SCIENCE

Vol. 104, No. 2694 Friday, 16 August 1946

Inequalities in Adult Capacity—From Military Data

Walter V. Bingham

Chief Psychologist, The Adjutant General's Office, War Department

TEN MILLION MEN HAVE TAKEN the Army General Classification Test. This is the most widely used of many tests of ability, aptitudes, trade proficiency, and technical knowledge developed by the Army as aids in classifying its personnel. During tense years of rapid mobilization and campaigning, this test has served a specific military purpose, namely, to indicate the rate at which each individual soldier might be expected to absorb training and the level of difficulty of the material he should be capable of mastering.

After describing this classification test and the inequalities of adult capacity which it measures, the frequencies of occurrence of these differences will be brought into relationship with certain other variables, educational and occupational. These data throw light not only on military problems; they illuminate also our conception of the American people, reminding us of our failure to conserve and capitalize to the full our national endowment of human resources.

The AGCT

The Army General Classification Test—AGCT for short—is not an IQ Test, if the words "Intelligence Quotient" are correctly used. An IQ is a ratio. It expresses the *rate* at which a child's mentality has developed. The ratio of his attained level of intelligence to his age is called his intelligence quotient. During infancy, childhood, and early youth this quotient not only tells the rate of development which has already taken place but also indicates in a general way, although not very precisely in individual cases, the probable rate of subsequent development and the level likely to be reached at maturity.

When a recruit joins up, it is not difficult to ascertain how intelligent he is; but it is too late to ascertain his intelligence *quotient,* because the information is not available as to whether he had reached his mental maturity at age 18, 17, 16, 15, or 14. This is one reason why the Adjutant General's Office in 1941 emphasized the inappropriateness of expressing adult capacity in IQ terms (*1*). Another reason, essentially practical, is that a personnel classification officer is not interested in knowing what a soldier's rate of development *has been.* Instead, the question is: How intelligent is he now? What can he learn and how fast can he learn it?

Neither is it a matter of practical concern to know what a soldier's *native* intelligence was at birth, before his mental development had been facilitated in any degree by stimulating surroundings or hampered by a stultifying environment. The assignment officer wants an index of what the new soldier can be expected to learn, rather than a figure which purports to tell what he might have been able to learn if only he had had a better home, no enfeebling illness, and a great deal more education.

What a soldier can learn to do is a reflection not only of his original endowment. It is conditioned also by what he learned while growing up—the mental furniture he accumulated. His adult capacity to learn the complicated arts of war is determined in part by the educational opportunities he has had and the way in which he has grasped those opportunities, and only in part by the native capacity with which he was born. It is not our intention to attempt the baffling task of disentangling the interwoven threads of heredity and environment; of family stock, ancestry, and race; of age, schooling, and experience; and of motivation as a contributor to achievement. Those are puzzles for the geneticist and the sociologist; but the facts reported here may furnish some grist for their mills.

The AGCT exists in several forms. Four alternate forms, known as 1a, 1b, 1c, and 1d, have been used in examining troops who are literate in the sense that they can read English. Two forms in Spanish are used in the Caribbean Command. A nonverbal test, known as 2abc, was developed for use with illiterates and troops literate only in some language other than English, such as Chinese. This can be given in pantomime, without spoken directions. A relatively new

Presented at the annual meeting of the National Academy of Sciences, 22 April 1946.

and longer examination, known as Form 3a, currently taken by all recruits and inductees at Reception Centers soon after they have been accepted and sworn into military service, yields in addition to the general classification score a profile of four other useful scores.

Forms H–1 and H–2, focused on the higher levels of ability, were designed for use in screening applicants for training in Officer Candidate Schools, while R–1 and R–2, more discriminating in the lower part of the scale, have been used in Recruiting and Induction Stations to identify the least promising applicants for enlistment. Mention should be made also of the Army Individual Test, an instrument of wide applicability where conditions call for a careful recheck of a soldier's standing.

