles in which authors published the scores of their students and compared them to the average scores of colleges that had previously made such information public. After 1923, this practice of making invidious distinctions based on the scores on entering students apparently became more subtle; the unabashed bragging disappeared from the journals. However, the practice of establishing institutional prestige based on aptitude test scores certainly continues. For a discussion of modern institutional comparisons based on average entrance examination scores see: Alexander Astin "An Empirical Characterization of Higher Education Institutions" Journal of Educational Psychology 53 (February 1962):224-225; and, Alexander Astin and Lewis C. Solomon "Are Reputational Ratings Needed to Measure Quality?" Change, (October 1981):15-19.

[39] Terman, "Intelligence Tests in Colleges and Universities," p. 482.

the average in the other." The new intelligence tests would allow psychologists to demonstrate conclusively that "comparatively speaking, some student bodies are made up largely of intellectual cream while others have drawn almost entirely from the lower levels."[40]

Consistent with his desire to impose a ranking, Lewis Terman published a chart showing the percentage of alumni from twenty-eight different institutions who were listed in "Who's Who."[41] For eleven institutions, he then listed their median alpha scores to make the point that the new alpha data could help educators to rank institutions.[42] Terman, in fact, saw potential value in Stanford's publicizing its scores aggressively while simultaneously selecting its students based on high scores on intelligence tests. Noting that the Palo Alto University's provincial location was keeping it from surpassing the status of the established Eastern schools, Terman proposed that selective admissions based on high "intelligence test scores would be, if made public, beneficial to the institution."[43] In the face of such public comparisons of institutions based on their students' intelligence scores, H.T. Hunter displayed uncommon candor when he indicated that "the scores, particularly the medians, obtained at Southern Methodist University are considerably lower than those obtained at other institutions

---

[40] Terman, Lewis, Editor's Introduction to Ben D. Wood, Measurement in Higher Education, p. 7.

[41] Terman, "Intelligence Tests in Colleges and Universities," p. 482. Terman reprinted these data from a Study by B.W. Kunkel, "The Distinction of college Graduates," School and Society 2 (August 28, 1915):316-324.

[42] Terman, "Intelligence Tests in Colleges and Universities," p. 482.

[43] Ibid.

reporting."[44] He then, however, defensively attributed the difference to "the method of administering the examination."[45]

In his Measurement in Higher Education, Ben Wood published a detailed chart of comparisons of intelligence test score averages of students taking the Thorndike Intelligence Test at various institutions . Derived from the reports of the institutions themselves, Wood listed the mean and standard deviations of ten colleges including Columbia, Brown, Pomona, Stanford, and the University of California.[46] He was not shy about interpreting the results: "By using these sigmas in connections with Means, it would appear that less than fifteen percent of the University of California freshmen reach or exceed the mean intelligence score of men who survive two years in Columbia College."[47]

Intelligence Tests for Comparison of Departments

In the interest of finding out what the tests could tell them about the students and their institutions, psychologists made many comparisons within single institutions. For example, Brown University used the results of 1918 administrations of both the Army Alpha test and Brown's own psychological examination to make comparisons among groups of students seeking different degrees. At a time when the Bachelor of Science degree was still considered inferior to the Bachelor of Arts, the tests confirmed administrators

---

[44] Hunter, "Intelligence Tests at Southern Methodist," p. 437.

[45] Ibid.

[46] Wood, Measurement, p.41.

[47] Ibid., p.42.

preconceptions.[48] Here too, Terman saw potential comparisons; he sought to compare departments and colleges within particular institutions. In his 1923 introduction to Ben Wood's book, Terman presented great detail on institutional and departmental comparisons.

> At Ohio State University the median army Alpha intelligence test score of students in the Liberal Arts department was 147; of students in agriculture, 133; of dental students, 115. At the University of Illinois the median Alpha score of law students was 163; of students in agriculture, 139; of music students, 121. At Miami University the median intelligence score necessary to earn the grade of A was 162 in philosophy classes, 155 in chemistry, 145 in history, 132 in drawing or music, and 126 in home economics.[49]

Intelligence Tests for Admissions

In retrospect, all of the experimental uses in discipline, advising, sectioning, and institutional comparisons are significant primarily because they increased educators' familiarity with and confidence in the new exams; they also set the stage for the most significant application of intelligence tests in America's colleges. Lewis L. Thurstone of Carnegie Institute of Technology noted in 1919 that, as psychologists, their "main concern so far has been to establish the diagnostic value of mental tests as a criterion for admission." Typically psychologists gave tests to students who had already been admitted. But the real goal was clear. Thurstone mused:

> If we asked ourselves this question: If we had been given complete control over admissions could we have predicted by means of our

---

[48] Steven S. Colvin, "Psychological Tests at Brown University," School and Society, (July 5, 1919):29. The results of the Brown study showed a strong correlation between the new Brown University Test and the Alpha. The ranking by degree were: A.B. 73.7/74.3 (Brown/Alpha); Ph.B. 68.7/67.8; Sc.B. 66.11/62.8. Colvin remarked that "This difference which is definite and consistent for both sets of tests" confirmed other information on candidates "tabulated over a period of years," Ibid.

[49] Terman in Wood, Measurement, p. 7.

> tests which students were to fail and which were to succeed in their college work?[50]

Thurstone was prescient on this issue. The tests were used for many purposes, but one purpose was to become the most significant. Colleges started to use the tests for admissions.

During the last months of the war, educators saw and discussed a connection between the new tests used for the selection of officers and the selection of college students. Writing in August of 1918, Herman C. Bumpus, President of Tufts College, asserted that the same qualities that the army was looking for in its officers were those qualities that should be sought by colleges. He noted that in assessing intelligence, the army is looking for "adaptability, ease of learning, capacity to apply knowledge, and ability to grasp and to overcome difficulties." He then called for "readjustments in our entrance requirements as will recognize a wider range of intellectual activity."[51] Columbia's admissions committee chair, Adam Leroy Jones, contended that "among the many things taught by the organization of the National Army are to be found methods of selecting men for special purposes."[52]

Reports from the army psychologists on the potential value of their new tests led to several tentative, pre-armistice calls for incorporation of intelligence tests into admissions procedures. Columbia's Jones, for example, in the Report

---

[50] Louis L. Thurstone, "Mental Tests for College Entrance," Journal of Educational Psychology 10 (March 1919):129.

[51] Herman C.Bumpus, "Broadening Without Lowering College Entrance Requirements" Journal of Education 88 (August 29, 1918):174. [Reprint of an "address" to an unspecified audience.].

[52] Adam Leroy Jones, "Psychological Tests for College Admission," Educational Review 58 (1919):272.

of the Director of Admissions for the year ending June 30, 1918, under a caption "Future Changes," declared:

> It seems highly probable that the conditions in colleges and technical schools after the war shall have ended may call for still more careful administration of entrance requirements and for modifications which will look toward the rejection of all who are not clearly fitted for the work which they wish to undertake. Economy of the funds to be devoted to education and economy of the student's time and energy will forbid his admission to a course of study for which he is not fitted. Probably new tests may be devised which will help in solving the problem. The tests for general intelligence discussed for several years passed and more recently used in the army cantonments have not yet proved themselves sufficiently precise to be of service in solving our problem; but further experience may show how they may be used.[53]

The First Step: Intelligence Tests for Admitting Veterans

The first use of the Alpha for admissions purposes was a direct consequence of the war. Institutions sought to make special allowances for the non-traditional older student veteran. Stanford sought a way to admit veterans without having them return to high school. The Universities of Nevada and Arizona made special arrangements for disabled veterans as long as they scored well enough on the army's "psychological examination." [54]

William Proctor, an admissions counselor at Stanford, noted in 1922 that it had been inappropriate to ask veterans who sought to matriculate at Stanford but who had not completed high school to return for a diploma. Therefore, since

---

[53] Adam Leroy Jones, "Future Changes," Report of the Director of Admissions: Columbia University (June 30, 1918).

