# The Use of the Selective Service College Qualification Test in the Deferment of College Students[1]

Henry Chauncey

*Educational Testing Service, Princeton, New Jersey*

EXPERIENCE, as Benjamin Franklin pointed out some time ago, keeps a dear school. Certainly as a nation we have paid heavily to acquire the wisdom on which present policies with respect to deferment of college students are based. In the press of immediate military needs during World War II, we lost sight of some important long-range requirements. Now that scientific and specialized personnel are in short supply to an unprecedented degree, and at a time when our technological superiority must be maintained at any cost, it constitutes an expensive lesson to realize that World War II training restrictions are responsible for a considerable measure of the present shortage. Estimates place the deficit attributable to this factor at 100,000 graduates with a bachelor's degree in science and 6000 scientists with the doctor's degree.

The Selective Service regulation which provides that college students high in educational aptitude may be deferred to continue their training is an attempt to avoid repetition of this same costly error during the present emergency. The Selective Service College Qualification Test, designed to implement this regulation, constitutes one means by which high educational aptitude may be demonstrated. Rank in class provides an alternate method. This article summarizes some points of interest in the development of the College Qualification Test and presents some of the findings now available from the first series of administrations in the spring and summer of 1951.

The decision by Selective Service Headquarters to utilize a nationwide test of scholastic aptitude in the draft deferment program, along with rank in class, was based on the recommendations of the six Scientific Advisory Committees, appointed in 1948 by Lewis B. Hershey, director of the Selective Service System. The committees' report, presented by M. H. Trytten, chairman, was made in the early fall of 1950. Shortly thereafter a subcommittee representing civilian and military testing experts prepared general specifications for a test of scholastic aptitude which would be suitable for use as a means of qualifying for deferment. The Educational Testing Service was designated by the subcommittee as the testing agency best equipped to carry out the construction of such a test and its administration on a nationwide scale—provided, of course, that the plans for the student deferment program, at that time still tentative, received final approval.

That the plans were approved, as publicly announced in an Executive Order on March 31, 1951, is a matter of history. Not so well known is the fact that the development of the test had been in progress for some months prior to that date. Realizing that a testing program like the one envisaged would require extensive preliminary preparation, in November 1950 Selective Service negotiated a planning contract with the Educational Testing Service, which provided for the design of test forms and procedures. Only because of General Hershey's farsightedness in authorizing us to proceed with the necessary arrangements while the program was still under consideration was it possible for the first nationwide administration of the SSCQT to be held less than two months after the date of the Presidential Order.

General Hershey's foresight also resulted in increased efficiency of execution of the necessary research studies by making it possible for us to tie the research directly into the operational activities connected with the administration and scoring of the test. From such details as the most appropriate design of the test answer sheets, and the setting aside of a representative sample of papers for detailed analysis, to the fairly complicated procedures involved in obtaining data on college performance for certain candidates and collating these with data on test performance, the advance planning made possible a nicety of integration of research with operational procedures that could not otherwise have been effected.

In the development of the test itself, we proceeded on the basis of directives specifying a test of scholastic aptitude that would yield a single composite score weighted about equally with linguistic and quantitative abilities, that would be composed of items selected for maximum validity in predicting general college achievement, that would have a sufficiently ample time limit so that it was not primarily a speed test, and, finally, that would be equated at two points to the Army General Classification Test. For freshmen, sophomores, and juniors, the passing score on the test (subsequently assigned a numerical value of 70) was

[1] Based on an address presented before the Conference on Scientific Manpower, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, Pa., Dec. 30, 1951.

to reflect a level of ability comparable to that represented by a score of 120 on the Army General Classification Test; for seniors, the passing score (subsequently assigned a numerical value of 75) was to be comparable to a score of 130 on the Army General Classification Test.

Following these directives, the staff made plans for a test consisting of 150 five-choice items, about equally divided among those measuring verbal ability and those measuring mathematical or quantitative ability. Some of the verbal ability items were designed to test reading comprehension; others were designed to test understanding of verbal relations. The quantitative ability items were also of two kinds: those drawing upon arithmetic reasoning, and those drawing upon data interpretation, for their correct solution. Detailed plans for standardizing the test and evaluating its effectiveness were also prepared.

