

# Early Views on Intelligence and Inheritance by Goddard and Terman

Michael Salib

May 16, 2003

Henry Herbert Goddard and Lewis M. Terman were pioneers in the early field of mental testing. Both claimed that human mental ability could be reduced to a single factor called intelligence measured on a linear scale. Furthermore, they advocated the belief that mental tests reliably measured intelligence, and that intelligence was heritable and largely insensitive to environmental change. Significantly, both men justified their beliefs by clothing them in the garb of science, and in particular, mathematical precision. Despite the strength of their claims, much of the science comprising their work and underlying their arguments has been discredited, thus raising the question of how and why did these seemingly bright, well trained men make such grievous errors. I will show that the flaws in their thinking result largely from uncritical acceptance of commonly held but unfounded social beliefs in addition to deep structural flaws at the boundary between science and society. That is, Terman and Goddard came to believe the things that they did because their role in society required them to do so; they both benefited handsomely because of the work they did, regardless of its

objective scientific merit. Both men had a need for intelligence to exist as a reified entity and a need for intelligence testing to accurately measure it. Both men also had a need to believe in the heritability of intelligence.

Much of Goddard's science borders on the ridiculous, and yet few people at the time doubted the legitimacy of his work. One reason was that he included copious amounts of data, threatening to overwhelm skeptical readers in an avalanche of misguided precision (Goddard 1913, chapter 2). In *The Kallikak Family*, Goddard traced the lineage of one of his school's feeble-minded students back across five generations by cataloging information about almost a thousand individuals (Goddard 1913, p. 116). Goddard's belief that copious measurement and observation confers scientific legitimacy is reminiscent of Lavoisier's belief that copious quantities of decimal digits make results unassailable. The need to collect vast quantities of heredity data, without first verifying the validity of the data collection method employed is similar to Lavoisier's striving for ever more precise measures, without first verifying that the extra precision actually had any meaning.

Modern observers claim that Goddard's data do not support the sweeping conclusions he makes. For example, Paul claims that "*The Kallikak Family* data were clearly spurious" (Paul 1995, p. 54) while Gould opines that Goddard's "study is little more than guesswork rooted in conclusions set from the start" (Gould 1996, p. 198). Goddard himself lends credence to this claim by writing in the preface that "it is true that we have made rather dogmatic statements and have drawn conclusions that do not seem scientifically warranted from the data . . . we have done this because it seems necessary to make these statements and conclusions for the benefit of the lay reader, and it was impossible to present in this

book all of the data that would substantiate them” (Goddard 1913, p. X). Goddard’s early admission of exaggeration may explain a rather puzzling feature of his book. According to the director of Photographic Services at the Smithsonian Institution, “the photographs of the Kallikak family members have been retouched [in order to] give the appearance of dark, staring features, sometimes evilness, and sometimes mental retardation” (Gould 1996, p. 201). Interestingly, only the photographs of the Kallikak family in the wild have been manipulated; no other photographs, including those of Deborah, the Kallikak family member at Goddard’s school have been altered (Gould 1996, p. 201).

Perhaps most distressing is Goddard’s continued insistence that Deborah is somehow defective. Although initially arriving at Goddard’s Vineland school for the Feebleminded at the age of 8 with a variety of problems, in time, “she learned to read, add, play the coronet, and sight-read music” (Paul 1995, p. 52). Paul goes on to describe her as “a charming and beautiful woman, an excellent gardener, seamstress, and woodcarver with a cheerful disposition and no obvious defects” (Paul 1995, p. 52). And yet Goddard dismisses the reports of her progress, ascribing it all to the wishful thinking of her instructors, by claiming that “Deborah’s teachers have worked with her faithfully and carefully, hoping for progress, even seeing it where at a later date it became evident that no real advance had been made . . . our work with Deborah convinces us that such hopes are delusions” (Goddard 1913, p. 11–12). While its true that Deborah will likely never graduate from Harvard or Yale, her gifts seem more than sufficient to enable her to lead a happy and productive life. At least her teachers seem to think so, since they write that Deborah can “do practically everything about the house. [She] has no noticeable defect. She is quick and observing, has a good

memory, writes fairly, does excellent work in wood-carving and kindergarten” (Goddard 1913, p. 7).

Goddard went beyond merely asserting that there existed a vast population of feeble-minded people; he ascribed all of society’s ills to them as well. Paul quotes Goddard as claiming that “hereditary feeble-mindedness is the cause of all these [social] problems” (Paul 1995, p. 59). Paul explains that Goddard believed the inability to distinguish right from wrong was caused by feeble-mindedness and “that explains why the rate of feeble-mindedness among criminals and misfits is so high” (Paul 1995, p. 59). Given the danger posed by the feeble-minded to society, Goddard advocated segregation, colonization, and sterilization, saying of the first two that he “would insist that segregation and colonization is not by any means as hopeless a plan as it may seem . . . if such colonies were provided in sufficient number to take care of all the distinctly feeble-minded cases in the community, they would very largely take the place of our present almshouses and prisons, and they would greatly decrease the number in our insane hospitals . . . they would become more or less self-supporting in their institutions, so that the expense of their maintenance would be greatly reduced” (Goddard 1913, p. 105). In effect, Goddard was proposing a vast, state financed and administered warehousing system for the feeble-minded. He estimated that there were 300,000 people in the United States who would need to be incarcerated in all but name to keep from breeding and harming society (Goddard 1913, p. 106). Given that his school for the feeble-minded would no doubt serve as the template for these institutions, Goddard and his associates stood to reap a windfall of nearly unimaginable portions should his proposal have been adopted.

