Modern Intellectual History, 6, 3 (2009), pp. 637–644
doi:10.1017/S1479244309990229

MEASUREMENT AND MERITOCRACY: AN INTELLECTUAL HISTORY OF IQ

THEODORE M. PORTER
Department of History, University of California, Los Angeles
E-mail: tporter@ssc.ucla.edu

John Carson, *The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics, 1750–1940* (Princeton, NJ: Princeton University Press, 2007)

Is *intelligence* a fit topic for *intellectual* history? The creation and institutionalization of IQ (the initials have become self-sufficient, and no longer stand for "intelligence quotient") have been a favorite topic in the history of psychology, and have even achieved some standing in social histories of class, race, and mobility, especially in the United States. The campaign to quantify intelligence tended to remove it from the domain of intellectual history, which after all has traditionally emphasized ideas and interpretations. Measurement, and not alone of the mind, was pursued as a way to rein in the intellect by making it more rigorous. What was pushed out the door, however, returned through the window in the form of debates about what intelligence means; in what sense and with what tools it can be measured; and how these measures relate to other ways of comprehending mind, thought, and reason. Quantification, a potent strategy for releasing science from the grip of history, is itself profoundly historical, as a half-century of modern scholarship has demonstrated. This historicizing of the antihistorical embodies what we may call counterreflexivity, and, as such, is partly about puncturing illusions, though it need not take a negative view of the social role of science. The perspective of history is all the more essential because the depoliticization of merit through science entails a consequential moral and political choice. Measurement, by rationalizing and stabilizing the idea of intelligence, enabled it more readily to enter everyday discourse and to be put to work in schools, businesses, and bureaucracies.

John Carson's deep and wide-ranging book gives a rich account of a concept that came to be quantified. It is exemplary for its alertness to the evolving meanings and changing patterns of usage of words like "talents" and "intelligence," and still more for the ways it situates these ideas historically and

shows how they were made to matter. As his title suggests, the problematic for Carson's study arises from the tension created by the persistence of inequality in a democratic age. Merit (as in "meritocracy," with its implication that superior individuals deserve their success) offers a plausible rationale for unequal power and wealth in a polity devoted to equality before the law. But more than mere rationalization, the changing definition of merit, along with new ways of operationalizing it, has had consequences.

These are apparent already in the working out of Carson's first main topic, the origin of the modern sense of intelligence as a more or less unitary thing. Through the eighteenth century and into the nineteenth, meritocrats in France and the United States alike spoke mainly of talents or abilities, which are plural, while the word "intelligence" was not yet used to refer to an encompassing mental capacity. He shows that the conception of intelligence as a magnitude, of which one being can be said to have more and another less, was used to order biological species before it was applied to human individuals. Pre-Darwinian naturalists took up ranking animals by their level of intelligence, and almost simultaneously applied the same metric to human races. Racial intelligence was established, usually by physicians, on the basis of cranial volume, facial angle, or the shape of the skull. And yet the naturalistic perspective did not necessarily support a unitary notion of intelligence, but could just as well disassemble it. Phrenology, an insurgent, materializing science of the brain that achieved considerable respectability in the mid-nineteenth century, identified numerous, independent human faculties and located each in a different region of the brain. Its differentiation of mental capacity followed the psychology of Scottish common sense philosophy, which emphasized plural talents rather than sheer brainpower.

The battle of unitary versus pluralistic intelligence persisted into the age of measurement, and has been one of the key political issues about IQ testing in modern times. Stephen Jay Gould's well-known essays in *The Mismeasure of Man* deal especially harshly with the idea that intelligence reduces to a "general factor," *g*, which is largely biological or hereditary. The statistical concept of *g* was put forward in 1904 by the British psychologist Charles Spearman, using the new method of correlation articulated by the eugenicists and statistical pioneers Francis Galton and Karl Pearson. Gould's scathing critique was directed partly at Spearman but mainly at more recent champions of unitary, biological IQ such as the Berkeley educationist Arthur Jensen. Gould assigned great significance to a controversy between American champions of IQ, such as the Stanford psychologist Lewis Terman, and critics, especially L. L. Thurstone of the University of Chicago, who applied a different statistical technique in order to break intelligence up into several overlapping components. Carson is less impressed by their debate, arguing that on most points the

American psychologists of IQ were in agreement. And yet a different version of this distinction between unitary and multifaceted intelligence reappears in his book as a contrast between American and French interpretations of quantified intelligence.

