A nation of morons

The intelligence tests introduced in America for army recruits during the First World War were seen as the way to bring the respectability that psychology yearned for as a new science. But the tests produced some surprising results, and influenced the decision to restrict immigrants in the 1920s.

Stephen Jay Gould  Robert M. Yerkes, about to turn 40, was a frustrated man in 1915. He had been on the faculty of Harvard University since 1902. He was a superb organiser of men, and an eloquent promoter of his profession. Yet psychology still wallowed in its reputation as a "soft" science, if a science at all. Yerkes wished, above all, to establish his profession by proving that it could be as rigorous a science as physics. Along with most of his contemporaries, he equated rigour and science with number and quantification. The most promising source of copious and objective numbers, Yerkes believed, lay in the embryonic field of mental testing.

But mental testing suffered from inadequate support and its own internal contradictions. It was, first of all, practised extensively by poorly trained amateurs whose manifestly absurd results were giving the enterprise a bad name. In 1915, at the annual meeting of the American Psychological Association in Chicago, a critic reported that the mayor of Chicago had scored as a moron on one test. Yerkes joined with critics in discussions at the meeting and proclaimed: "We are building up a science, but we have not yet devised a mechanism which anyone can operate." Furthermore, available scales gave markedly different results even when properly applied. And support had been too inadequate, and coordination too sporadic, to build up a pool of data sufficiently copious and uniform to compel belief.

Wars always generate their retinue of camp followers with ulterior motives. Many are simply scoundrels and profit-seekers, but a few are spurred by higher ideals. As the First World War approached, Yerkes got one of those "big ideas" that propel the history of science: could psychologists possibly persuade the army to test all its recruits? If so, the philosopher's stone of psychology might be constructed: the copious, useful and uniform body of numbers that would fuel a transition from dubious art to respected science. Yerkes proselytised within his profession and within government circles, and he won his point. As Colonel Yerkes, he presided over the administration of mental tests to 1.75 million recruits during the First World War.

Yerkes brought together all the major figures from the hereditarian school of psychology—those who believed that everything important about intelligence and, indeed, behaviour, is inherited and on the whole unaffected by the environment. From May to July 1917 he worked with Lewis Terman, H. H. Goddard and other colleagues at Goddard's Training School in Vineland, New Jersey, writing the army mental tests.

Their scheme included three types of tests. Literate recruits would be given a written examination, called the Army Alpha. Illiterates and men who had failed Alpha would be given a pictorial test, called the Army Beta. Failures in Beta would be recalled for an individual examination. Army psychologists would then grade each man from A to E (with plusses and minuses) and offer suggestions for proper military placement. Yerkes suggested that recruits with a score of C— should be marked as "low average intelligence — ordinary private". Men of grade D are "rarely suited for tasks requiring special skill, forethought, resourcefulness or sustained alertness". D and E men could not be expected "to read and understand written directions".

The Alpha test included eight parts, the Beta seven; each took less than an hour and could be given to large groups. Most of the Alpha parts presented items that have become familiar to generations of test-takers ever since: analogies, filling in the next number in a sequence, unscrambling sentences, and so forth. This similarity is no accident; the Army Alpha was the grandaddy, literally as well as figuratively, of all written mental tests.

These familiar parts are not especially subject to charges of cultural bias, at least no more so than their modern descendants. In a general way, of course, they test literacy, and literacy records education more than inherited intelligence. Moreover, a schoolmaster's claim that he tests children of the same age and school experience, and therefore may be recording some internal biology, did not apply to the army recruits—for they varied greatly in access to education and recorded different amounts of schooling in their scores.

A few of the items in the tests are amusing in the light of Yerkes's assertion that the tests "measure native intellectual ability": Consider the Alpha analogy: "Washington is to Adams as first is to . . ." But one part of each test is simply ludicrous in the light of Yerkes's analysis. How could Yerkes and company attribute the low scores of recent immigrants to innate stupidity when their multiple-choice test consisted entirely of questions like:

Crisco is a: patent medicine, disinfectant, toothpaste, food product

The number of a Kaffir's legs is: 2, 4, 6, 8.
Christy Mathewson is famous as a writer, artist, baseball player, comedian.