The test outline and the standardization and research plans prepared by the staff were then reviewed and approved, with slight modifications, by the Educational Testing Service Standing Committees on Research and on Tests and Measurements, as well as by representatives of the research testing agencies of the Departments of the Army, Navy, and Air Force, and the Selective Service Scientific Advisory Committees. In accordance with the approved outline, test questions were constructed, assembled into trial forms, and pretested on large groups of young men not eligible for deferment. The items chosen for final use were drawn from among many others that had been tried out in this fashion. Since four different forms of the test were required, one for each of the testing dates, additional administrations were necessary to establish score equivalences among the various forms, as well as to establish the comparability between SSCQT scores and Army General Classification Test scores. The procurement of suitable groups of examinees constituted a considerable problem, particularly in view of the time limitations. The assistance of the Air Force Human Resources Research Center, the U. S. Merchant Marine Academy, and the U. S. Military Academy in this respect was most valuable. We are also indebted to the Adjutant General's Office of the Department of the Army for making available to us a restricted form of the Army General Classification Test for purposes of equating the SSCQT to the AGCT.

A total of more than 339,000 candidates took the SSCQT during the first series of administrations in the spring and summer of 1951. About half of these took the test on May 26, and most of the remaining candidates were tested during the next two sessions, on June 16 and June 30. Only a small number—a little over 3000—participated in the administration on July 12, the one non-Saturday testing date established for the convenience of students whose religious convictions prevented their taking the test on a Saturday.

The 339,000 students who were tested approximate 21 per cent of the total number of male college students in the country, estimated at 1,569,000 in the fall

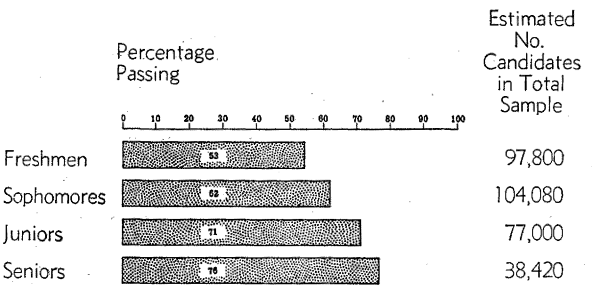


FIG. 1. Percentage of students in each undergraduate year of study achieving a score of 70 or more on the SSCQT. (Based on data for a 10 per cent sample of all candidates tested in the spring and summer of 1951.)

of 1950. More than half the latter total, however, is made up of students who were either ineligible to take the test, or who had no reason to do so. Among these, for example, were some 560,000 veterans. In addition, more than 260,000 of the nonveterans were part-time or nondegree students, and consequently ineligible; also ineligible were an estimated 80,000 nonveterans in the full-time student category who were under 18 or over 26. The total for these three groups comes to more than 900,000; adding in ROTC-deferred students and those classified as 4-Fs brings it to well over 1,100,000, leaving only around 450,000 who were both eligible and not already deferrable on other grounds. Exactly what proportion of the total number of students who might have had something to gain from taking the test actually took it is not possible to determine, but a rough estimate would be between 70 and 80 per cent.

Studies of the test results that have been completed to date are of two types. Those of the first type might be termed demographic, in that they provide information concerning regional differences in test performance, differences among students in various major fields, and the like. These studies have been based on data for a 10 per cent sample of all students to whom the SSCQT was administered, randomly selected within each testing center. Studies of the second type have been concerned with the relationship between test performance and college performance as measured by rank in class, for students at different kinds of institutions and at different stages of their college training. These have been based on data from the May 26 administration for some 5500 students at 23 institutions, selected to represent the kinds of colleges and universities most frequently found in each geographic region.[2] Studies of these two types received high priority because the information they furnished was most important to Selective Service for planning purposes. Other studies of less immediate, but considerable long-range, significance are still in progress and will be reported at some future time.

[2] In the sample of 23 institutions, each major type of college and university was represented in proportion to its number of male graduates in 1947–48; each geographic region was represented roughly in proportion to its male graduates that same year; and the sample was set up by selecting a college or university of the type most often found within a region to represent that region in the sample.