The Measure of Merit is a comparative study of the vocabulary and, eventually, the machinery for assessing talents and intelligence in France and the United States. Although the book is full of original material as well as new interpretations, this comparative dimension is what really sets it apart from existing scholarship on mental testing, of which there is much, especially on the United States. His comparison, moreover, is at a fundamental level, and has far-reaching implications. Carson ranges over two centuries in order to get at deep differences in intellectual and political cultures, and shows with admirable specificity how they played out in school systems and military organizations, even at the level of the design and administration of mental tests. The book is formidably learned and densely documented, and takes some time to assimilate, though the writing is always clear.

In this as in many works of comparative history, the resemblances tend to be kept in the background as the author works to articulate and document contrasts. Yet it is important to note that a sequence of shared ambitions in France and the United States runs through the whole period of the book: first a conviction in the eighteenth century and the early nineteenth that differential "talents" merit unequal positions and rewards; then a pervasive concern with racial distinctions, intensifying in the 1850s and 1860s; and culminating in efforts in both countries to create question-based tests of intelligence starting about 1900. The book makes perhaps a stronger case for the similarities than its author seems to intend. He mentions several times that the French, in contrast to the Americans, never supposed that a single test of intelligence could replace a more nuanced and comprehensive look at the whole person. The point has its validity, but Carson is hard-pressed to find American psychologists asserting in a direct way the contrary proposition. Certainly the US military's reliance on IQ results during the First World War—one of the seemingly well-worn topics on which Carson has come to original conclusions on the basis of important new materials—never went beyond what French psychologists could have accepted. US schools and colleges were indeed more inclined than French ones to use IQ as an automatic basis for sorting students or admitting and rejecting applicants. Carson includes rather little about this topic, perhaps because there is already much excellent scholarship. I would mention especially Kurt Danziger's book *Constructing the Subject* (1990), which links IQ testing and the psychology of individual differences to mass education based on age-graded schools. The sufficiency of IQ as a measure of ability or intellectual promise was, I think, more often a practice than an explicitly argued scientific position.

The French, Carson argues, were inclined to prefer a clinical perspective to the objectifying one of the standardized mental test. This is one of his chief points of contrast between French and American uses of "intelligence." He makes the argument in an appropriately subtle way, showing that Alfred Binet, the French pioneer of the intelligence test, went some way toward an objective, quantitative ideal in the scale of intelligence that he worked out with Théodore Simon. This test had become the basis in France for administrative routines of classification by 1911. Carson contrasts the bureaucratic aspect with Binet's earlier, lovingly detailed book examining the intelligence of his two daughters. American assessments of intelligence, too, had an important clinical aspect in the first decade of the twentieth century. Leila Zenderland brought out the mixture of clinical and impersonal characteristics of American testing in her book *Measuring Minds* (1998). Her protagonist, Henry Goddard, worked at an institution for mental defectives in Vineland, New Jersey, and became the most influential champion in America of Binet's tests before Terman. It was the army tests of 1917–18, applied in great haste on a scale of millions, that marked the decisive departure of testing in America from a clinical ideal.

To be sure, the French educational system also depended on examinations that were administered on a vast scale. These were perhaps not so fully standardized, and at least could not be scored automatically like an American multiple-choice examination, but admission into *lyçées* and *grandes écoles* was, by 1900, determined primarily by examination results. In much of Europe and beyond, tests like the English "eleven-plus" were used through the 1950s to determine which students would be placed on a university track, and which on a vocational one. The results of these tests could be far more decisive for the fate of a child than any American IQ test. The crucial difference was that the European examinations were based on a relatively uniform curriculum, and the French ones in particular on a national one. That is, the national tests were in a way intrinsic to the educational system, and reinforced its standards. Spearman's understanding of *g*, general intelligence, was also linked to examinations in school subjects, and he argued that the most prestigious school subjects, especially Greek and Latin, were the best indicators of general intelligence. American mental testing (for college admissions), by contrast, was designed to get at something independent of the curriculum. American schools, after all, were locally controlled and extremely heterogeneous. Aptitude tests, which closely resembled IQ tests, provided selective universities with a relatively uniform basis for comparing students from high schools of which they knew nothing with those from the elite boarding schools and grammar schools with which they had developed close working relationships.