Most of the data here presented were secured with one of the four forms first mentioned. Each of these poses three kinds of tasks: verbal, arithmetical, and spatial. The verbal items, increasing in difficulty, sample a person's grasp of precise meanings. The arithmetical items require numerical computation and solution of problems smoothly graded in difficulty from easy to hard. The spatial items test ability to visualize and to think about the relationships of things in space. All items are of the multiple-choice sort, requiring no writing of answers but only the selective marking of answer sheets which can then be evaluated objectively, accurately, and rapidly with an ordinary scoring stencil or with an electrical test-scoring machine. The time allowance, following a preliminary period of instruction and fore-exercise to make certain that the directions are understood, is 40 minutes, which is long enough for most men to reach their upper limit of difficulty. The reliability of these forms lies between .94 and .96, as estimated by the conservative Kuder-Richardson formula (*4*).

The Standard-Score Scale

Scores are expressed in terms of a standard scale on which 100 represents the midpoint of the distribution and 20 is the standard deviation (the root-mean-square deviation) from this central tendency. The extreme range of standard scores is from 39 to 163. The middle zone, called Army Grade III, includes all scores from 90 up to but not including 110—that is, within the range ± 0.5 standard deviation from the mean. The next higher zone, called Army Grade II, extends from 110 to 130. All scores of 130 or higher are in Army Grade I. Similarly, below the middle zone are Army Grades IV and V[1] as seen in Fig. 1.

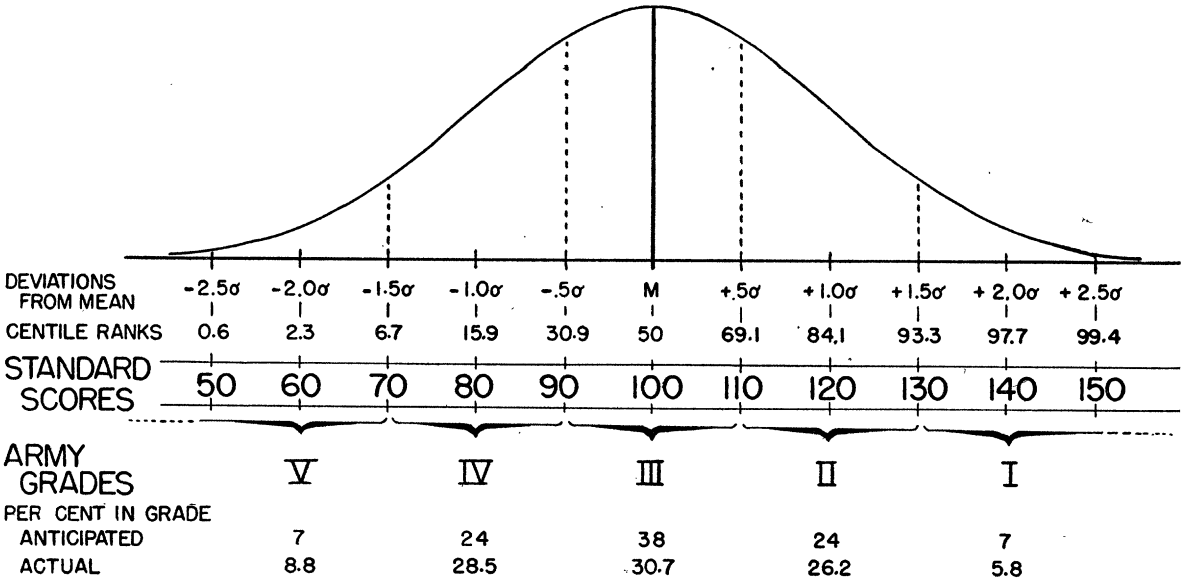

FIG. 1. The Army's mental footrule, showing standard scale, Army Grades, anticipated proportions in each grade, and actual distribution of scores on the Army General Classification Test among approximately 10,000,000 men examined after acceptance and induction into the service.