Fig. 1 shows the number of freshmen, sophomores, juniors, and seniors estimated to have taken the test, on the basis of the data in the 10 per cent sample, and the percentage of candidates in each undergraduate year of study who achieved a score of 70 or higher: 53 per cent of the freshmen; 62 per cent of the sophomores; 71 per cent of the juniors; and 76 per cent of the seniors. This increase in percentage equaling or surpassing a score of 70 on the SSCQT seems in large part attributable to progressive drop-outs of less able students from the freshmen through the senior year, although some of it may be associated with increased mental maturity. It should be noted that the passing score for seniors was not 70 but 75. Half the seniors in the 10 per cent sample scored at or above 75, which was the critical score necessary for deferment to enter graduate study.

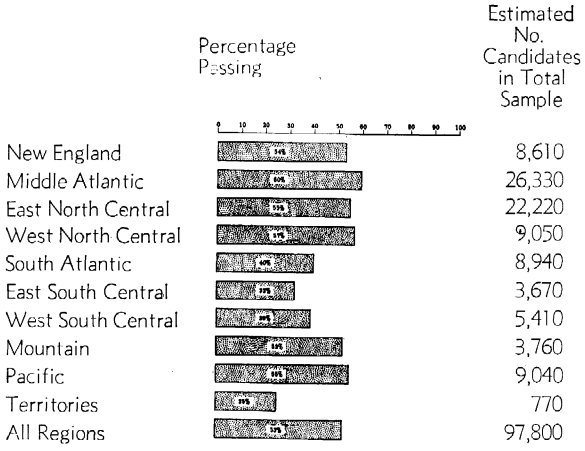


FIG. 2. Percentage of freshmen in each region achieving a score of 70 or more. (Based on data for a 10 per cent sample of all candidates tested in the spring and summer of 1951.)

The differences in performance among students residing in the various geographic regions[3] of the country are rather interesting. Fig. 2 gives the percentage of freshmen in each region who equaled or exceeded a score of 70 on the SSCQT. Despite minor shifts in relative rank order of the various regions, the regional differences in freshman performance are similar to those for sophomores, juniors, and seniors (not shown). In all four comparisons, the percentage of candidates from the New England, Middle Atlantic, East North Central, West North Central, and Pacific regions who passed the test was somewhat higher than that for the country as a whole, whereas the percentage of candidates from the South Atlantic, East South Central, and West South Central regions, and from the territories, who passed the test was appreciably lower than that for the country as a whole. For candidates residing in the Mountain region, the percentage passing was equal to or very slightly below the national average for each year of study.

[3] Geographic divisions utilized were those used by the U. S. Bureau of the Census: *New England*—Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut; *Middle Atlantic*—New York, New Jersey, Pennsylvania; *East North Central*—Ohio, Indiana, Illinois, Michigan, Wisconsin; *West North Central*—Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, Kansas; *South Atlantic*—Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida; *East South Central*—Kentucky, Tennessee, Alabama, Mississippi; *West South Central*—Arkansas, Louisiana, Oklahoma, Texas; *Mountain*—Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada; *Pacific*—Washington, Oregon, California; *Territories*—Alaska, Hawaii, Puerto Rico.

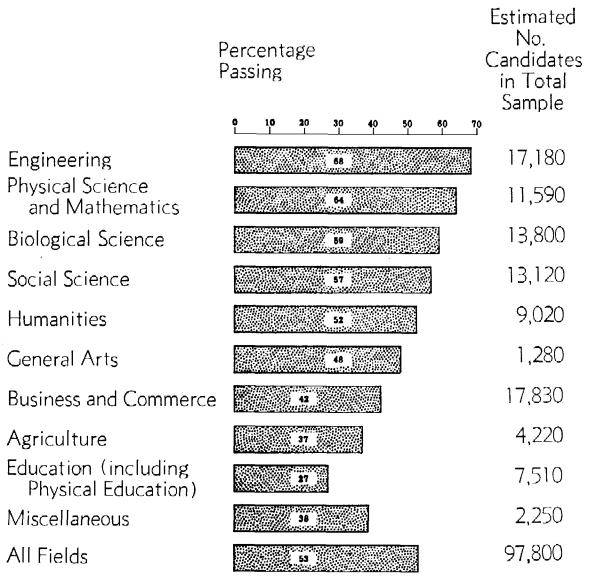


FIG. 3. Percentage of freshmen in each major field of study who achieved a score of 70 or more. (Based on data for a 10 per cent sample of all candidates tested in the spring and summer of 1951.)

