Echoes of a Forgotten Past: Eugenics, Testing, and Education Reform

by Alan Stoskopf

The children of today must be viewed as the raw material of a new State; the schools as the nursery of the nation.

—Ernest Bryant Hoag and Lewis M. Terman (1914)

The metaphors employed by educational psychologist Lewis Terman formed part of an eugenic ideology that guided educational reform in the early 20th century. Echoes from this past can also be heard in today's calls for more high-stakes testing. An investigation of the role eugenics played in the history of testing is more than an arcane study of a forgotten chapter in U.S. education. It raises fundamental questions about where today's reform efforts are heading.

Sir Francis Galton, an English mathematician, invented eugenics. He conceived of it as an interdisciplinary field of scientific inquiry with a civic purpose. As Galton (1919, 1) wrote, "Eugenics is the study of the agencies under social control that seek to improve or impair the racial qualities of future generations either physically or mentally." What Galton saw as a new branch of scientific inquiry would become a dogmatic prescription in the ranking and ordering of human worth. His ideas found their most receptive audience in the first decades of the 20th century in the United States.

When eugenics is mentioned in the history of education, it is usually as a footnote to the social-efficiency and scientific-management movement in U.S. schooling. In fact, it was also a widely accepted theory of human differences that deeply influenced the thinking of such reformers as Terman, Henry Goddard, and Edward Thorndike. These three educational psychologists made important contributions in the construction and administration of standardized tests; they also were part of the mainstream thinking of their time, embracing unexamined beliefs about human potential. These beliefs would influence the direction of school reform then and create legacies for us to confront today.

Eugenic ideas would need the right social context to emerge and thrive. The United States at the turn of the century was experiencing large-scale immigration from Southern and Eastern Europe as well as the beginning of African-American migration from the Jim Crow South to Northern cities. Competition for jobs intensified existing frictions along class and racial lines. It was out of this cauldron of social upheaval that the U.S. eugenics movement began. For many native-born European Americans, these changes threatened their social status. Eugenic advocates articulated those
fears through the language of academic research and social reform (Beckwith 1993).

However, eugenicists used a flawed and crude interpretation of Gregor Mendel's laws on heredity to argue that criminality, intelligence, and pauperism were passed down in families as simple dominant or recessive hereditary traits (Cravens 1988). Many eugenicists believed that some individuals and entire groups of people (such as Southern Europeans, Jews, Africans, and Latinos) were predisposed to carry a disproportionate number of defective traits in the general population. These ideas began to permeate different disciplines in higher education in the 1910s, especially biology and psychology (Selden 1999). One of the first educational psychologists to become an ardent proponent of eugenics was Henry Goddard (1866–1957). His innovation was to infuse eugenics into educational theory and practice.

**GODDARD: THE THREAT OF THE 'FEEBLEMINDED'**

Goddard was the Director of the Vineland Training Center for Feebleminded Boys and Girls (in New Jersey) between 1906 and 1918. When he assumed leadership of the institution, Goddard was already interested in eugenics. It seemed to provide a scientific explanation for why some students were "slow learners."

Vineland served as a sort of holding center for children and adolescents who had been deemed "feebleminded." This term was a popular eugenic catch phrase for anyone thought to be mentally deficient. The feebleminded were considered hereditarily prone to crime, poverty, and a host of moral improprieties. For example, feebleminded girls were believed to be at risk for prostitution (Kevles 1995). The problem for Goddard and other eugenicists was the absence of an accurate device to identify the feebleminded. This changed with the IQ test.

Goddard had become familiar with the work of French psychologist Alfred Binet on school children in Paris, France. Binet had developed a series of protocols to assist teachers in identifying students who were not performing to grade level. Goddard saw Binet's work as a perfect tool for measuring the eugenic worth of an individual. Binet was not a eugenics advocate; he was much too uncertain about what intelligence was and believed that a person could adapt and intellectually grow over time (Gould 1996). Goddard discounted Binet's reservations. To him, intelligence was fixed and finite. Goddard thought a score on a standardized test would be the best predictor for what a person was capable of doing in life.