[54] William Martin Proctor, "Intelligence Tests as a Means of Admitting Special Students to Colleges and Universities." School and Society 16 (October 21, 1922): 474. Stanford was not alone in making unique provisions for veterans. The University of Pennsylvania made announcements "throughout the army" that it would be admitting "war specials." By 1919, one hundred and eighty-six veterans had been admitted to Pennsylvania based on their army "intelligence examinations" without meeting regular admissions requirements. "Intelligence Tests at the University of Pennsylvania," School and Society (November 8, 1919):548.

Stanford considered the Alpha test to be an indicator of ability, veterans were admitted based on their test scores.[55] Lewis Terman explained Stanford's action:

> Certainly a college is justified in permitting the exceptionally able candidate who is short in some of the usual academic requirements to enter by the test route. Properly safeguarded, the plan involves no risk whatever of lowering academic standards. Instead, it puts the emphasis on ability where it belongs. The candidate who can earn an exceptionally high test score in spite of inadequate training is the best possible bet as regards scholastic promise."[56]

Terman's remark here is interesting because he acknowledges a relationship between opportunity for education and performance on the tests while simultaneously contending that the tests measure aptitude or capacity for education.

Intelligence Test For Admission: Beyond Veterans

The use of the examinations in admission processes quickly expanded beyond just the considerations of veterans. The demographics of higher education and the psychologist's infatuation with the tests soon led to the insinuation of intelligence tests into admissions decisions generally. In explaining its introduction of the SAT, the College Board would, in 1926, note that during the

---

[55] Proctor, "The Use of Intelligence Tests," p. 472.

[56] Lewis Terman, in Editor's Intro to Ben D. Wood, Measurement In Higher Education, p. 8. Institutions other than Stanford enacted policies consistent with Terman's assertion; some even went further than applying them just to veterans. Toledo University reported in 1920 that any student over twenty-five years old could substitute an acceptable intelligence test score for "five entrance credits" "The Arts College and the City," School and Society, (December 25, 1920):632. Shortly thereafter, a mini-debate emerged in that periodical. A Northwestern University faculty member bitingly inquired, "I wonder whether an intelligence test is to be accepted as a substitute for one third of the graduation requirements, also." Roy C. Flickinger "Intelligence Tests and College Credits" School and Society, 13 (January 15, 1921):87. Then, a Carl Holliday advocated just such a plan of intelligence tests as a substitute for graduation requirements. Holliday argued, "Good Idea! I have long wondered what plan could be devised to 'choke off' some portion of the annual crop of B.A.'s." Carl Holliday "Intelligence Tests and Academic Credit" School and Society 13 (January 29, 1921):145.

thirty year period from 1890 to 1920, the number of students enrolling in American colleges and universities increased at a pace five times more rapid than the increases in population.[57]

The experience with the returning soldiers on a few campuses convinced admissions officers generally that using new tests could help them discover a student's aptitude and capacity for learning rather than simply relying on the extent of that student's previous schooling. Stanford reported that it had observed that the veterans who lacked requirements but were admitted to Stanford on their test scores alone performed better than students who had achieved similar test scores and had met the high school curriculum requirements.[58]

The number of high school graduates increased from 240,000 in 1915 to 528,000 in 1925.[59] This dramatic growth of the public high schools created a consistently expanding pool of college applicants. In the first decades of the century, enrollments in American colleges increased rapidly. In 1922 American colleges granted four times as many degrees as they had in 1890.[60]

At the close of World War I, admissions committees across the country faced unprecedented decisions; some educators recognized the change in the demographics they faced. In a February 19, 1919 edition of the Princeton Alumni

---

[57] "The Scholastic Aptitude Test of the College Entrance Examination Board," The Work of the College entrance Examination Board (Boston: Ginn and Co., 1926), pp. 46.

[58] Proctor, "The Use of Intelligence," Ibid. p.472.

[59] Valentine, The College Board, p. 31.

[60] Carl Campbell Brigham, with Roswell P. Angier, Andrew H. MacPhail, David C. Rogers, and Charles L. Stone, "The Scholastic Aptitude Test of the College Entrance Examination Board." (Report of the SAT Committee) New York: College Entrance Examination Board, 1926.

Weekly, a member of Princeton's Graduate Council wrote, "Princeton is facing some serious important problems in adjusting her entrance requirements and curriculum to meet the needs of the country after the war."[61]

The number of students applying for admission to college after the war exceeded the capacity of those colleges and universities. The increasing pressure on admissions is reflected in a comment by Princeton's President, John Grier Hibben: "The question of the limitation of the number of undergraduates required immediate solution."[62] Princeton had always simply grown as requests for enrollment grew. However, in January 1922, the Trustees decided that "it was the judgment of the board that with our present resources the number of undergraduates that can be adequately accommodated and properly taught is approximately 2000."[63] Radcliffe Heermance explained Princeton's new selective admissions criteria to an alumni audience in Atlanta in the following words: "The state university must under its charter produce quantity education, but there are certain national institutions, in the East particularly, that because of their charter, their traditions, and so forth, can produce quality education."[64]

---

[61] Andrew C. Imbrie, "Entrance Requirements and the Curriculum" The Princeton Alumni Weekly 19 (February 19, 1919):375. Imbrie was a New York City Princeton alumnus of the class of 1995 and a member of the Graduate Council. Imbrie presented portions of correspondence he had received from fifteen other members of the Graduate Council.

[62] Quoted in Radcliffe Heermance, "The Growth and Development of Selective Admission at Princeton" (A privately published report to the President and Trustees of Princeton University, 1952), p. 2.

[63] From a 1922 report of the Special Faculty Committee on Limitation of Enrollment. Cited in George E. Tomberlin, "Trends in Princeton Admissions," Senior Thesis, Department of Sociology, Princeton University, April 1971.

[64] Heermance, Radcliffe, "The Operation of the Plan of Selective Admission: An Address at the Atlanta Meeting of the National alumni Association" Princeton Alumni Weekly 24 (April 9, 1924):549.

Princeton's Alumni Weekly reported in October 1924, that "the Committee on Admission seems to have had even a harder job than last in keeping the number of new students admitted down to the limit of the University's facilities."[65]

Princeton was not alone in facing this new admissions equation; in the early twenties, the elite among America's colleges faced the need to limit enrollments. As David O. Levine has demonstrated in the postwar period, "a small but critical number of liberal arts colleges enjoyed the luxury of selecting their student bodies for the first time."[66] President Lowell of Harvard noted that "a problem confronts many colleges from the large growth in the number of young men seeking higher education since the war."[67] The proper approach to this problem was not clear. Lowell contended that, since entrance examinations had been used simply to determine "whether a boy is or is not fit to do college work,"[68] to attempt to use the existing tests for competitive purposes would be "quite a different thing."[69] Lowell was groping for tools to use in competitive admissions policies but shrank from the possibility of using tests; the "means of ascertaining the qualities of an applicant" should, in Lowell's view, include "the school record, the advantages or difficulties under which his education has been

---

[65] Princeton Alumni Weekly 25 (October 1, 1924):1.