Fig. 3 gives data on freshman performance, broken down according to major field of study. The results for sophomores, juniors, and seniors (not shown) are, again, surprisingly similar to those for freshmen. In engineering and in physical science and mathematics, the percentage of candidates passing the test is well above average throughout. Social science is also consistently above average, but not to the same extent. In business and commerce, agriculture, and education (which includes physical education), the percentage passing is well below average in all four comparisons. "Miscellaneous," covering a variety of specialties from carpentry, textile design, and power laundry, through aeronautical administration, law enforcement, and speech pathology, also makes a uniformly poor showing.

The percentage of humanities candidates passing the test remains at or slightly below average in the four comparisons. Like humanities, biological science and general arts appear to be "middle-bracket" fields, although they tend to shift their relative position rather systematically over the four years of study. Biological science, which ranks above social science for freshmen, drops to a position only slightly above that of business and commerce in the senior comparisons. General arts, which ranks below humanities and biological science for freshmen, moves up above both of these to a position just below social science for seniors.

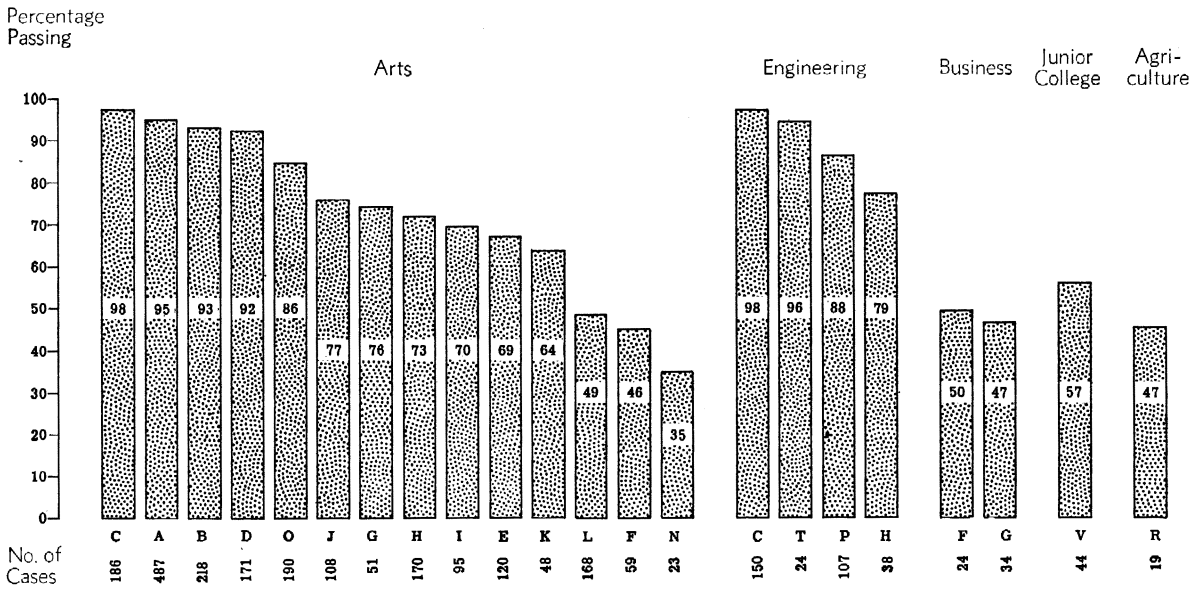


FIG. 4. Percentage of freshmen at various institutions achieving a score of 70 or more. (Based on data for sample of 23 institutions.)

In interpreting the differences among students in various major fields of study with respect to performance on the SSCQT, it should be kept in mind that the test gave equal weight to the linguistic and quantitative aspects of scholastic aptitude. The relative rank order of the several fields might well have changed somewhat if the verbal and numerical aspects of scholastic aptitude had been weighted differently.

The results presented to this point have been based on data for the 10 per cent sample. Those now to be summarized are based on data concerning class standing and test performance for 69 homogeneous groups of students at the 23 selected colleges and universities—i.e., arts freshmen at University A, engineering juniors at University Q, etc. Data for a total of 5527 students—2526 freshmen, 1659 sophomores, and 1342 juniors—were utilized in this analysis. Seniors were not studied.