On a scale so defined, a normal distribution of scores would have been bell-shaped, symmetrical with respect to the midpoint, as follows: in Grade V, the lowest, 7 per cent of the population; in Grade IV, 24 per cent; in Grade III, 38 per cent; in Grade II, 24 per cent; and in Grade I, 7 per cent. The actual distribution was somewhat flattened and skewed by peculiarities of population sampling, such as those introduced by policies of Selective Service regarding deferments. Then, too, the original scaling and calibration of the

[1] The dividing point between Grades IV and V, originally set at standard score 70, was changed in 1943 to 60, for administrative reasons.

test in 1940 had been accomplished during the earliest stage of mobilization, under pressure for haste, with data from less than 3,000 subjects, a sample neither large enough nor sufficiently representative of the total population, rural and urban, East and West, North and South. In this sample were too few of the very stupid and the very brilliant, of the wholly uneducated and the highly educated. Under the circumstances it is not surprising that the anticipated distribution of scores has been approximated only roughly. The proportions found are: in Grade V, 8.8 per cent; IV, 28.5 per cent; III, 30.7 per cent; II, 26.2 per cent; and I, 5.8 per cent.

The Population Sample

The population of approximately 10,000,000 on whose performance these actual proportions are computed includes all enlisted men taken into the Army during a five-year period beginning on 1 March 1941. This is an enormous sample, but it differs from the entire national population of the same age and sex in important ways. All who took this test had met minimum standards for acceptance into the military service. The physically or mentally disqualified had already been screened out. Missing, too, are those registrants who were not called for training because of dependents, civilian occupation, or other reason; while only those officers are represented in the tabulation who had first served as enlisted men and earned their commissions in Officer Candidate Schools or, in a few instances, on the battlefield. To be sure, graduates of schools for training officer candidates, warrant officers, and aviation cadets constituted more than 68 per cent of the officer population. Of the remaining 32 per cent (the officers who did not take the AGCT), three-fifths were National Guard, Reserve, and Regular Army Officers and the remainder were doctors, dentists, chaplains, and other specialists commissioned in the Army of the United States directly from civil life. It is not unwarranted to assume that if all these officers had been included in the population tested, the proportion scoring in the top bracket would have been a little larger, the curve of distribution less skewed to the right.

The proportion of the Army population in each of the Army grades varied considerably from one year to another. Changes in minimum standards of acceptability had their effect, as did nationwide policies regarding deferments and to some extent, perhaps, the attitudes of local boards.

In 1940 and until 17 April 1941 every recruit inducted within the continental limits was supposed to be able to understand simple orders in English and to read and write in English or in some other language; but Army regulations did not prescribe very precisely the required degree of literacy or the procedures by which it was to be ascertained. During this period, however, it was established that soldiers in training who could not read English about as well as the average man who had completed four years of school were having so much difficulty in learning that they were holding up the progress of their units; so a half-dozen simple paper-and-pencil tests were developed and standardized for use at Recruiting and Induction Stations whenever the interviewers were in doubt as to a man's ability to read at what came to be called the fourth-grade level.

During 1942 available man power became less abundant. In particular, draft boards in certain communities where the population was largely colored found it hard to meet their quotas of literate inductees. So, beginning on 1 August, acceptance of a certain proportion of illiterates was permitted. The rule was that on each day not more than 10 per cent of the white men and 10 per cent of the colored men processed at an Induction Station might be illiterate provided they were physically fit and were able to pass a specified nonverbal mental test. These regulations remained in effect until 1 February 1943, when the maximum proportion of illiterates was cut to 5 per cent of each day's intake. Then, after 1 June 1943 illiteracy as such was no longer a bar to induction. Examiners at Recruiting and Induction Stations were furnished a new set of screens—a battery of tests designed to admit to military training all bright and physically fit men even though they did not know how to read, while excluding the very dull even though they were literate. At the same time the 93 Induction Stations were provided with qualified examiners familiar with the new techniques, 140 psychologists having been commissioned and trained in the Army's procedures of examining. After further experience and experimentation a revised battery of induction station tests was introduced in June 1944, and these standards and procedures prevailed until after V-J Day when, on 23 September 1945, the induction of illiterates was discontinued. These changes had their inevitable effect on the proportions of men in the service scoring in the different Army grades.