[66] Levine, The American College, p. 136.

[67] "Limitation of Enrolment[sic] at Harvard" The Princeton Alumni Weekly 24 (January 30, 1924):341.

[68] Ibid.

[69] Ibid.

acquired, his character so far as it can be ascertained, and not lest in importance what can be learned by a personal talk with the boy, whenever possible."[70]

Thus, unlike the pre-war admissions committees that had been primarily concerned with the problem of filling classes and with the question of assuring that particular candidates possessed the necessary skills for that institution, the postwar committees had to make distinctions among even the fully qualified candidates.[71] Adam Leroy Jones, Director of Admissions at Columbia, noted in November, 1919: "Increasing numbers of candidates for admission to college and an increasing variety of motives actuating these candidates have made proper selection both more difficult and more important."[72] Furthermore, public secondary education, in the immediate postwar period, increasingly sought to define its own curricular requirements. For example, in 1918 the Commission on the Reorganization of Secondary Schools of the National Education Association issued the following statement:

> In view of the important role of secondary education in achieving the objectives essential in American life, it follows that higher institutions of learning are not justified in maintaining entrance requirements and examinations of a character that handicap the

---

[70] Ibid.

[71] William E. Allen, "College Admissions," School and Society 14 (October 1, 1921):235-240. Allen phrased it "under these competitive conditions, who should be given a chance in college?" p. 236. See also: Pillsbury, Walter B. "Selection--An Unnoticed Function of Education," Scientific Monthly 12 (1921):62-74. The change in admissions function was difficult for some educators; David Levine cites the case of a Dartmouth trustee who inquired of President Hopkins, "Mr. President, do I understand rightly that you seriously propose sometime in the future to decline the application of somebody who really wants to enter Dartmouth? . . .Well, now I guess this is all right and I'll probably vote for it, but by God, I've got to have a little time on it after forty years of watching Dartmouth grab and hogtie every prospect that wandered into town with a hazy idea of sometime going to college somewhere." See Levine, The American College, p. 140.

[72] Adam Leroy Jones, "Psychological Tests for College Admission," Educational Review 58 (1919):271.

> secondary school in discharging its proper functions in a democracy.[73]

The changing nature of college admissions--the need to choose among many applicants who possessed minimal qualifications-- led to the use of intelligence tests in admissions.[74]

Columbia was the first major university to use intelligence tests for admission of its general student body. In February 1919 the Literary Digest reported that "a radical educational experiment in applied psychology, whose results will be watched with intense interest by institutions of learning all over the world, is announced to go into effect at Columbia University, New York in September."[75] The Literary Digest described the Columbia innovation as "subsitut[ing] psychological tests, to measure the students' general intelligence and mental alertness for the old system of entrance examinations which gaged only his scholarship."[76] The New York Times reported that behind Columbia's move was the conviction that "there is a considerable waste in using the expensive plant and operating force of an institution of higher education on great

---

[73] "Cardinal Principles of Secondary Education: A Report of the Commission on the Reorganization of Secondary Education, Appointed by The National Education Association," (Washington, D.C.: U.S. Department of Interior, Bureau of Education, Bulletin , 1918, No. 35.):19-20.

[74] Brigham, et. al. "The Scholastic Aptitude Test, p. 46. The Brigham committee noted the growth in applicants and asserted that "The natural consequence is that many institutions have sought to develop more adequate means for selecting from among the applicants those best fitted to profit by the opportunities offered." Another, most sinister, impulse to introduce the tests has been presented by Harold Wechsler, who contends that the tests were intended to exclude or justify the exclusion of Jewish students. See Wechsler's chapter seven "Repelling the Invasion: Columbia and the Jewish Student," in: The Qualified Student: A History of Selective College Admission in America (New York: John Wiley and Sons, 1977).

[75] "Psychological Tests for College Entrance," Literary Digest 60 (1919):26.

[76] Ibid.

numbers of students who are incapable of profiting seriously by higher education."[77]

Columbia's Director of Admissions, A.L. Jones, explained the purpose of the new tests to the New York Times in terms that continued the themes raised in the "new plan" discussions:

> The most radical departure will be the entire doing away with the old-style examinations that were given to establish the applicant's knowledge of the subjects required for admission to college. This will be covered by his school record, and the psychological tests will demonstrate whether he is qualified to continue his schooling."[78]

In another publication that same year, Jones explained further the need for the tests. He noted that "a good many stupid boys have succeeded in one way or another in acquiring degrees from reputable colleges. A reliable test of intelligence would have excluded them to the advantage of all concerned."[79]

As was the case with others in the testing community, Jones was not bothered by the lack of a definition for this intelligence that he sought to use as a criterion for admissions to college. He commented:

> Without trying to give an exact definition we may at any rate recognize the fact that intelligence implies the presence in some degree of the following: power of attention, memory, ability to discriminate among facts and ideas, and range of information."

Jones confidently asserted that although "no one would attempt to say with finality precisely what each question measures. . . in the tests which have been

---

[77] New York Times. Cited in "Psychological Tests for College Entrance." Literary Digest 60 (1919):26-27.

[78] Quoted in "Fewer and Better Students," New York Times (February 22, 1919), p. 10. The general thrust of this article was to question the wisdom of Columbia's plans to adopt these tests.

[79] Adam Leroy Jones, "Psychological Tests for College Admission," Educational Review 58 (1919):272.

devised each question or problem calls for the exercise of certain of these capacities."[80]

Among those colleges and universities that introduced intelligence tests for considerations in admissions, most initially treated the scores as simply one piece of information about applicants. However, in 1922 Colvin conducted a study at Brown to project the impact of a plan to "refuse to register candidates who score low psychologically."[81] Colvin contended that the results of the study made it clear that "the method of using low psychological scores (before the men have begun academic work at all) as a criterion for exclusion is superior to dismissing men at the end of the first semester because their work is unsatisfactory."[82]

These intelligence tests were new to educators. They were consistent, however, with increasing interest in determining a student's aptitude rather than solely assessing his or her previous preparation. By the end of World War I, calls for colleges to admit students with "high intellectual qualities" were not new. President of Dartmouth Ernest N. Hopkins expressed the views of many with his statement that "requirements for entrance to American college have long been subject to criticism as so mechanical and formal as in many cases to exclude the men with capacity for superior intellectual accomplishment, while admitting the

---

[80] Ibid., p. 273.

[81] Colvin and MacPhail, "The Value of Psychological Tests at Brown University," p. 120.

[82] Ibid.

men of less potential ability, but with a record of more precise conformity to technical routine."[83]

There was nothing novel about an argument such as "there are a variety of habits of study, capacities to use reference material, ability to analyze, organize and present data, the possession of which should be regarded as indispensable to the effective pursuit of college studies."[84] The idea that an instrument could measure intellectual qualities directly, however, was new.