In advance of the test administration, it was thought that students who met the requirements for deferment on the basis of their rank in class (i.e., the upper half of freshmen, upper two thirds of sophomores, and upper three fourths of juniors) would be less likely to take the SSCQT than students whose rank in class was too low for deferment on that basis, and whose sole opportunity to qualify consequently rested on their test performance. That is not what happened, however. In each year of study, relatively more of those in the upper than in the lower portions of their respective classes took the test. Of the freshmen in the selected samples, 59 per cent were from the top half of their classes, and 41 per cent from the bottom half. Of the sophomores, 73 per cent were in the top two thirds; 27 per cent in the bottom third. Again, 82 per cent of the juniors in the samples studied were in the top three fourths of their respective classes, and 18 per cent in the bottom fourth. The variability among institutions, however, with respect to the percentage of students from the upper and lower portions of their classes who participated in the test, was considerable. Since 23 institutions comprise a relatively small sample, it has seemed best, in arriving at estimates presented later in this paper, to use the theoretical values of 50 per cent of freshmen from the lower portion of their class participating in the test, 33⅓ per cent of sophomores, and 25 per cent of juniors. The obtained percentages, in any event, do not differ significantly from the theoretical ones.

One finding of particular importance emerging from the analysis of data for the 23 institutions is the tremendous variability among these colleges and universities in the percentage of students achieving the critical score on the SSCQT. This finding is in accord with the results of other studies, which have revealed wide variability in test performance among institutions, both throughout the country as a whole and within each geographic region. Fig. 4 shows the percentage of freshmen passing the test in each group consisting of 19 cases or more. At the highest ranking institution, C, 98 per cent of both liberal arts and engineering freshmen passed the test; at the lowest, N, only 35 per cent of the arts freshmen passed. The data for sophomore and junior groups at the various institutions (not presented here) are similar to those for freshmen, although they reveal somewhat less variability. Sophomore and junior groups from institutions that ranked high in the freshmen comparisons continue to stand relatively high; those from institutions that ranked low in the freshmen comparisons continue to stand relatively low.

The institutional differences in performance on the

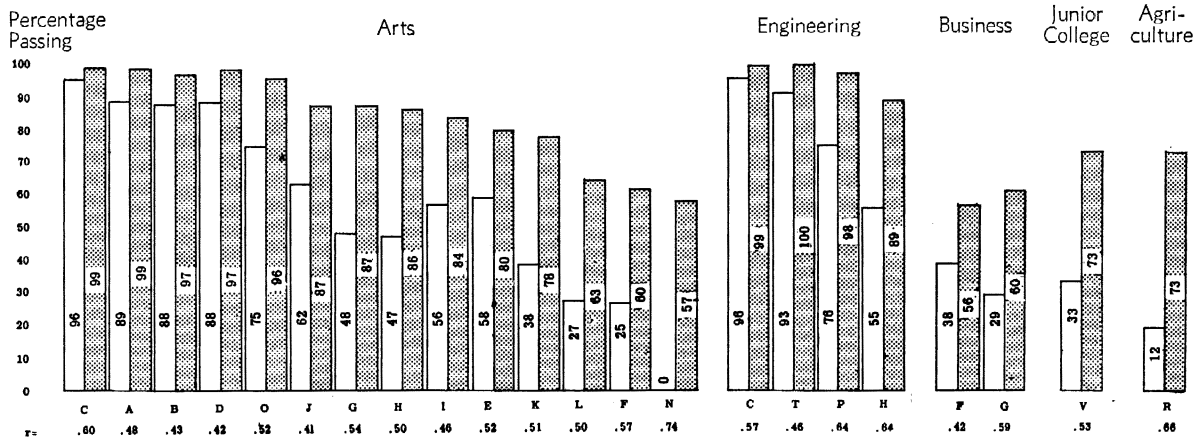


FIG. 5. Percentage of freshmen in upper and lower portions of their respective classes who achieved a score of 70 or more. (Based on data for sample of 23 selected institutions.)