Goddard translated Binet's test into English and decided to use it. He first experimented with it on inmates at the

---

*Alan Stoskopf is Director of the Education for Democracy Initiative at Facing History and Ourselves, a national not-for-profit education foundation in Brookline, Massachusetts. His research interests address innovative ideas and practices in preservice education.*
Vineland Center, but he wanted a larger sample to demonstrate its wider applicability. Ellis Island provided the solution. Between 1912 and 1913, Goddard and his field workers, in conjunction with the Public Health Service, administered the first version of a standardized IQ test to a large population in the United States (Kraut 1994). The results of his study on immigrants produced headlines in academic journals and the popular press. According to Goddard’s (1917, 252) findings, “83% of the Jews, 80% of the Hungarians, 87% of Russians, and 79% of the Italians tested below the 12 year old limit and were therefore feebleminded.” Goddard (1917, 266) went on to say, “Not only are these figures representative of these ethnic groups as a whole they are probably too small.” It was an indication of the racial and class attitudes of the time that these claims were well received among many in the general public and the academy.

Goddard pushed his test for helping with student placement in schools and in teacher training. He was hired as a consultant by the New York public schools to advise on the placement of students into specialized learning tracks, which the city was already pursuing (Travers 1983). Goddard’s (1914) IQ tests could help determine which students belonged in special classes and which students should be sent to separate schools.

Goddard was on the cusp of something big. Eventually, the Vineland Center trained more than 1,000 teachers in test administration and interpretation. As Wallin (1953, 54) commented, “Most of the teachers became trail blazers in their home communities. Many rose to positions of leadership.” These teachers learned to apply terms like moron and defective learner to children who scored low on the tests, separating the feebleminded student from “normal” learners (Trent 1994).

**Terman: Tests Tell the Truth**

Terman (1877–1956), another educational psychologist, built upon Goddard’s initial work. He spent the majority of his academic career in the College of Education at Stanford University (1910–22) and then as Chair of the Psychology Department (1922–42). Terman, a pioneer in the development of intelligence and achievement tests, was one of the first to use them for longitudinal research. In his most famous study, he tracked 1,470 students with high IQs into their adult lives, dispensing the popular belief that gifted children were more emotionally unstable than anyone else in the general population (Hoff 1999).

Terman also embraced eugenics. This was evidenced in his work with Goddard and Robert Yerkes, administering IQ tests to more than 1.7 million army recruits beginning in 1917. Their questions revealed blatant class, cultural, and ethnic bias (Brigham 1923; Gould 1996). The results of that work would have a tremendous impact on U.S. public policy, especially in influencing the passage of the Immigration Restriction Act of 1924 (Chase 1976). It also catapulted testing into the forefront of educational reform in U.S. schools.

When Terman arrived at Stanford, schools were already under attack in the popular press for being inefficient and unaccountable when compared to modern businesses (Callahan 1962). The “efficiency expert” represented a new breed of educational reformer. The scientific-management model was concerned with eliminating waste and training students to find their productive place in the world (Kliebard 1986; Spring 1986).

Terman saw a natural fit between an industrial model and his eugenic-inspired work with testing. So did Ellwood Cubberley, the man who hired Terman at Stanford. Cubberley (1916, 338), Dean of the College of Education, was already known
as an advocate for cost accountability in education and for his vision of schools as “factories in which the raw products [students] are to be shaped and fashioned into finished goods.”

At Stanford, Terman developed a new set of scoring scales for schools. These scales, based on a small sample of 982 native-born children from European-American, middle-class, Protestant homes in the Palo Alto area (Gersh 1981; Cravens 1988), formed the basis of the well-known Stanford / Binet intelligence test. The assumptions that went into the creation and use of the test were fundamentally eugenic. As Terman (1916, 91) wrote, there were thousands of feebleminded among “laboring men and servant girls... . As far as intelligence is concerned, the tests have told the truth... . No amount of school instruction will ever make them intelligent voters or capable citizens in the true sense of the word.” Terman (1916, 92) then advocated “that the whole question of racial differences in mental traits will have to be taken up anew and by experimental methods,” due to the prevalence of deficiencies among “Indians, Mexicans, and negroes.” In frank language, Terman (1916, 92) concluded, “Children of this group should be segregated in special classes and be given instruction, which is concrete and practical. They can not master abstractions, but they can often be made efficient workers, able to look out for themselves. There is no possibility at present of convincing society that they should not be allowed to reproduce, although from a eugenic point of view they constitute a grave problem because of their unusually prolific breeding.”

In a few years, thousands of schools took this view to heart. By 1925, a survey of 215 cities found that nearly 65 percent of the districts used standardized IQ tests to classify students into tracked classes (Marks 1976–77). Terman (1923) also invented a series of achievement tests used to calibrate the correct placement of students into homogeneous classes.