Private, independent college and universities were free to use the intelligence tests for selective admissions. Public state universities often felt constrained. The Ohio State University, a consistent innovator in the use of intelligence tests, attempted to resolve the tension between a desire to be more selective and the need to appear to serve all high school graduates. In 1929 Ohio State published "A Suggested Program of Pre-College Guidance for High Schools in Ohio."[85] In the introductory words of President Rightmire, the document was "designed to assist high-school principals, counselors, and

---

[83] Reported in "Dartmouth College Entrance Requirements" School and Society 11 (March 1920):314-315. Administrators and psychologists were not the only groups that suggested changes in admissions. At the University of Oregon, the student body noted in a 16 page publication expressed the opinion "we do not believe that every young person in the State of Oregon should attend an institution of higher learning." They indicated that "there still remains a flotsam population of students who do themselves no good by remaining in college, whose presence has an undesirable effect on others . . . and whose time wasted here is a distinct economic loss to the state." Noting that other institutions such as Dartmouth and Michigan were developing selective admissions, the student committee, with the support of student run Oregon Daily Emerald called for changes including the introduction of the use of "psychological examinations." "An Estimate of the Intellectual Activity within the University of Oregon. Suggested Changes in Administrative Policies." University of Oregon Bulletin 25 (May, 1926).

[84] G.H. Colebank, "Rational College Entrance Requirements," Journal of Education 89 (February, 1919):150.

[85] "A Suggested Program of Pre-College Guidance for High Schools in Ohio." The Ohio State University Press, Columbus, Ohio 1929.

teachers who are called upon to advise pupils and their parents with reference to readiness for, and success in, a college career."[86] This large university that dominated secondary education in Ohio both by its unchallenged position in higher education and through its school of education was attempting to "humanize the educational process."[87] The bulletin suggested that while small, privately endowed colleges "have adopted selection and exclusion to stem the tide in their direction, the doors of the state university, perforce, have remained open."[88] Noting that `humanization' must again replace standardization in college education, the University's recommendation to the high schools began: "This fundamental recommendation and request is made at the outset: Give at least one 'Ohio College Association Intelligence Test' to each senior in the high school."[89] Laced throughout the remaining thirty-five pages were correlation coefficients and endorsements of the predictive validity of intelligence tests.

Such widespread use of the examinations as that recommended by Ohio State would, in the view of the University, help the high school guide capable students toward higher education, while discouraging students who "are probably unfitted for successful work in college as the colleges are now organized."[90] Proscribed by its mission within the state from moving overtly toward selective

---

[86] Ibid., p. 4.

[87] Ibid., p. 5.

[88] Ibid., p. 5. This perceived pressure to maintain non-selective admissions was not an illusion. For example, in 1919, The New York Times had editorialized that "in the newer parts of the country" universities performed an essential public service by being inclusive in their admissions policies. "Fewer and Better Students," The New York Times, (February 22, 1919):10.

[89] Ibid., p. 10.

[90] Ibid., p. 9.

admissions, the Ohio State University could achieve some of the same result if high school counselors would administer intelligence test and use the results to advise students.

The Benefits of Intelligence Tests in Admissions

Educators characterized the impact of the changes in admissions "regarding the entrance tests as a service to the student as well as a protection to the institution."[91] Further, they discussed the benefits to the secondary schools. Speaking to the Connecticut Headmaster's Association, Professor Thomas H. Briggs of Teacher's College characterized the introduction of the new admissions plan as the "result of gradual change." The first step, he noted, was Columbia's abandonment of its own examinations in favor of those of the College Entrance Examination Board.

Briggs applauded the success of the early College Board essay examinations but indicated his sympathy with schoolmasters who felt that the need for intensive preparation for examinations limited their freedom to teach what they considered to be important. He cast Columbia's decision to introduce the intelligence tests in terms of the University's desire to give more responsibility to the secondary schoolmen. He continued, "But now, in order to grant the greatest freedom possible to the secondary schools, an alternative is proposed to the formal examination in subject-matter--an alternative, be it noted, not a required substitute."[92] Briggs indicated to the schoolmasters that

---

[91] Carl E. Seashore, "Progressive Adjustment Versus Entrance Elimination in a State University," School and Society, (January 13, 1923):29.

[92] Thomas H. Briggs, "The New Columbia University Admissions Plan," Education 39 (1919):474.

> what the tests are it is unnecessary to state. They have so little relation to the ordinary type of examination and they are being prepared so in multiple that it will be impossible for the most skilled coach or crammer to add more than one per cent to a students ability to pass them. This estimate, I think, is conservative, if not generous. Consequently, every principal who has an ideal of education--and who has not?--may feel perfectly free at once to abandon all thought of specific preparation for college entrance examinations, so far as Columbia is concerned, and to concentrate by all the means in his power on producing the best type of American citizen."[93]

Briggs noted that Columbia had been considering some change in admissions for several years. According to Briggs, Columbia hoped to achieve two things from the new plan: "first to select more surely the young men who can profit best by the courses that it offers; and, second, to free the secondary schools to teach as they have the vision, with no interruption for drilling their students on any details that do not seem in the fullest sense of the word essential for education."[94]

Criticisms of the Innovation

Not all psychologists were sanguine about the benefits of using intelligence tests in admissions. Some were simply reserved in their judgements of the usefulness of the tests; others were clearly opposed to their use. Princeton, a University that would later become a leader in testing and whose name is often associated with the SAT, was slow to introduce psychological examinations.[95] In 1926 Brigham wrote:

> Such tests have been given for four years at Princeton and we are beginning to understand how to use them. The amount of weight

---

[93] Ibid., p. 477.

[94] Ibid., p. 480.

[95] It is a common mistake for people to infer that Educational Testing Service, located in Princeton, New Jersey, is connected with Princeton University. Except for the historical tie of Carl Campbell Brigham this connection is spurious.

> we assign to a test score depends on the purpose for which is used. For the purpose of admission the test score does not earn, and therefore does not deserve as much weight as the regular Entrance examinations of the College entrance Examination Board. As an additional bit of evidence available in borderline cases, the test score is a very great help to the admissions committee."[96]

By 1925 Princeton used a concept of three criteria for admission. "We have three principal tests. First, character and promise. This must be emphasized because it is the most important." To measure this quality, Princeton used the "complete school record of every applicant." Second, Princeton used the College Entrance Examination Board examinations. After these examinations, candidates with adequate records and scores were admitted. The third evaluation was used in marginal cases. "Our third test (only used in doubtful cases) is the psychological. We are now giving these examinations to all entering men. Our psychological department is not yet sufficiently satisfied with these tests to make them a final criterion of fitness but we do use the results to guide us in admitting those who are otherwise short in their requirements."[97] Yale University demonstrated a similar ambivalence regarding the usefulness of the examinations. Yale predicted grades using the Princeton Bogey concept, but then admitted students who were predicted to fail. Yale's former director of admissions recalled in 1985, "in the earlier days there was quite a bit of skepticism about prediction and quite rightly so, of course. A

---

[96] Brigham, Carl C. "Intelligence Tests: The Third of the Present Series of Princeton Lectures by Members of the University Faculty" Princeton Alumni Weekly 26 (May 5, 1926):792.