SSCQT are considerably larger, and more consistent, than might be expected on grounds of sampling fluctuation. They reflect marked differences among these colleges and universities in the general level of academic ability of their student bodies. The differences in general level of performance are, in turn, reflected in the number of students from both the upper and the lower portions of their classes at these institutions who passed the test. Fig. 5 gives these data for the same freshman groups that were shown in Fig. 4. As Fig. 5 reveals, more of the high-standing than of the low-standing students in each group achieved the passing score, even though at certain institutions the number who failed comprised only a small fraction of the total. Differences in the percentage of students from the lower portions of their class who passed the test are especially marked. At Institution C, for example, 96 per cent of both arts and engineering freshmen from the lower half of their class passed the test; at Institution N none of the freshmen from the lower half of their class passed.

Values along the bottom row of Fig. 5 are the correlations between test score and rank in class for each group. Despite the wide differences among colleges and universities in caliber of student body, the variation in the correlation coefficients (ranging from .41 to .74) is no greater than would be expected on the basis of sampling fluctuation. In other words, although the passing score set for the SSCQT eliminates widely varying percentages of students at the different institutions, the test itself appears to be as equally good a predictor of freshman grades at one institution as at another.

The same is true for the various sophomore and junior groups studied. The test predicts sophomore and junior performances at any one college or university about as well, within sampling error limits, as at any other. For students at a given stage of training, then, the SSCQT appears to be an equally good predictor of rank in class, regardless of the institution involved. There are, however, differences in the degree to which it can predict performance at different stages of training. Based on data for the 23 institutions studied, it would appear that scores on the SSCQT are somewhat more closely related to performance in the freshman year than in the sophomore or junior year. This is also true of the Scholastic Aptitude Test of the College Entrance Examination Board and of the American Council on Education Psychological Examination. Since data for these two examinations were obtained for the same groups used to study the relationship between the SSCQT and college standing, pertinent figures are available for all three tests. These are summarized in Table 1.

As Table 1 shows, there were six groups of freshmen who took both the SSCQT and the SAT. In these six groups, the average correlation found between SSCQT and rank in class was .52, and the average correlation found between SAT and rank in class was .53. There were thirteen groups of freshmen who took

TABLE 1

CORRELATION SUMMARY

| | No. of groups | No. of cases | Average correlation* of rank in class with: Selective Service College Qualification Test | Scholastic Aptitude Test of the CEEB | ACE Psychological Examination |
|---|---|---|---|---|---|
| Freshmen | 6 | 1321 | .52 | .53 | — |
| | 13 | 1001 | .53 | — | .41 |
| Sophomores | 4 | 609 | .47 | .52 | — |
| | 7 | 641 | .48 | — | .34 |
| Juniors | 2 | 562 | .37 | .41 | — |
| | 5 | 357 | .41 | — | .33 |

* Weighted mean $z$-score transformation.

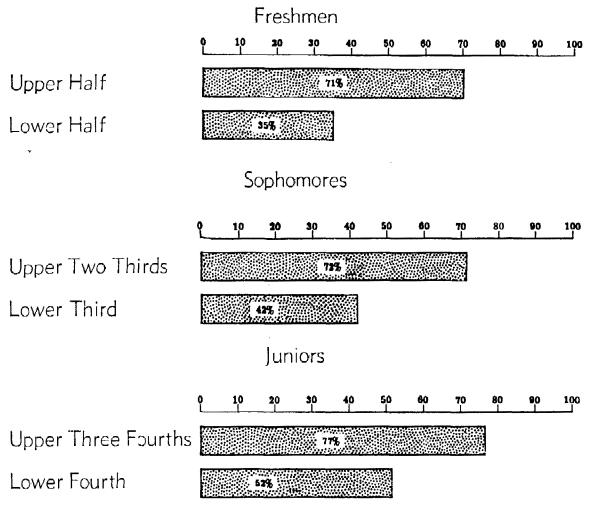


FIG. 6. Percentage of freshmen, sophomores, and juniors in the upper and the lower portions of their respective classes who passed the test. (Based on data for sample of 23 selected institutions, adjusted to conform with known characteristics of the nationwide group.)

both the SSCQT and the ACEPE. In these thirteen groups the average correlation found between SSCQT and rank in class was .53, whereas the average correlation found between ACEPE and rank in class was .41. Analogous data are given for sophomores and juniors. It may be seen from Table 1 that for all three tests the correlations with rank in class tend to decrease somewhat from the freshman to the junior year.