Terman’s influence on education was reflected in the offices and awards he received. In 1922, he was elected President of the American Psychological Association; by the middle of the decade, he became the editor of six journals relating to educational research and psychology. Clearly, Terman advanced the understanding of the technical possibilities in the creation and administration of tests, an achievement acknowledged to this day. At the same time, his eugenic arguments affected how he interpreted test scores, which also impacted education.

By the end of the decade, more than 130 different standardized tests were used to classify the fast-growing student population in U.S. schools (Hildreth 1933). What seemed efficient and natural to some teachers and administrators turned into diminished educational expectations for hundreds of thousands of school children. Disproportionate numbers of African-American, Latino, and immigrant children from Southern and Eastern Europe were put in non-academic tracks because of test scores, especially in urban school districts (Gersh 1981). However flawed the methodology might have been in constructing some tests and interpreting scores, they increasingly played a major role in deter-
mining the academic fate of school children—and the fates were unequal. Both the learning tracks established within schools and the differences on test scores between unequally funded school districts have been closely correlated with the racial and class backgrounds of students (Oakes, Gamoran, and Page 1992).

**Thorndike: The Cult of the Expert**

According to Thorndike (1874–1949), Professor of Educational Psychology at Columbia University, this state of affairs was not a problem. Rather, it presented a solution for where schools needed to go. Thorndike had close ties with Terman. Both had worked on the Army tests in 1917 and collaborated on the first National Intelligence Tests. Thorndike would use Teachers College as his base of operations from 1899 until the time of his death (Jončich 1968). During his long career at Columbia, he wrote 50 books and 450 articles. Thorndike did much to establish the notion of a norm-referenced test. This was a valuable contribution to the field of assessment, and normed tests have since been widely used in all sectors of society.

Thorndike institutionalized Terman’s myth that the “tests have told the truth” in U.S. schools. This effort sprang from a eugenic conception of human worth. As Thorndike (1940, 957) declared, “By selective breeding supported by a suitable environment we can have a world in which all men will equal the top ten percent of present men. One sure service of the able and good is to beget and rear offspring. One sure service (almost the only one) which the inferior and vicious can perform is to prevent their genes from survival.”

As Karier (1975) has noted, the school’s job was to make sure that the “inferior,” meaning the lower classes and races, found their proper slot in the educational mill and the “able” were allowed to fulfill their destiny to lead. Thorndike’s prestige made these ideas appear unassailable to educators working in the schools (Gersh 1981).

Thorndike argued against the “common man” trying to outthink the expert when it came to educational policy. As Thorndike (1920, 670) suggested, “Whenever there is the expert . . . should we not, in fact, let him do our thinking for us in that field?” These ideas became a selling point for a vision of educational reform that required very little original thinking on the part of the teacher and administrator. Their job was to execute the technical functions of helping students realize their proper station in life.

Superintendents quickly endorsed these views in schools and at national conferences (Chapman 1988). David Corson (1920, 87), superintendent of schools in Newark, New Jersey, proclaimed, “All children are not born with the same endowment of possibilities; they cannot be made equal in gifts or development or efficiency. The ultimate barriers are set by a power inexorable. The radical and absolutely necessary action is to accept a classification of children according to ability and attainment, not according to physiological age.” Paul Stetson (1921, 353), a superintendent in Michigan, envisioned that as “intelligence tests are perfected . . . it may be possible to classify pupils entirely on the basis of intelligence tests.”

**Dissident Voices**

There was, however, resistance to this version of education reform. Across the country in the 1920s, reports and surveys indicated lukewarm support for the use of tests among many teachers and principals (Chapman 1988). Furthermore, a lively dissent among some educators and journalists in the popular press about the uses and effects of these tests on school life ensued. A few well-known educators, including
John Dewey and Otto Klineberg, attacked some of the assumptions and methodology of the tests (Travers 1983). Though largely unnoticed in mainstream academic journals, African-American educators like Horace Mann Bond and Martin Jenkins also criticized this work. Their efforts laid the foundation for better-funded research efforts in the 1930s and later that would attack the hereditary assumptions of people like Terman and Thorndike (Stoskopf 1999).

Nonetheless, eugenic notions of human differences were in the ascendancy along with the use of IQ and achievement tests to stream students into ability-group tracks (Wigdor and Garner 1982). Goddard, Terman, and Thorndike were not the exceptions; they were closer to the rule at this critical time of school growth and reorganization. By 1928, 376 courses in eugenics were offered in colleges and universities across the country (Cravens 1988).