Parallel with alterations of minimum standards the Army introduced significant changes in the organization of training. At first the assimilation of Mexicans, American Indians, Chinese, Puerto Ricans, and others who knew little or no English even though educated in their own language had posed a serious problem in the training centers. To meet this situation Special Training Units were established where unique procedures were developed for teaching such men rapidly. These units provided 12 weeks of well-conceived and effectively supervised instruction in reading, language

expression, arithmetic, and the most elementary military subject matter. Progress was astonishing. With such a head start, most of these men were then able to pick up the ordinary routines of basic training and carry on. So, in June 1943, Special Training Units of this sort were established in each Reception Center. Here, immediately on acceptance into the service, were concentrated the illiterates, the non-English-speaking literates, and all who scored in Army Grade V on the AGCT. After 12 to 16 weeks of this introductory training nearly 85 per cent of this otherwise unusable personnel was salvaged. Many whose test performance at entrance was in Army Grade V were able to raise their standing to Army Grade IV or even higher. Such are some of the facts that account for the changes observed from year to year in the distribution of AGCT scores.

Validity

The AGCT is valid in the military situation. It tells, at least approximately and in a usable way, what it purports to tell about a man's capacity to learn. For example, in the training camps it was found that the population scoring in the upper ranges of the test, namely, in Army Grades I and II, furnished a large majority of the pace-setters and of those who were able to master the more advanced military specialties. More than half of the men scoring in these grades became skilled technicians, noncommissioned officers, or specialists; and practically all of the hundreds of thousands of enlisted men who eventually went to Officer Candidate Schools and earned commissions were drawn from this upper range. There were exceptions. For example, although a standard score of at least 110 was ordinarily prerequisite for assignment to an Officer Candidate School, a War Department regulation permitted enlisted men to attend if they had had two years of Junior ROTC training. Thirty-five such soldiers taking officer candidate training at Fort Benning were found to have scored below 110 on the AGCT, and of these, three passed the course and received their commissions. The other 32 failed. Instances also occurred in which commands were required to fill quotas of Officer Candidate School candidates even though they did not have men enough who could meet the specifications. Most of these failed the course, either for poor scholarship or lack of leadership.

The middle zone, Army Grade III, furnished a substantial number of corporals, sergeants, clerks, and technicians and a great many capable privates.

In Army Grade IV also were large numbers of men who absorbed basic training thoroughly in the short time available and then could be relied upon to exhibit the resourcefulness and skills which warfare now demands of every individual fighter in the ranks. Some in this grade became noncoms and specialists; and at the other extreme some qualified only as unskilled laborers.

Of those who scored in Army Grade V, about one-third demonstrated capacity to learn the art of combat. For the others there were many kinds of useful work in training camps, ports of embarkation, and transport services, but some of the men proved to be more of a burden than a help during mobilization.

On each soldier's Qualification Card were entered the essential data about his age, physique, occupational experience, schooling, hobbies and sports, familiarity with foreign languages, previous military experience, if any, and his score in the AGCT. Here, too, were recorded his standing in any special trade tests, aptitude tests, or achievement tests the Army examiners found it expedient and feasible to use as supplements to the data obtained by interview. Such personal data were invaluable when selecting men for unusual duties or for training in a military specialty (*2*).

The utility of these data was most vividly in evidence whenever a new division was being activated and an arriving trainload of novices or partly trained troops had to be distributed quickly to the various regiments, companies, batteries, and service units in such a way that each unit would have its due proportion of the available skills and abilities.

AGCT data helped in selecting the men who were to take special training (*4*). Three examples illustrate the degree of validity of predictions from such scores. Among 1,042 men in training as weather observers, to cite one instance, it was found that the chances that a soldier would do average or better in the course were only 3 in 100 if his AGCT standard score was 80; 12 in 100 if his standard score was 100; 35 in 100 if his standard score was 120; and 65 in 100 if his standard score was 140.

The chances in 100 that a man receiving a certain standard score in the AGCT will achieve average or better in a course of training for general clerical work were ascertained to be as follows: 1, if his score is 60; 5, if it is 80; 20, if it is 100; 47, if it is 120; and 76, if it is 140.