These findings are probably due in considerable part to the increase in the number of elective courses at the more advanced stages of training. Once the freshman year is over, and increasingly so thereafter, the students are no longer running the same race. The grading in some of the elective courses in which few students are enrolled also tends to be somewhat unreliable. In addition, as the less able students are weeded out along the way, difference in ability becomes smaller, which also tends to reduce the correlations.

From a policy point of view, the most important aspect of the relationship between the two measures —SSCQT score and rank in class—is the manner in which they operate together as bases for deferment. Although test performance data were, of course, available for all students participating in the spring and summer administration of the SSCQT, data on rank in class were available only for the students in the 23 selected institutions. Figures showing the joint operation of the two bases for deferment in the country as a whole are, consequently, estimates obtained by adjustment of data for the 23 institutions to conform with known characteristics of the nationwide group with respect to performance on the test. Fig. 6 gives the percentage of students in the upper and the lower portions of their respective classes estimated to have passed the test in the country as a whole, utilizing the theoretical values of 50 per cent of freshmen from the lower portion of their class participating in the test, 33⅓ per cent of sophomores, and 25 per cent of juniors. As revealed in Fig. 6, the estimates are: for the freshmen, 71 per cent in the top half of their class and 35 per cent in the bottom half; for the sophomores, 72 per cent in the top two thirds and 42 per cent in the bottom third; for the juniors, 77 per cent in the top three fourths and 52 per cent in the bottom fourth.

Of obvious importance also, from a policy standpoint, is the joint operation of the two bases for deferment within each region and within each major field of study. By utilizing appropriate combinations of data in Fig. 6 with those presented earlier, and making certain reasonable assumptions, it is possible to derive this information. Fig. 7 gives the estimated number of freshmen in each region who could qualify for deferment on the combined basis of test score and rank in class. The regional differences shown in Fig. 7

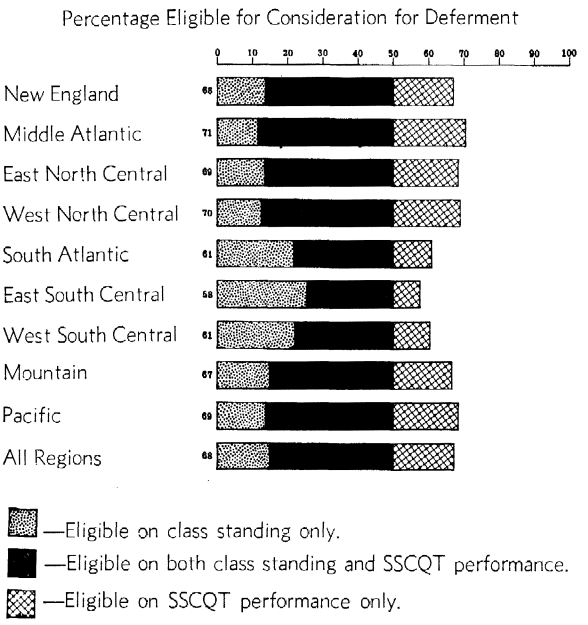


FIG. 7. Percentage of freshmen in each region eligible for consideration for deferment on combined basis of test score and class standing. (Based on combination of data for the 10 per cent sample and the 23 selected institutions.)

are smaller than those shown in Fig. 2, but nevertheless reflect significant variation in the general level of scholastic ability of students in the various regions. Estimates for sophomores and juniors (not shown) follow a similar pattern, although the differences among regions are smaller for sophomores than for freshmen, and still smaller for juniors, because of the increase in percentage of students eligible for consideration for deferment on the basis of rank in class.