Eugenics began to wane in the 1930s. Advances in scientific thought undermined its simple genetic determinism. The Great Depression also impoverished many middle-class families, making it hard to accept that only those born with tendencies toward “pauperism” could become poor. Slowly, increasing diversity within the research community and social activism in the larger society made it more difficult for eugenic ideas to remain prominent in education. Yet, the use of standardized IQ tests and the accompanying hereditary beliefs behind the tests did not go away so easily.

LEGACIES AND QUESTIONS FOR TODAY

Today, standardized tests are no longer used to measure the eugenic worth of students. They can provide important diagnostic information for educators. These tests can supplement ongoing and authentic assessment of student work. Despite periodic controversies about built-in biases in various commercial tests, researchers have worked hard to eliminate the obvious prejudices that test makers committed in the 1920s (Hoff 1999). Yet legacies from this past haunt the current wave of education reform.

The Business Model

In the 1920s, analogies from business were uncritically applied to school organization and instruction. Little attention was paid to the developmental needs of individual students and to pedagogical strategies needed to reach diverse learners. Students were seen more as products to be shaped than as active contributors to their own learning.

The contemporary standards movement linked to high-stakes testing began in the 1980s with fears of the United States losing its global competitive edge in the international marketplace. A Nation at Risk (National Commission on Excellence in Education 1983), issued by the Reagan administration, was its clarion call. Since then, we have been inundated with state testing mandates demanding greater accountability.

Once again, analogies from the business world are pushing aside the educational discussion on how students learn and teachers teach. Architects of “reform” tend to be business leaders, governors, and test-preparation experts who rarely involve a diverse group of educational researchers or, more importantly, teachers from our classrooms in the discussion of what reform should look like (Sacks 1999).

Single-Number Scores

In the era of scientific efficiency and eugenic reform, students were seen as capital, each with a different worth. Some were considered more deserving of a college-preparatory curriculum than others.

Today, reform does not lead with these beliefs. Instead, many who embrace high-
stakes testing urge its use to elevate standards for those most underserved and provide them with more demanding academic opportunities. These intentions are worthy, but the translation of the current reform effort into the lives of students and teachers can be a very different story.

Simply boosting test scores does not equal substantive educational improvement. Students may not have acquired skills that lead to a deeper engagement in critical thinking or lasting competencies in a particular academic discipline (Thorkildsen and Nicholls 1991; Popham 1999). Also, poor minority-group students often do not benefit from this version of reform. They are retained and drop out in greater numbers than other cohorts of students (Haney 2000).

**Doing Reform to Students and Teachers**

In the 1920s, little attention was paid to the motivational factors behind how students learned and the attitudes and practices of teachers in enhancing student learning. We know more today, but the current vision of reform seems to hearken back to the days when a teacher was the handmaiden of external experts.

In contemporary test-driven reform policies, many teachers complain about a loss of control of their curriculum and the value of local assessments (Harris and Wagner 1993). Rather than a genuine collaborative effort between outside policy makers and teachers in schools, we see a reemergence of outside experts doing the real thinking.

**Stratifying Learning Opportunities**

Academic tracking remains as an institutional legacy. Today, however, tests are not purporting to be sorting devices based on a racial ranking; instead, student, teacher, and school performance is sorted. Yet the current incarnation of high-stakes testing may accomplish the same goals. Furthermore, students who do poorly on these tests may be given crash courses in test preparation, but they often do little to develop the habits of mind needed for sustained critical inquiry (McNeil 2000).

**Assessing Our Past**

These concerns should serve as a springboard for honest reflection and open debate about what we have inherited from the past and what has changed. Certainly, these points could and should be contested in public forums. They also should be discussed within a historical context, something that is not usually done in the adoption of education reforms in this country.

Supporters of high-stakes tests might recoil at the assumptions underlying the use of standardized tests earlier in the century and see no relevance for our time. However, that history forces us to rethink what we mean by reform and the educational assumptions behind any movement for change. Eugenic reform leaders in the past genuinely wanted to improve schools. Indeed, there were major problems in public education. Just having good intentions and tapping into people’s fears, however, is not enough. If left unexamined, this approach creates more problems that it purports to solve. We have seen such problems in the past, and we are beginning to see them today.

In the end, we need a real questioning of what we mean by education reform. Discussion must occur with educational policy makers and representatives from diverse communities in open forums. That is what is supposed to happen in a democracy. If anything, the eugenics movement is a reminder of what can happen when the assumptions and consequences of educational reform are not put to the test of real-life experience. We can not afford to let that happen again.
REFERENCES


Callahan, R. E. 1962. Education and the cult of efficiency: A study of the social forces that have shaped the administration of the public schools. Chicago: University of Chicago Press.