[97] Smith, H. Alexander, "Entering College and remaining There Under the New Programme," The Princeton Alumni Weekly 25 (April 22, 1925):682.

certain number of instances arose in which the prediction didn't work out, either the student beat it markedly or else he fell flat on his face."[98]

Beyond holding reservations about the general application of the new instruments, some psychologists expressed specific criticisms. M.J. Van Wagenen, a psychologist at the University of Minnesota, contended that the Alpha tests were of little use: "It would seem that the tests prove least useful just where reliable results from their use is needed; namely, in eliminating those most likely to fail in their college work."[99] Johnston at the University of Minnesota acknowledged that the Army Alpha had value in predicting future scholastic achievement, but added that it was "on the whole inferior in predictive value to high school records."[100]

The criticisms were not simply about the Alpha. In a paper read before the Psychology section of the American Association for the Advancement of Science, James W. Bridges, a psychologist at the University of Toronto, criticized the use of any of the existing intelligence examinations in the admissions process. Noting that because of its length and variety of item types, one would expect that the Thorndike Intelligence Test would have the highest correlations with academic success, Bridges contended that, "the Thorndike Columbia tests are little if any better as a basis for prediction of academic achievement than other shorter examinations." Further, he argued that "all available data therefore point

---

[98] Burnham, "Oral History," p. 16. Burnham reports that as time passed and admissions officers gained faith in prediction, students were not admitted who were predicted to fail.

[99] M.J. Van Wagenen, "Some Results and Inferences Derived from the Use of Army Tests at the University of Minnesota." Journal of Applied Psychology 4 (1920):59-72.

[100] J.B. Johnston, "Test for Ability before College Entrance," School and Society 15 (April 1, 1922):345-353, 347.

to the same conclusion, namely, that the results of intelligence tests do not line up college students in the same order, anywhere near the same order, as they are lined up by academic achievement measured by grades earned."[101] Bridges concluded that, "either we must broaden the concept of intelligence to include all factors that make for achievement, or we must select some other objective criterion."[102]

To criticize the use of intelligence tests in the admissions process was not synonymous with overall rejection of the tests. For example, Bridges saw great potential value in the tests for diagnosing academic difficulties and for tracking in the public schools. He made clear that, "the value of their general use in colleges and universities is here seriously questioned. Specifically in regard to admissions decisions, their use is absolutely contraindicated."[103]

Some educators accepted some role for the tests in admissions processes but feared possible overemphasis. George C. Chambers, Director of Admissions at the University of Pennsylvania, strongly stated that "under no circumstances should the results of an intelligence examination, even a series of such examinations, be the sole criterion in deciding upon an application for admission to college."[104] Chambers indicated that "of all the various types of data

---

[101] Bridges, "The Value of Intelligence Tests," p. 298.

[102] Ibid., p. 300.

[103] Ibid., p. 302.

[104] George C. Chambers, "Intelligence Examinations and Admissions to College." Educational Review 61 (1921):128.

concerning prospective college students, the quality of his high school work is the most significant."[105]

Even among critics of the intelligence examinations, however, there was an impulse to move away from tests of knowledge. Chambers acknowledged that the student who has a good knowledge of the facts of mathematics will be able to

> create a better impression upon his instructors, but it is almost certain that he will not continue his lead more than a few weeks as compared with the boy with a higher grade of intelligence, even though the latter may not have his mathematical facts so fresh in mind when he starts his college freshman works.[106]

His solution was for higher education to refine the "certificate" system. He proposed that high schools be ranked in four categories of quality. Those schools in the "Class A" category should be given the privilege of certificating any of their graduates . . . to do college work successfully."[107] Schools in the lower three categories of quality would be able to certify students from decreasing factions of their student body. Students whom the schools could not certify would be given the option to take either the College Board Comprehensive Examinations or a combination of an English usage examination and an intelligence examination.[108]

---

[105] Ibid., p. 131.

[106] George C. Chambers, "Intelligence Examinations and Admissions to College." Educational Review 61 (1921):137.

[107] Ibid., p. 134.

[108] Ibid., p. 135.

The Major Tests Used Prior to the SAT

The tests used in the this period can be classified into a pattern. Initially, psychologists at various institutions used the Army Alpha examination in its original form. Because of frustrations with the Alpha, psychologists developed three other categories of examinations: at some institutions the testers made minor modifications in the Alpha; in other cases the changes were so significant that the test was essentially a different examination from the Alpha and took on a new name; and, in a few cases, notably those of Thorndike, Toops, Thurstone, and a consortium of scholars in the National Research Council, psychologists wrote new instruments addressed at national rather than individual institutional consumers. Thus in the first half of the decade of the twenties, American psychologists were using 1) the Alpha, 2) the Alpha with minor modifications, 3) institutional specific intelligence examinations, and 4) national examinations such as the Thorndike, Thurstone, Terman Group Test, and the National Intelligence Test.

The Alpha

Immediately after the war, most psychologists in higher education experimented with and applied the Army Alpha test in its original form. The test was easy and quick, requiring less than thirty minutes to administer. Yale University began using the Alpha test in 1919 and by 1920 reported correlations of the test scores and academic performance. Yale officials noted the difficulty in getting statistically significant correlations because Yale students clustered so near the top end of the scale; ninety-nine percent of Yale freshmen in 1919 would have qualified as officers had they taken the tests in the army. The Yale

admissions officers, nonetheless, believed that "tests [enabled them] to eliminate in advance many of the men who are misfits in our educational scheme.[109]

Soon, however, psychologists saw serious limitations inherent in the Alpha. Most significantly, it was scaled so low that it had no "statistical" or "predictive power" among tests takers earning top scores. Some psychologists criticized the Alpha but used it nonetheless. Charles Leonard Stone of Dartmouth College expressed a common sentiment among psychologists: "Despite the fact that the Alpha examination was designed for purposes very unlike academic functions, much interesting material has been gathered from colleges and universities in the past two years." Stone concluded that "all in all, the present Alpha would, from the standpoint of elective advisory purposes, seem to be as random an agent as the traditional campus method of selecting courses.[110]

Psychologists on many campuses made minor revisions to the Alpha to try to make the test more difficult and to give it more validity among the population of college students.[111] Early revisions to the Alpha often involved simply adding more items or shortening the time allotment to make speed more a factor .[112] By

---

[109] John E. Anderson., "Intelligence Tests of Yale Freshmen," School and Society 11 (April, 1920):418.

[110] C.L. Stone, "The Significant of Alpha in College," p. 298.

[111] Toops, "The Status of University Intelligence Tests," Toops was just one among many psychologists who recognized the shortcomings of the Alpha.

[112] See for example first Toops revision.

1920 psychologists were adding entire sections to the alpha; in these cases they would often rename the test.[113]

The Institutional Examinations

In November of 1922, David Camp Rogers, a professor of psychology at Smith College administered, an intelligence examination to 1,967 students out of the College's student body of 1,984. Rogers' test, The Smith College Intelligence Examination, reflected its Alpha ancestor; however, it included more difficult items and new formats for items. Rogers had previously made clear his position on the possible uses of intelligence tests in admissions processes; his purpose in giving the tests to the entire student body was to discover ways in which the examinations could "be an aid toward making education processes more effective."[114] His innovations were intended to create an instrument that would have more power to distinguish among capable test takers.