Similarly, if we assume that so far as grades are concerned there is no difference among the various major fields, it is possible to estimate the number of students in each field of study who could qualify for deferment on the combined basis of test score and

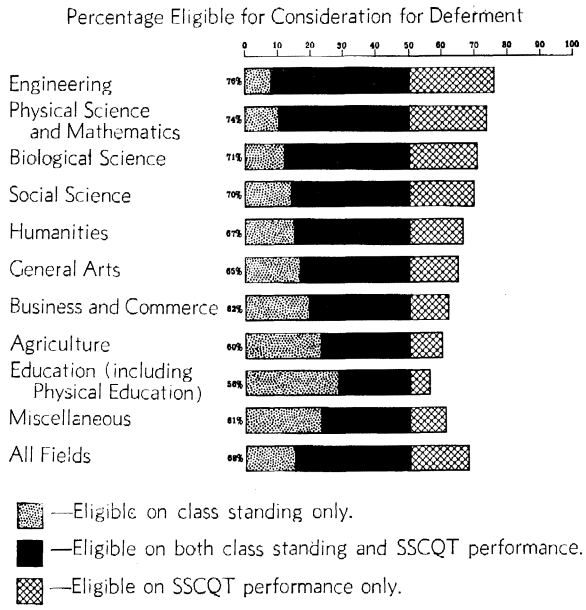


FIG. 8. Percentage of freshmen in each major field of study eligible for consideration for deferment on combined basis of test score and class standing. (Based on combination of data for the 10 per cent sample and the 23 selected institutions.)

rank in class. Fig. 8 gives these estimates for freshmen. It is evident that the high-ranking fields mentioned earlier, engineering and physical science and mathematics, still maintain considerable relative advantage, although the margin has decreased. Education and agriculture maintain their relatively poor status, but again by a smaller margin. Although the differences among the major fields are smaller for sophomores and juniors than for freshmen, the estimates for these two classes (not shown) follow a similar pattern. (In this discussion of the joint operation of the two bases for deferment, within region and within major field of study, I am sure it is understood that it is not mandatory for the local boards to follow the class standing or test score criterion for deferment. At the same time, the individual or his institution has the right of appeal in the event it is felt that the student has been unfairly classified.)

The results that have emerged from the various analyses are clearly in conformity with the rationale on which the original decision to include a test of scholastic aptitude in the draft deferment program was based. As the data presented here reveal, there are tremendous variations among institutions with respect to the general caliber of their students. A student who stands relatively low at one institution may compare very favorably, both in ability and in level of academic performance, with a student who stands relatively high at another institution. Without the correction factor introduced by the test results, individuals from institutions that maintain high standards in the selection of their student body would be unfairly penalized in comparison with those from institutions whose standards are less rigorous.

Within any given institution, also, it would appear from the results given here that the test score supplies information of relevance and importance beyond that afforded by rank in class alone. As indicated earlier, college grades are influenced by numerous factors in addition to the student's level of ability. Differences among students in the level of difficulty of courses chosen, in opportunity for study as affected by demands of self-support and extra-curricular activities, and in value attached to the importance of grades per se are all reflected in college grades, as are differences among instructors in grading standards and in sheer ability to evaluate quality of student performance. Although the test scores, too, are subject to a certain amount of measurement error, they are almost entirely free from the other nonrelevant influences that enter into grades, and therefore provide valuable supplementary data for the purpose at hand.

The findings available to date confirm that the SSCQT effectively serves the purpose of giving capable students everywhere an opportunity to qualify for deferment to continue their training. Further, they augur well for the future operation of the program in still another way. The particular critical scores in use at the present time were selected because they seemed best to meet existing needs. From the studies that have been made, it is apparent that the two measures—rank in class and test score—function with joint effectiveness throughout a wide range. The findings indicate that, if manpower requirements change, appropriate changes in the critical scores can be made without injuring the joint efficiency of the two measures.

In closing, I should like to return to the most important point of all. Students selected for deferment under the college qualification program are permitted to continue their training, not for their own benefit, and not for the benefit of the institutions they attend, but for the benefit of the country as a whole. They are deferred at this time in order that they may be able to serve the military and the country as a whole more effectively later on, after their training has been completed. The basic aim of the program is to select those men whose potential contributions as specialists are of sufficient significance to warrant postponement of their service.

Nevertheless, it is vitally important that a program of this nature be fair to all. Although the draft deferment program was undertaken for the benefit of the nation, and not for that of the individual or the institution, its operating provisions must ensure just treatment for both the individual and the institution. From the analyses that have been made, it would seem that the plan developed by Selective Service with the counsel of the advisory committees appointed by General Hershey was succeeding in this as in the other objectives which it was designed to meet.