These findings surprise no one. Less generally recognized is the validity of such a test as a predictor of average or better success in learning to be an airplane and engine mechanic:

Standard score—	60	80	100	120	140
Chances in 100—	5	17	40	67	88

This prediction is almost as good as that from a paper-and-pencil test of mechanical aptitude given to the same soldier-learners. But neither of these tests came up to the validity, for this purpose, of a Tech-

nical Trade Test consisting of the best items from a general technical test and a trade information test. Most striking was the relative futility, in this situation, of all the many performance tests of manual dexterity, mechanical ingenuity, and manipulative skill that were proposed and tried. What has to be acquired in this as in most other difficult mechanical trades is the know-how, the ability to diagnose what is wrong, good judgment in selecting the right tool, and knowledge of how, when, and where to apply it. These indispensable qualifications do not reside in an artisan's hands but in his head. Dexterity, motor control, and manual proficiency are of distinctly secondary importance in mechanical occupations such as this one.

In situations like these the AGCT demonstrated its validity, within reasonable limits, as an indicator of ability to learn a soldier's duties. The inequalities of capacity which it roughly measures range all the way from an intelligence level characteristic of able officers and technicians to the minimal mental equipment required for performing the most simple useful service in the military establishment. This range is wide.

Relation of Test Score to Schooling

The individual differences revealed by this Army test are related positively but not very closely to differences in education and in occupational level. The higher a soldier's educational level, the greater his likelihood of scoring in Army Grades I or II. But note: vast numbers of the men in these top zones of intelligence never went beyond high school. Among those who did not even finish grammar school were about 5,000—.0005 per cent of the total population examined—who scored in Army Grade I.

Most of the soldiers who had graduated from a college scored in Army Grades I or II. Indeed, their average score was close to 130, at the dividing point between Grades I and II. So far as intellectual capacity is concerned it is not unreasonable to assume that nearly all young men in Army Grade I, at least, would be pretty good risks for college and for professional training provided they had had a chance to finish their secondary schooling. Yet, of all the soldiers in Army Grade I, only one-fourth are college graduates.

In Army Grade II the approximate distribution by educational level is as follows: 184,000 are college graduates; 1,666,000 finished high school but not college; 858,000 completed eighth grade but not high school; and 56,000 did not complete eighth grade.

Altogether in Army Grades I and II there were about 2,000,000 men who had finished high school but were not college graduates. This is one-fifth of the entire population examined. Here is a vast acreage of human capacity. Has it been cultivated as intensively as it should?

Some of these men actually may have obtained a better education off campus than they would have obtained in a college and gone farther up the ladder of subsequent achievement. However, the question that remains is: Are there not several hundred thousand who had the requisite capabilities but did not accomplish what they might have, in peace and in war, because their education was truncated?

Test Score and Occupation

The fact of America's prodigality with her human resources comes into focus again when we examine the distribution of AGCT scores in each of several fields of occupation.

Looking at the average test performance within each occupational group, it is obvious that the occupations arrange themselves in a sort of hierarchy. At or near the top are the learned professions and such occupations as chemist, accountant, auditor, and business manager; while laborer and farmhand are close to the other extreme. This is what anyone would expect. The great variability or spread of scores within each occupation is not as commonly appreciated nor is the amount of overlap when any two occupations are compared. In one study, for instance, of the men whose civilian occupation was driving a heavy truck, the upper one-fourth outranked in test score the lower one-fourth of the managers and officials. Among the professions the law ranked close to the top; but 9 per cent of the boilermakers did as well as the average lawyer.

So it is with several hundred callings which furnished man power to the Army. In a majority of them were men who, given the will and the opportunity, might have prepared themselves to follow successfully an occupation or profession more valued than the one they chose or into which they drifted or were forced by circumstance.