Smith College was not alone in creating its own test. Toops' study in 1923 found colleges using over thirty intelligence tests that were distinct enough to have separate names.[115] Each of these tests, however, was a descendent of the Army Alpha. With few exceptions the examinations included analogy items, sentence completion items, and general knowledge items. The formats of item types paralleled those on the army test. The Alpha section on general knowledge, for example, directly influenced the similar sections on such tests as the Ohio State Intelligence Test, the Smith College Intelligence Test, and the

---

[113] See for example, Toops subsequent revision renamed the Ohio State Intelligence Test and David Camp Rogers, Smith College Intelligence Test.

[114] Rogers, "The Intelligence Examination and College Problems" p. 269.

[115] Toops, "The Status of University Intelligence Tests," p. 28.

Princeton Intelligence Test. In some cases the items on these tests were almost identical to those on the Alpha. For example, the Alpha question "Sapphires are usually a) blue b) red c) green d) yellow" became "The topaz is usually a) red b) yellow c) blue" on the Ohio State Test.[116]

As the various psychologists attempted to refine or modify the Alpha and to produce their own exams, they varied in the degree of emphasis they gave to different aspects of the Alpha. In a few cases, they introduced completely new types of items. The Princeton Intelligence Test, for example, included items addressing mechanical ability and spatial reasoning. Depicting a series of gears and pulleys, the Princeton examination asked students to answer such questions as "If [spindle] A makes 2100 revolutions per minute with the belt in position 4 on step con A, what will be the speed of K, when K has 48 teeth."[117] Items such as this, i.e. items that lacked a precedent in the Alpha test, were however, the extreme exception. For the most part, the test directly reflected their common wartime ancestor.

Tests ranged in length. Some instruments could be completed in under thirty minutes; most significant among these examinations were the Army Alpha, the Terman Group Test, The University of Washington Test, and the Illinois General Intelligence Test. Most university tests, however, were longer than the Alpha because of added sections. The approximate one hour length of the Yale Test, the University of Texas Test, The Ohio State Test, and the Johns Hopkins Test was typical. The longest major test was the Thorndike which, among the

---

[116] For examples of item types from the major examinations, see appendix A.

[117] Princeton University Intelligence Test, Section Three, question 9.

major tests, deviated furthest from the Alpha in general format and in item format. Thorndike's test took three and one-half hours.[118]

The Thorndike Intelligence Test

The Thorndike Intelligence Test for High School Graduates, published in 1919, was the first intelligence test other than the Army Alpha to be used for admissions purposes. Edward L. Thorndike of Columbia's Teachers College was, in the 1920s, generally acknowledged as the leader in the mental measurement movement.[119] What eventually became the Thorndike Intelligence Test began as an instrument called the Thorndike Tests for Mental Alertness used by Columbia in 1918 to admit students specifically to its Army Training Corps. In 1919 Columbia adopted the modified version of the tests for admissions of regular students. Other institution followed, and by 1924, the refined Thorndike test was the most widely used test, displacing the Alpha's earlier lead.[120]

In 1919 the Literary Digest and the New York Tribune informed the public of the new Thorndike examinations at Columbia. Noting that details were not yet available, the two publications indicated that Thorndike would be drawing heavily from the work of Lewis Terman, whom the Tribune described as a "Professor of Education at Leland Stanford University, California, where experimental work in

---

[118] Toops, "The Status of University Intelligence Tests," p. 28.

[119] Geraldine Joncich, The Sane Positivist: A Biography of Edward L. Thorndike (Middletown, Connecticut: Wesleyan University Press), p. 443-444.

[120] Toops, "The Status of University Intelligence Tests," p. 121. The Name of the Thorndike exam was generally shortened to the Thorndike Intelligence Examination.

mental-measuring code first elaborated by Alfred Binet perhaps has been carried further than at any other institution in this country."[121]

The Thorndike, like the Army Alpha, did draw from Terman's work. Unlike the numerous institutional variations on the Alpha, The Thorndike examination also drew from item types that had been used at least since 1900. Also, unlike the institutional examinations, funding for the Thorndike examination came from a foundation. In 1914 the Carnegie Foundation, which, from that point forward would play a major role in admissions and achievement testing, asked Professor Thorndike to develop a "trial set of tests for students entering schools of engineering which might be standardized in respect to form, difficulty, administration, and scoring, and which might be more a measure of the pupil's future promise and less a measure of his previous educational advantages than the ordinary entrance examinations."[122] During the war, Thorndike refined his instrument and applied it to military situations. To meet the need for an examination for the Air Service, which would be more difficult than the Army Alpha examination, Thorndike introduced questions on arithmetic operations, syllogisms, "recognition of identities" and "marking absurdities." Ben Wood contends that the Thorndike was "designed to measure differences in a small part of the total range of human intellect. In this respect it contrasts most

---

[121] Literary Digest, Feb. 1919. and New York Tribune, February 6, 1919.

[122] Thorndike's work was reported in a several Carnegie Foundation publications. Including Charles Riborg Mann. A Study of Engineering Education: Carnegie Foundation for the Advancement of Teaching: Bulletin Number 11. [Chapter VIII, pp. 46-54, deals specifically with Admissions.] Mann highlights Thorndike's experiments with "mental tests" for admissions to Columbia, M.I.T., the University of Cincinnati, and Wentworth Institute. He then calls for further validity studies based on correlations with scores on the tests and future occupational success in the field of engineering, but noting that such correlations are not yet available, settles for correlations with school marks, "combined judgment of the students concerning one another's intellectual ability" and the judgment of teachers. Ibid., p. 50.

violently with the Army Alpha test which was designed to classify human intelligence into only five divisions." [123] The Thorndike examination divided the top two categories of the Alpha examination--A and B men or "superior and very superior"--into twenty different classes.[124]

The item types and configuration of various intelligence tests used by colleges, thus reflected their descendancy from the Binet. In general this heritage was passed on through the Army Alpha; in the case of the Thorndike, the heritage was direct as the author went back to the Binet itself.

Intelligence Tests and Statistical Correlations

Intelligence tests had more in common than the ancestry of their formats and items; they are all products of statistical correlations. This emphasis on correlations is also a direct descendent of Binet's work. The distinguishing aspect of Binet's test as opposed to the psychometric work of Wundt and Tichener was that Binet correlated his tests with academic performance.

Beginning with the first use of the Army Alpha for admissions, educators were interested in possible statistical correlations with the new tests; the quantified society was clearly triumphant.[125] Psychologists promoted the tests based upon the significance of their correlations with attributes that the educational community in general valued.

---

[123] Wood, Measurement, p. 29.

[124] Ibid.

[125] James Earl Russell, "Contributions of Measurement," National Eduction Association Addresses and Proceedings (Washington D.C.: NEA, 1926), p.200. Russell noted that "No contribution to the scientific movement has been of greater practical worth than the method of statistical data."