Recruits had to be allocated to their appropriate test. Men illiterate in English, either by lack of schooling or foreign birth, should have taken examination Beta, either by direct assignment, or indirectly upon failing Alpha. Yerkes’s corps tried heroically to fulfill this procedure. But standards for the division between Alpha and Beta varied substantially from camp to camp. The problem cut far deeper than simple inconsistency among camps. The persistent logistical difficulties imposed a systematic bias that substantially lowered the mean scores of Blacks and immigrants. For two major reasons, many men took only Alpha and scored either zero or next to nothing, not because they were innately dumb, but because they were illiterate and should have taken Beta by Yerkes’s own protocol. First, recruits and draftees had, on average, spent fewer years in school than Yerkes had anticipated. Lines for Beta began to lengthen and the entire operation threatened to clog at this bottleneck. At many camps, unqualified men were sent in droves to Alpha by artificial lowering of standards. Schooling to the third grade sufficed for Alpha in one camp; in another, anyone who said he could read, at whatever level, took Alpha.

Secondly, and more important, the press of time and the hostility of regular officers often precluded a Beta retest for men who had incorrectly taken Alpha. As the pace became more frantic, the problem worsened. The chief tester at Camp Dix complained “In June it was found impossible to recall a thousand men listed for individual examination. In July Alpha failures among Negroes were not recalled.” The stated protocol scarcely applied to Blacks who, as usual, were treated with less concern and more contempt by everyone. Failure on Beta, for example, should have led to an individual examination. Half the Black recruits scored D— on Beta, but only one-fifth of these were recalled and four-fifths received no further examination. Yet we know that scores for Blacks improved substantially when the protocol was followed. At one camp, only 14.1 per cent of men who had scored D— on Alpha failed to gain a higher grade when retested on the Beta.

The effects of this systematic bias are evident in one of E. G. Boring’s experiments with the summary statistics. Boring, later a famous psychologist himself, but then Yerkes’s lieutenant, culled 4893 cases of men who had failed both Alpha and Beta. Converting their scores to the common scale, he calculated an average mental age of 10.775 for Alpha, and a Beta mean of 12.158. He used only the Beta scores in his summaries; Yerkes’s procedure worked. But what of the myriad who should have taken Beta, but received only Alpha and scored abysmally as a result—primarily poorly educated Blacks and immigrants with an imperfect command of English, the very groups whose low scores caused such a hereditary stir later on?

Academicians often forget how poorly or incompletely the written record, their primary source, may represent experience. Some things have to be seen, touched and tasted. What was it like to be an illiterate Black or foreign recruit, anxious and befuddled at the novel experience of taking an examination, never told why, or what would be made of the results; expulsion, the front lines? In 1968, an examiner recalled his administration of Beta: “It was touching to see the intense effort...put into answering the questions, often by men who never before had held a pencil in their hands.” Yerkes had overlooked, or consciously bypassed something of importance. The Beta examination contained only pictures, numbers and symbols. But it still required pencil work and, on three of its seven parts, a knowledge of numbers and how to write them.

The internal contradictions and a priori prejudice thoroughly invalidated the hereditary conclusions that Yerkes was to draw from the results. Boring himself called these conclusions “preposterous” late in his career. But I had not understood how the Draconian conditions of testing made such a thorough mockery of the claim that recruits could have been in a frame of mind to record anything about their innate abilities. In short, most of the men must have ended up either utterly confused or scared shitless. I believe that the conditions of testing, and the basic character of the examination, make it ludicrous to believe that Beta measured any internal state deserving the label intelligence. Despite the plea for geniality, the examination was conducted in an almost frantic rush. Most parts could not be finished in the time allotted, but recruits were not forewarned.

Still, the tests did have a strong impact in some areas, particularly in screening men for officer training. At the start of the war, the army and national guard maintained 9000 officers. By the end, 200 thousand officers presided, and two-thirds of them had started their careers in training camps where the tests were applied. In some camps, no man scoring below C could be considered for officer training.