Whereas Goddard was the first man to bring Binet's intelligence test to America (Paul 1995, p. 59), translating it along the way, Terman radically altered it, turning it into the precursor of all modern IQ tests (Gould 1996, p. 205). Working with a cadre of graduate students at Stanford University, Terman normalized Binet's tests to ensure that "average" children would always attain a score of 100 (Gould 1996, p. 207). He also added many other tests, calling the resulting test the Stanford-Binet test, but his most significant innovation was in making the test suitable for mass consumption. While Binet's test had been designed to be given by a trained examiner to one student at a time (Gould 1996, p. 179), Terman's test was designed to be administered en-masse, to large groups of subjects at a time (Gould 1996, p. 205). This difference stemmed from different beliefs about the role of mental testing and the question of who should be tested. Binet was commissioned by the French Ministry of Education to find techniques for detecting students who were behind their classmates and in need of help (Gould 1996, p. 182). He deliberately refused to relate the resulting test score to innate intelligence, since the test was designed strictly as a diagnostic tool (Gould 1996, p. 182).

In contrast, Terman envisioned testing everyone. Specifically, he advocated for the use of IQ tests as the basis for grade promotion, saying that "hitherto the school has had to rely on tests of information because reliable tests of intelligence have not until recently been available; as trained Binet examiners become more plentiful, the information standard will have to give way to the criterion which asks merely that the child shall be able to do the work of the next higher grade" (Terman 1916, chapter 1). Terman also felt that IQ tests could be invaluable for industry, claiming that soon "intelligence tests will become a

recognized and widely used instrument for determining vocational fitness” (Terman 1916, chapter 1). Indeed, in Terman’s hopeful world view, eventually, even resource distribution decisions made by charities would fall under the purview of mental testing as he describes here: “a little psychological research would aid the united charities of any city to direct their expenditures into more profitable channels than would otherwise be possible” (Terman 1916, chapter 1).

Despite the rosy predictions Terman made for his new baby, the Stanford–Binet test, there remained serious problems with both the test itself and the underlying science. Binet’s original test did not suffer from those same problems because it was much more modest in scope and objective. The most serious criticism levelled against Terman is that the Stanford–Binet tested conformity to social expectations rather than abstract reasoning (Gould 1996, p. 206). This is a more general example of the charges of cultural relativism levelled against modern day IQ tests, which all derive from the Stanford–Binet. The cultural relativism problems are particularly acute when the tests are used to measure differences in ability between groups rather than individual differences within the same group. This problem was well demonstrated by Goddard when he used IQ tests to compare the average intelligence of incoming immigrants at Ellis Island, and then pronounced that immigrants from southern and central Europe were clearly defective (Gould 1996, p. 196). Later on, Yerkes demonstrated the same problem after tests of almost 2 million US army soldiers during World War I by making claims about the relative intelligence of different races (Gould 1996, p. 218).

But Terman faced other problems in trying to legitimate his IQ test. In one example, he measured the IQ of a large group of hobos, expecting to find uniformly low score as compared

with measurements he had made for many professions. Unfortunately, he found that with an average score of 89, the hoboes as a group ranked above many blue-collar workers. In order to sweep this messy fact under the rug, Terman took advantage of the fact that the hoboes had a much larger variation than other groups, and sorted his data by the average score of the bottom quartile (Gould 1996, p. 212). The result was to penalize the hoboes, thereby ranking them closer to “where they belonged”. This case serves as a classic example of rearranging data to fit established theories. Unfortunately, Terman sunk still lower. He actually undertook a study to estimate the IQs of past geniuses, as chronicled in the five volume *Genetic Studies of Genius* (Gould 1996, p. 214). These estimates were to be based on a detailed analysis of various historical genius’ biographical entries. Gould claims that much of the differences in estimates “are a methodological artifact of the varying quality of information [they] were able to compile about the childhood of their subjects” (Gould 1996, p. 214). Indeed, those historical figures whose childhoods were not well documented receive rather low IQ estimates (Gould 1996, p. 215).