The working out of this contrast between the relatively unified administrative structures and secure authorities in France and the more ramshackle American

state is the greatest contribution of Carson's comparative approach. The two systems gave rise to tests claiming different forms of objectivity, for what the French grounded in the legitimate authority of the state was more often, in America, explicitly marked off from the contingent, localized domains of politics and culture. American intelligence was commonly understood as biological, so that its system of testing, which accommodated great diversity in schooling, tended also to fix every individual on a single scale of ability. American psychologists were prone to interpret ethnic and racial differences in mean test scores as evidence of biological hierarchy, akin to the sense of hierarchy that had given validity to quantified "intelligence" in the first place. This aspect has been particularly controversial, and has had a lot to do with the outpouring of historical attention to IQ and aptitude testing.

Paradoxically, then, the American search for objective intelligence, free from everything contingent and local, was an accommodation to distinctive institutional circumstances. The diversity of schools in the United States, as much as the diversity of its population, meant that any definition of intelligence and any scheme for measuring it could become controversial. And measures of intelligence have in fact been adjusted in response to such pressures, sometimes by trying to neutralize discontent by emphasizing the technical demands of the field and insisting on the objectivity of the measures, but also sometimes by trying to accommodate criticism. Although tests of IQ and of "aptitudes" have been particularly vulnerable to the charge that they give advantage to children of the prosperous and members of the most powerful ethnic and social groups, their consequences have not been so simple. Some immigrant populations who were mostly poor and uneducated when they arrived in the United States were helped by the tests in their socioeconomic ascent. An abundance of ironies and surprises were inevitably brought forth by this effort to quantify a thing so elusive as intelligence and to make it stand still. The irony, indeed, is heightened by the typical defense of IQ as a psychological and biological fact, the outcome of objective scientific investigation.

While sympathetic to the critiques, Carson does not swallow them whole. Contingency and complexity have been watchwords of historical writing since the 1970s, when the model of quantitative social science fell out of favor among historians. That rising sensibility no doubt had something to do with the more critical attitude that historians brought to science at just this time. Yet the idea of hegemonic and "totalizing" discourses, especially when they presumed that science now ruled the roost, involved simplification as extreme as any "pretensions" of science. A growing corpus of scholarship, Carson's book now preeminent within it, has shown that the science of IQ was not monolithic, and that the institutions targeted by educational psychologists, such as schools, were never putty in anyone's hands, but had their own interests and traditions.

The uses of mental testing were promoted and constrained by the demands and expectations of parents, businesses, urban reformers, and various levels of government.

The critics of IQ preferred to see it otherwise. Gould's *Mismeasure of Man*, published in 1981 and translated into many languages, quickly became and has remained the most influential account of the history of mental measurement, among nonspecialist scholars as well as the general public. It illustrates by way of a sequence of exemplary cases a stubborn ambition to measure and thereby to demonstrate scientifically the importance of biological inequality, racial as well as individual. Although Gould's scientific critique has itself been controversial, I find his arguments regarding the heterogeneity of mental abilities and the hazards of reifying a unitary measure of mental ability to be moderate and sensible. His historical plot line, however, consists mainly of reiteration rather than of mapping out a real trajectory, and the frame of his historical argument is not much more than a universalizing of what he finds in the episodes that he features. Gould portrayed his anthropometricians and psychologists as part of their society in that they shared its racism and acceptance of inequality, but in a notably simple way. Thus his book omits the cultural richness that historians normally expect in accounts of this kind.