But the major impact of Yerkes’s tests did not fall upon the army. Yerkes may not have brought the army its victory, but he certainly won his battle. He now had uniform data on 1.75 million men, and he had devised, in the Alpha and Beta exams, the first mass-produced written tests of intelligence. Inquiries flooded in from schools and businesses. In his massive monograph in 1921 on Psychological Examing in the United States Army, Yerkes buried a statement of great social significance in an aside. He spoke of “the steady stream of requests from commercial concerns, educational institutions and individuals for the use of army methods of psychological examining or for the adaptation of such methods to special needs.” Tests would now rank and stream everybody; the era of mass testing had begun.

Boring selected 160 thousand cases from the files and produced data that reverberated through the 1920s with a hard hereditary ring. The task was a formidable one.
The sample, which Boring culled himself with the aid of only one assistant, was very large; moreover, the scales of three different tests (Alpha, Beta and individual) had to be converted to a common standard so that racial and national averages could be constructed from samples of men who had taken the tests in different proportions (few Blacks took Alpha, for example).

From Boring's ocean of numbers, three "facts" rose to the top and continued to influence social policy in America long after their source in the tests had been forgotten.● The average mental age of White American adults stood just above the edge of moronity at a shocking and meagre 13. Terman had previously set the standard at 16. The new figure became a rallying point for eugenicists who predicted doom and lamented our declining intelligence, caused by the unconstrained breeding of the poor and feeble-minded, the spread of Negro blood through interbreeding, and the swamping of an intelligent native stock by the immigrating dregs of southern and eastern Europe.
● European immigrants could be graded by their country of origin. The average man of many nations was a moron. The darker people of southern Europe and the Slavs of eastern Europe were less intelligent than the fair people of western and northern Europe. Nordic supremacy was not a jingoistic prejudice. The average Russian had a mental age of 11-34; the Italian, 11-01; the Pole, 10-74. The Polish joke attained the same legitimacy as the moron joke—indeed, they described the same animal.
● The Negro lay at the bottom of the scale with an average mental age of 10-41. Some camps tried to carry the analysis a bit further, and in obvious racist directions. For example, at Camp Lee, Blacks were divided into three groups based upon the intensity of skin color; as might be expected the lighter groups scored higher!

The grand average of 15 had political impact, but its potential for social havoc was small compared with Yerkes's figures for racial and national differences; for hereditarians could now claim that the fact and extent of group differences in innate intelligence had finally, once and for all, been established. Yerkes's disciple Carl Brigham, then an assistant professor of psychology at Princeton University, published in 1923 a book, short enough and stated with sufficient baldness (some would say clarity) to be read and used by all propagandists. A Study of American Intelligence became a primary vehicle for translating the army results on group differences into social action. Yerkes himself wrote the foreword and praised Brigham for his objectivity.

Once he had proved that the tests measure innate intelligence, Brigham devoted most of his book to dispelling common impressions that might threaten this basic assumption. The army tests, for example, assessed Jews (primarily recent immigrants) as quite low in intelligence. Does this discovery not conflict with the notable accomplishments of so many Jewish scholars, statesmen, and performing artists? Brigham conjectured that Jews might be more variable than other groups; a low mean would not preclude a few geniuses in the upper range. In any case, Brigham added, we probably focus unduly on the Jewish heritage of some great men because it surprises us: "The able Jew is popularly recognised not only because of his ability, but because he is able and a Jew." "Our figures, then, would rather tend to disprove the popular belief that the Jew is highly intelligent."

One persistent correlation threatened Yerkes's hereditarian convictions, and his rescuing argument became a major social weapon in later political campaigns for restricting immigration. Test scores had been tabulated by country of origin, and Yerkes noted the pattern so dear to the hearts of Nordic supremacists. He divided recruits by country of origin into English, Scandinavian and Teutonic on one side, and Latin and Slavic on the other, and stated "the differences are considerable (an extreme range of practically two years mental age)"; favouring the Nordics, of course.