Thorndike proposed the specifications "for an ideal examination of the intelligence for college entrance" in May of 1920. These specifications were all couched in statistical terms with emphasis on correlations. His first requirement was statistically significant predictive validity. Given the "restricted range" of applicants to American universities, Thorndike acknowledged that Pearson "r" values would not be as high as .75, but indicated that educators should not settle for less than values of .5. In addition to high correlations with academic success, i.e., predictive validity, Thorndike called for high indices of reliability and for formal mechanisms of equating scores earned on different forms of the test.[126] Intelligence testing is an exercise that does not necessarily have validity to the lay person; non-psychologist educators had to be sold on the usefulness of the tests.[127] Making the nature of correlations understandable, even to educators, was not easy. Attempting to put the correlations in a context that quantitatively unsophisticated educators would understand, Stephen Colvin at Brown explained that

> from this data one may conclude that honor students need not be sought among those who make low psychological scores, that the chances are practically all set against the student who scores "medium" on the psychological tests getting more than one honor, and finally, that the men destined to achieve more than one honor

---

[126] E. L. Thorndike, "Intelligence Examinations for College Entrance." Journal of Educational Research 1 (1920):330. Not surprisingly, Thorndike's own Thorndike Intelligence Examination for High School Graduates met all of his specifications.

[127] For an example of approaches to informing non-psychologists of the salient points of descriptive and predictive statistics, see: Carl Campbell Brigham, "The Scholastic Aptitude Test of the College entrance Examination Board" in: College Entrance Examination Board, The Work of the College Entrance Examination Board 1901-1925 (New York: Ginn and Company, 1926), pp. 51-54. Brigham's chapter is a reproduction, with relatively minor omissions, of the first section of a manual prepared by the committee in charge of the Scholastic Aptitude Test.

> are, almost without exception, to be sought among those who make high psychological scores.[128]

Writing for an audience of educators, both at Columbia and at other institutions, Ben D. Wood also tried to explain the all important correlations. His book, Measurement in Higher Education, became a landmark in literature on the introduction of testing in our colleges and universities. Unlike Colvin, who had explained correlations in non-quantitative terms, Wood presented what were at that time esoteric concepts of predictive and descriptive statistics. Specifically, Wood attempted to inform his educator audience of the meaning of the then recently introduced "Pearson Product Moment R Correlation Coefficient," as well as the concepts of "validity," "reliability," "error of measurement," and "probable errors of estimates of true scores."[129]

Correlations with Academic Success

The obvious and important criterion of correlation was with the student's scholarship. Thurstone noted, in 1919, that

> to use any other criterion constitutes an indictment against the scholarship marks as a criterion for retention and promotion of

---

[128] Colvin and MacPhail, "The Value of Psychological Tests at Brown University," p.122.

[129] Wood, Measurement, pp.22-66. Fortunately for potential students from abroad, Wood was not a man of prejudice or xenophobia; he used his knowledge of statistics to warn of possible inappropriate conclusions from correlations. Simultaneous with the hearings on immigration restriction that were based in part on conclusions reached from the Yerkes studies, Wood stressed that "The intelligence tests is a highly specialized instrument, designed mainly to measure the one factor of mental alertness. It does not, unfortunately, measure the intellect directly, but something else which under certain conditions is almost if not quite synonymous with intellect. This "something else" that is measured in lieu of intelligence is very complex, and includes such elements as achievement in English language, information in specific fields, ability to think in terms of facts and relations in specific fields of information expressed in a certain language." Wood recognized, clearly to his credit, that students from foreign cultures should not be subjected to the examinations. "Differences in opportunity to learn English alone would disqualify him; and when we add the differences in general informational background, cultural habits,etc., the disqualification becomes complete and unquestioned."

[129] Ibid.

> college students. If our tests agree well with estimated intelligence but fail to agree with faculty action concerning dismissal of poor students and the promotion of good students, then our tests are of little use to the administrative officers of the college, unless they are willing to be guided by our statements of the student's intelligence rather than their own scholarship ratings.[130]

Such a correlation was not, however, without problems. Not all psychologists were as willing to concede that the comparative foil for their correlations should be the instructor's grade. Toops pointed out a major problem with these correlations--the lack of reliability of college grades themselves--because he could find only an average reliability of .66 between first and second semester grades. "Their improvement [reliability of college grades] is without a doubt, one of the next most important steps in raising the validity of predictions from college intelligence tests."[131]

Some psychologists developed more faith in their own exams than in the grading process. In his discussion of this particular correlation, Ben Wood, for example, criticizes the "unreliability in the measurement of college success," i.e. the college grades themselves. Wood points to many causes of the unreliability of grades but most clearly fears the subjective response to the "glib talker and artful bluffer."[132] He notes that "there is evidence that some very high college grades may be due to the ubiquitous expression of superficial observations in a

---

[130] Louis L. Thurstone, "Mental Tests for College Entrance," Journal of Educational Psychology 10 (March 1919):132.

[131] Toops, "The Status of University Intelligence Tests," p. 122. [It is noteworthy that, according to Robert Klitgaard, that correlation has not changed. Writing in 1985, Klitgaard notes that "college GPA's in adjacent semesters correlate only 0.6 to 0.7." Robert Klitgaard Choosing Elites (New York: Basic Books, 1985), p.203.

[132] Ben D. Wood, Measurement in Higher Education (Yonkers-on-Hudson, New York, World Book Company, 1923), p. 55.

loud voice and interested manner."[133] Further, he cautions that the weakness of objective grading weakens the important correlation of grades with intelligence scores.[134] He contended that since "a test cannot predict a criterion better than the latter forecasts itself, all hope of improvement in the intelligence tests prediction depends upon improvement in the reliability and significance of the criterion [college grades]."[135]

Despite these misgivings about the reliability of grades, psychologists continued to compute correlations, not only with the grades earned in college but also with the previously earned high school grades. The questions raised about the consistency of high school marks were even more common than those about college grades. In a 1925 Master's thesis, Harold Abelson of Teacher's College proposed a correction factor for high school marks. Feeder schools with large populations matriculating at particular colleges could be evaluated by the manner in which their graduates performed in college. Abelson found that he could improve the correlation coefficient between the Thorndike Intelligence Test and high school grades significantly by "correcting" high school grades.[136]

---

[133] Ibid., p. 55. Concern with reliability of grades was a consistent theme throughout Wood's career. He reiterated his criticisms in a speech before the American Association of Teachers Colleges in St. Louis, Missouri, on February 23, 1940. [Reprinted as Ben D. Wood, "Making Use of the Objective Examination as a Phase of Teacher Selection" Harvard Educational Review 10 (May 1940):278.] Further, Wood stressed his doubts about grades in an interview with this author conducted on May 23, 1985, Selection indicating that professor's grades "are sometimes unreliable because you can be influenced by the slick con-artist." Transcript and Tape of the Oral History are on file, for restricted use, at the Educational Testing Services Archives.

[134] Wood, Measurement in Higher Education, p. 123

[135] Ibid., p. 139.