Especially in intellectual history, we often speak of that crucial dimension, shortchanged by Gould, as "context," though oftentimes the considerations in question are conspicuously evident in the texts themselves. Just because the old history of ideas often held itself aloof from the more mundane institutions and practices of life, more recent historians of science and art, of philosophy, scholarship, and rhetoric, have found it all the more exciting to reintegrate what had been torn away. In doing so they (we) also are chipping away at the public image, indeed the self-conception, of "intellectuals" and scientists, who typically have argued in universals. These individuals speak and write of what should be valid for everyone, of what is true or just or good. Intellectual history and history of science aspire to undermine these universals by recovering their specificity and locality. The drive for generality and techniques for achieving it, which must be given their due, are always, we insist, displayed in more particular forms and adapted in part to specific sites.

Much of the writing in Carson's *Measure of Merit* is about the ideas and the work of individual thinkers and researchers, from Helvétius, Rousseau, and Jefferson to Samuel G. Morton, Paul Broca, Alfred Binet, Hippolyte Taine, Edward Thorndike, Robert Yerkes, and Lewis Terman. They write at their desks and work in their laboratories, while addressing sometimes their disciplinary peers, sometimes their political or bureaucratic patrons, and sometimes what by about 1800 was becoming known as "the public." They labored disinterestedly in ways that were curiously self-interested, trying to reform, or sell their wares to, schools,

armies, kings, presidents, and legislators. They worried about democratization and rising tides of mediocrity, but also about a self-perpetuating elite that was failing to recruit the most intelligent and meritorious into its ranks. They were increasingly conscious of themselves as "scientists"—a word that only came into common usage toward the end of the nineteenth century—yet did not wish to be written off as mere technical specialists. As late as the early twentieth century, the leaders of science sought a place within or even atop a more encompassing political, economic, and cultural elite. Many also envisioned their disciplines as fulfilling a vital social role. Psychologists, for example, abstracted from the assumptions of their class and nation to frame universalistic arguments. The highest social purpose of intelligence measurement was the establishment of an efficient system for selecting and training up leaders.

It is never the place of a new academic discipline, or even a well-established one, to settle matters so weighty as this, but neither were psychologists powerless. They were able to gain some leverage where the authorities were divided or uncertain how best to proceed, and were recruited where they could advance programs favored by the powerful. They cultivated these authorities as patrons, yet it is too simple to treat them as automatically aligned with a dominant class. Carson shows how discourses of intelligence and merit took shape on a terrain of contestation. He also is alert to the role of incipient disciplines and professions in shaping the social role of social science. In these respects, he follows the best practice of modern intellectual history. What I find most original and interesting in this book, along with its comparative dimension, is Carson's close attention to the institutions for which measures of intelligence and merit were designed, such as schools, armies, and homes for "defectives." The best historians of psychology and psychiatry, including Danziger and Zenderland, had already mapped out this territory, but nobody has been more skillful than Carson in showing how the ideas and practices of mental measurement grew up in alliance with the institutions they served.

The influences, if we are to use this term, were reciprocal. Even when psychology seemed to be serving in a purely technical capacity, its role was also creative, and even if academic experts often looked with disdain on practitioners of applied social measurement, the tools and concepts of "basic science" were shaped in part by these contexts of application. Such success as psychologists had in their expert role depended greatly on their ability to adapt to the institutions they served, and in many respects the masters in measurement did succeed. Even the US Army, hardly an institution that we would expect to let academic psychologists judge the abilities of its men, took seriously the results of the hundreds of thousands of tests administered to new recruits. Carson scrutinizes carefully the interactions between psychologists and Army brass, and shows that IQ measures mattered more than most historians have supposed.

Interactions of this kind, more variegated that what we generally associate with the great philosophers and social theorists, are perhaps characteristic of more institutionalized bodies of knowledge (and practice) than of free-floating intellectuals. Even the latter cannot really float so free, but operate on an intricate terrain defined by patrons, audiences, publishers, genres, and colleagues. Neither for intellectuals nor for scientists should historians aim to reduce the significance of thought and research to the local sites where, in the narrowest sense, they take place. There are, still, big pictures to be drawn and grander narratives to develop. These, however, should incorporate the nuances of historical specificity. Within the complex and detailed exposition of Carson's book, there is an important story of historical developments on a grand scale, of the development of tools and concepts for measuring intelligence, and of the changing role of intellect in public life from the eighteenth-century Enlightenment to the Second World War.