Of immediate concern to the National Academy of Sciences is the current shortage of young scientists. The supply of physicists, mathematicians, biochemists, psychologists, and specialists in other branches of natural science is distressingly insufficient. Most conspicuous is the deficit of scientists capable of filling top positions in research and in university departmental administration. Similar shortages are evident in such applied sciences as electronics, mechanical engineering, pharmacology, and scientific agriculture. These occupations compete for talent with other areas of specialization: with the social sciences and the humanities; with the fields of literature, history, art, education, social work, politics, public administration, and business management; as well as with the tradi-

tionally learned professions of law, medicine, and the ministry. In a confused and shaken world, the national defense also must absorb a larger fraction of the country's brains than were concentrated in the military profession during the 1920's and 1930's.

For these reasons America must no longer be prodigal of the capacities of her more promising youth.

Conclusion

We have seen that the AGCT measures a wide range of inequalities in adult capacity. With this yardstick, data have been collected from a massive population for a practical military purpose. These data, when brought into relation with facts about schooling and civilian occupation, reveal a vast pool of talent now only partly drawn upon. Large numbers of men have been identified whose levels of accomplishment, in both education and vocation, have fallen short of their potentialities.

These facts are a challenge to conserve the national heritage. This consists not only of fertile soil, mineral wealth, oil, water power, forests, and wildlife, but also of that most precious of resources, the intellectual capacities of our young people. Does it not follow that the number of scholarships available to really superior applicants for higher education should be greatly increased? Gates of opportunity in colleges, technical institutions, and graduate schools should be opened more widely or—wherever such facilities are overcrowded—more discriminatingly. To select the most promising from among multitudes is no longer a baffling task. It has been done. It can be done widely.

Experience of the military services during mobilization and war dictates an expansion of civilian facilities for intervewing, testing, record keeping, and counseling of young men and women. Staffs of qualified specialists such as those a few universities now maintain should be available in every educational institution to identify the most promising students and to facilitate their advancement. Incentives must be widely and plentifully supplied, and new ambitions stirred in any and all who aim at vocational targets less worthy than the best of which they are capable.

References

1. Bingham, W. V., *et al.* *Science,* 1941, **93**, 572–574.
2. Bingham, W. V. *Ann. Amer. Acad. polit. soc. Sci.,* 1942, **219**, 18–28.
3. Bingham, W. V. *Science,* 1944, **100**, 275–280.
4. Staff, Personnel Research Section, Adjutant General's Office. *Psychol. Bull.,* 1945, **42**, 760–768.
5. ———. *Personnel classification tests.* (War Dept. Technical Manual TM 12–260, rev.) Washington, D. C.: U. S. Government Printing Office, April 1946.

Obituary

Frank Michler Chapman
1864-1945

Frank Michler Chapman was born in Englewood, New Jersey, on 12 June 1864. Following graduation from the Englewood Academy in 1880 he entered the banking profession as an employee of the American Exchange National Bank of New York City, but his growing interest in birds impelled him to resign in 1886 in order to devote his life to ornithology.

He spent a little over a year as an independent student, during part of which he worked as a volunteer assistant to J. A. Allen, then head of the Department of Mammals and Birds in the American Museum of Natural History. Shortly afterward, he was offered a permanent position in the Museum and on 1 March 1888 began his official duties as assistant in Dr. Allen's Department. He was made assistant curator the same year, associate curator in 1901, and curator in charge of birds in 1908, and, when the Department of Birds was formally separated in 1920, he was appointed its first curator, a position which he held until his retirement in 1942.

The intervening years were busy ones. Chapman's principal interests were always concerned with reaching the public, and the exhibition halls of the Museum gave him wide opportunity to achieve this end. He gave his early attention to the possibilities of the "habitat group," and his first experimental display of the bird life of Cobb's Island, Virginia, proved so successful that he devoted a large part of the following decade to the collection of material for a large series of comparable groups of the birds of other parts of North America. These were followed, in turn, by more elaborate exhibitions of birds of the world, the domes of flying birds surmounting certain other halls, and still other displays.

His pen, meanwhile, was likewise busy. In 1899 he obtained the cooperation of the Audubon Society and commenced the publication of the popular magazine, *Bird-Lore,* which he continued to edit through 1934, when it was taken over by the Society. Through