[136] Harold H. Abelson, High School Marks Versus Psychological Test Scores in Power of Predicting College Success. A Master's Thesis submitted to the Faculty of Education, Columbia University, 1925, p. 5. [Available at Teacher's College Library] The contrast between

From the beginning of their use in admissions and in other roles, the defense of the intelligence test in higher education has been in its statistical correlations. After using the Thorndike Intelligence Test for one year, Columbia's Director of Admissions Albert L. Jones felt compelled to explain a relatively modest Pearson Product Moment r value of 0.65:

> The correlation between the work of the entire Freshman year for the students who entered by the New Plan and their marks on the mental test is 0.65. The most reliable data available indicate that the highest correlation that can be expected between the work of the Freshman year and the results of the usual college entrance examinations is about 0.45 to 0.55 . . .Every indication points to the mental as a most useful addition to our machinery of admission. [137]

Although psychologists were thus finding higher correlations between their new tests and student success than between the older examinations and student success, it is noteworthy that Herbert Toops, a statistician and industrial psychologist, had previously noted that "in the selection of employees in industry the results of tests for vocational aptitude must give a correlation of over .75 or .80 with actual achievement before said tests begin to be of financial assistance to the employers."[138]

This interest in the correlation of tests with performance led Carl Brigham to experiment with various combinations of intelligence test scores, essay type entrance examination scores and high school grades. Brigham prepared charts that depicted projected grades for incoming students, referred to as Bogey grades, and compared his projections to actual performances. The concept of

---

Abelson's approach and present day admissions officers is significant. In that period of "testing the tests" educators were investigating the correlation of tests with high school grades. Today, a school's average scores on the SAT are sometimes used to determine the significance of grades.

[137] Adam Leroy Jones, Report of the Director of Admissions, 1919-1920.

[138] Bridges, "The Value of Intelligence Tests," [Bridges paraphrases a communication with Toops.]

this statistical prediction became the foundation of the Scholastic Aptitude Test which Brigham developed for the College Entrance Examination Board in 1926.[139]

Correlations With Faculty Perceptions of Ability

Louis Thurstone focused several studies on the correlation between student performance on intelligence examinations and faculty perceptions of the student's ability. Thurstone introduced to the dialogue the concept of "critical scores," justifying their use by noting that "the method of critical scores has been applied successfully in standardizing Army Trade Tests."[140] Thurstone's purpose in using the "critical score," which he contended "emphasized prognosis with respect to each individual student rather than the abstract statement of relationship between the variables concerned," was to demonstrate that, among the 114 students he tested at Carnegie Institute, no "average or good student would have been eliminated by the mental test rating."[141]

Thurstone sought to demonstrate to faculty that their perceptions of students would be verified by objective tests. Drawing items from Walter Dill Scott's Completion Test, The Robinson Range of Interest Test, Opposites Test, Gordon Directions Test, and an analogies test, Thurstone created his own examination for this study.[142] He then interviewed faculty about students and correlated the results with those of the tests.

---

[139] Downey, Carl Campbell Brigham, p. 26.

[140] Louis L. Thurstone, "Mental Tests for College Entrance," Journal of Educational Psychology 10 (March 1919):141.

[141] Ibid.

[142] Thurstone, "Mental Tests for College Entrance," pp. 130-131.

Even educators outside the discipline of psychology soon became caught up in considerations of correlations of variables with the intelligence tests. At The Ohio State University in 1932, for example, historian Arthur H. Noyes devoted several pages to presenting correlations with The Ohio State University Psychological Examination in an article on "Some Recent Developments in History Instruction." Noyes, who at that time was an assistant professor and supervisor of the introductory courses in European history, concluded that: "there is a closer similarity between history marks and intelligence-test scores when the students are instructed in discussion sections than when they are in lecture-quiz sections."[143]

Correlations with Student Self-Perceptions

The emphasis on correlations with the new examinations was not limited to predicting academic performance. Wood was creative and far-reaching in the correlations that he presented. Beginning by correlating the Thorndike Intelligence Test with such academic variables as grades in college work and scores on other examinations, Wood then calculated correlation coefficients with "indicated morphological indices" including length of limbs, volume of trunk, lung capacity, height, weight, the ratio of height to weight, and the ratio of the length of the limbs to the volume of the trunk.[145]

---

[143] Arthur H. Noyes. "Some Recent Developments in History Instruction: The Lecture Versus the Discussion Method." in: Ralph W. Tyler, Service Studies in Higher Education Ohio State University Studies 15, (Columbus, Ohio), p. 152. Noyes reported correlation coefficients (Pearson "r") of .36 for discussion sections as opposed to .28 for lecture sections.

[145] Wood, Measurement, pp. 63-91. To his credit Wood acknowledges that the latter correlations with morphology are unimportant: "The results are interesting, but not convincing enough to form the basis of any significant remarks. Our main interest as college administrators is not the relation between morphology and intelligence." Ibid., p. 91.

Psychologists also correlated intelligence test scores with behavior patterns of students. For example, the University of Washington's William R. Wilson developed "correlation arrays" for 348 students that tested the relationship between intelligence test scores and how students apportioned their time. Wilson's concern was to find ways to assure that intelligent students had to work up to their potential.[146] Psychologists led in the use of correlation coefficients and did so with the tests that they were introducing to the college campuses. Investigations based on predictive validity and statistical correlations expanded to permeate both those publications relating to all aspects of college admissions and, in fact, educational research in general.[147]

Conclusion

The Army Alpha test as administered in World War tI stopped temporarily the experimentation in mental measurement. The period immediately after the war, however, saw extensive experimentation and application, particularly in higher education. As psychologists introduced the tests to their institutions, they experimented to discover what the tests could accomplish and how they should be best configured.

Psychologists compared the power of different item types and refined the configuration of their tests. Terman himself discussed his recognition of the value of one type of verbal item as opposed to another. He noted, for example

---

[146] Wilson, "Mental Tests and College Teaching," p. 633. Wilson found that students in the "upper quartile of intelligence" studied only 23 hours per week while student in the bottom quartile studied 49 hours per week.

[147] An early example of the use of statistical correlations without emphasis on psychological tests was Charles W. Odell's Ph.D. Dissertation at Teacher's College. "Predicting the Scholastic Success of College Students." Published as Bureau of Educational Research Bulletin No. 52, (Urbana, Illinois: University of Illinois, 1930).

that, "I have found that the ability to identify two words as synonyms is an appreciably better measure of the kind of intelligence to do school work than is the ability to identify two words as opposites."[144]

In response to institutional interest in using the new tests for admissions purposes, the College Board introduced the Scholastic Aptitude Test in 1926. Carl Campbell Brigham, in his 1926 manual that accompanied the introduction of the Scholastic Aptitude Test, pulled the themes of the psychologists together. Discussing the issues of admissions in light of "statistics concerning higher institutions [that]plainly show the numerical increase of college populations," he showcased the Scholastic Aptitude Test as a necessary component of the admissions decision.[145] Further, in that same document, Brigham carefully explained the methods of statistical correlation and even devoted a paragraph to clarifying "a popular misconception" about correlation coefficients.[146] His introduction of the SAT portrayed it as a response to new demographic factors, a direct outgrowth of the institutional examinations, and a complement to other College Board activities.[147]

The Board retained its dominance in the field of testing for admissions but did so at the expense of further significant experimentation. Testing took its early steps toward becoming a product of commodity science.

---

[144] Terman, "Intelligence Tests in Colleges and Universities," p.483.

[145] Carl Campbell Brigham et. Al. Scholastic Aptitude Tests: A Manual for the Use of Schools, Prepared by the College Entrance Examination Board. p. 5. [Document is housed in Educational Testing Services Archives.]

[146] Ibid., p. 8.

[147] Ibid., p. 1.