But Yerkes acknowledged a potential problem. Most Latins and Slavs had arrived recently and spoke English either poorly or not at all; the main wave of Teutonic immigration had passed long before. According to Yerkes's protocol, it should not have mattered. Men who could not speak English suffered no penalty. They took

The results of the Army intelligence tests influenced immigration policy in the States in the 1920s. The quotas were based upon the arrivals before 1890. Immigrants up to that year were mainly Nordic, and supposedly more intelligent than the southern and eastern Europeans who arrived later

Jews came out badly in the tests. Notable Jews were explained by the fact that the public noticed the few great ones (for example, Einstein) because they were so rare

Beta, a pictorial test that supposedly measured innate ability independent of literacy and language. Yet the data still showed an apparent penalty for unfamiliarity with English. Yerkes had to admit "There are indications to the effect that individuals handicapped by language difficulty and illiteracy are penalized to an appreciable degree in Beta as compared with men not so handicapped."

Another correlation was even more potentially disturbing. Yerkes found that average test scores for foreign-born recruits rose consistently with years of residence in America. Did this not indicate that familiarity with American ways, and not innate intelligence, regulated the differences in scores? Yerkes admitted the possibility, but held out strong hope for a hereditarian salvation. The Teutonic supremacists would soon supply that decision:
recent immigration had drawn the dregs of Europe, lower-class Latins and Slavs. Immigrants of longer residence belonged predominantly to superior northern stocks. The correlation with years in America was an artefact of genetic status.

The army data had their most immediate and profound impact upon the great immigration debate, then a major political issue in America, and ultimately the greatest triumph of eugenics. Restriction was in the air, and may well have occurred without scientific backing. But the timing, and especially the peculiar character, of the 1924 Restriction Act clearly reflected the lobbying of scientists and eugenacists, and the army data formed their most powerful battering ram.

Henry Fairfield Osborn, trustee of Columbia University and president of the American Museum of Natural History, wrote in 1923, in a statement that I cannot read without a shudder when I recall the gruesome statistics of mortality for the First World War:

I believe those tests were worth what the war cost, even in human life, if they served to show clearly to our people the lack of intelligence in our country, and the degrees of intelligence in different races who are coming to us, in a way which no one can say is the result of prejudice . . . We have learned once and for all that the Negro is not like us. So in regard to many races and sub-races in Europe we learned that some which we had believed possessed of an order of intelligence perhaps superior to ours [read Jews] were far inferior.

Congressional debates leading to passage of the Immigration Restriction Act of 1924 continually invoked the army data. Eugenacists lobbied not only for limits to immigration, but for changing its character by imposing harsh quotas against nations of inferior stock—a feature of the 1924 act that might never have been implemented, or even considered, without the army data and eugenistic propaganda. In short, southern and eastern Europeans, the Alpine and Mediterranean nations with minimal scores on the army tests, should be kept out. The eugenacists battled and won one of the greatest victories of scientific racism in American history. The first restriction act of 1921 had set yearly quotas at 3 per cent of immigrants from any nation then resident in America. The 1924 act, following a barrage of eugenistic propaganda, reset the quotas at 2 per cent of people from each nation recorded in the 1890 census. The 1890 figures were used until 1930. Why 1890 and not 1920 since the act was passed in 1924? 1890 marked a watershed in the history in immigration. Southern and eastern Europeans arrived in relatively small numbers before then, but began to predominate thereafter. Cynical, but effective. “America must be kept American,” proclaimed Calvin Coolidge as he signed the bill.

Brigham had a profound change of heart six years after his book, but he could not undo what the tests had accomplished. The quotas stood, and slowed immigration from southern and eastern Europe to a trickle. Throughout the 1930s, Jewish refugees, anticipating the Holocaust, sought to emigrate, but were not admitted. The legal quotas, and continuing eugenic propaganda, barred them even in years when inflated quotas for western and northern European nations were not filled. Estimates suggest that the quotas barred up to six million southern, central, and eastern Europeans between 1924 and the outbreak of the Second World War (assuming that immigration had continued at its pre-1924 rate). We know what happened to many who wished to leave but had nowhere to go. The paths to destruction are often indirect, but ideas can be agents as sure as guns and bombs.