Strictly speaking, Terman’s agitation for near universal IQ testing did not necessitate his beliefs that IQ was immutable and inheritable, but the two did dovetail nicely. In fact, his immutable hereditarian beliefs were integral to the broader appeal of IQ tests. After all, if IQ could easily change, it would have little value to schools for class placement and companies for job selection. But if IQ was a fundamental immutable trait like eye color, then what controlled its initial setting? Heredity was the obvious answer since it reinforced commonly held beliefs of the time. To say that the rich are rich merely because of an accident of birth was far too close for comfort to collectivist sentiment in an era when labor struggles

reached epic proportions. Terman demonstrated his hereditarian beliefs when he wrote that “practically all of the investigations which have been made of the influence of nature and nurture on mental performance agree in attributing far more to original endowment . . . children from successful and cultured parents test higher than children from wretched and ignorant homes for the simple reason that their heredity is better” (Terman 1916, p. 115). In his autobiography, Terman reasserted this claim, making a bold prediction for the future that “the major differences between children of high and low IQ, and the major differences in the intelligence test scores of certain races, as Negroes and whites, will never be fully accounted for on the environmental hypothesis” (Terman 1930, p. 330).

Terman took his hereditarian beliefs to their logical conclusion by ascribing all social pathology to feeble-mindedness (Gould 1996, p. 211). Thus, most criminals owed their evil nature to their feeble-mindedness. In (Terman 1916, chapter 1), Terman wrote:

“But why do the feeble-minded tend so strongly to become delinquent? Morality depends upon two things: (a) the ability to foresee and to weigh the possible consequences for self and others of different kinds of behavior; and (b) upon the willingness and capacity to exercise self-restraint. In other words, not all criminals are feeble-minded, but all feeble-minded are at least potential criminals. That every feeble-minded woman is a potential prostitute would hardly be disputed by any one. Moral judgment, like business judgment, social judgment, or any other kind of higher thought process, is a function of intelligence. Morality cannot flower and fruit if intelligence remains infantile.”

Having examined the logical conclusion of Terman’s beliefs, we now turn to the logical

origin of those same beliefs. Terman developed a very early interest in the extremes of human ability; describing his college years, he says that “two reports for Lindley, one on ”Degeneracy” and the other on ”The Great-Man Theory” caused me to read almost everything I could find in the library” (Terman 1930). He goes on to explain that his “choice of an experimental study of leadership for a master’s thesis was influenced by the reading [he] did on these reports” (Terman 1930). While pursuing his doctorate, Terman notes that he “became more and more interested in the method of tests” in his “efforts to find a solid footing for research with gifted and defective children” (Terman 1930). For two different reasons, those tests were increasingly mental. First of all, as shown by Clark Wissler, physical tests do not correlate well with academic ability; Wissler opined that “the importance attributed to such measurements by many investigators is not justified” (Wissler 1956, p. 444). A more pressing problem for Terman may have been that he hated working with the equipment used to measure physical responses in experimental psychology. Terman himself claims that “neither [the interest of the subject] nor the gifted teachers I had could make me enjoy working with apparatus . . . three years of work in the laboratory at Indiana and Clark Universities did not enable me to overcome my mechanical ineptness. My dislike of apparatus doubtless had something to do later in turning me to tests and measurements of the kind that make no demands upon mechanical skill” (Terman 1930). He also describes his master’s thesis as “one which gave a chance to do an experimental study without apparatus” (Terman 1930).

Gould portrays Terman as a ruthless hereditarian, desperate to pervert noble Binet’s work for the purpose of proscribing strict limits of achievement for each individual. Indeed,

Gould claims that Terman “relentlessly emphasized limits and their inevitability” (Gould 1996, p. 209), “believing that class boundaries had been set by innate intelligence” (Gould 1996, p. 213). Gould castigated Terman for his plans to use IQ testing to identify the feeble-minded and bring them under the guardianship of the state, and contrasted that vision with Binet’s “desire to segregate and help” (Gould 1996, p. 209). But was Terman really the monster Gould tried to portray him as, at least compared with Binet? Terman did write everything that Gould ascribes to him, but he also wrote more that Gould does not cite. The most serious example is Terman’s view on education reform. Terman believed that both bright and dull children suffer greatly when all are forced to learn at the same rate. According to him, the traditional “rule of brawn” schooling system that groups children of radically different abilities together based on chronological age failed everyone. Terman’s work with children and IQ testing convinced him that “among those classed as normal, vast individual differences have been found to exist in original mental endowment, differences which affect profoundly the capacity to profit from school instruction” (Terman 1916, chapter 1). Except for his innatist beliefs, Terman sounds remarkably progressive. And he did occasionally back away from his innatist convictions; one example is when he wrote that “every child who fails in his school work should be given a mental examination . . . it is necessary to determine whether a given child is unsuccessful in school because of poor native ability, or because of poor instruction, lack of interest, or some other removable cause” (Terman 1916, chapter 1).

Unfortunately, the issues discussed above are all too relevant today. Much of the same fallacies committed by Terman and Goddard have been committed anew by the modern day evangelists for innate, immutable, and hereditarian IQ. The quintessential tome for

neo-hereditereans, *The Bell Curve* by Charles Murray and Richard Herrnstein, has had a substantial and enduring impact on social policy (Gould 1996, p. 367). Beyond that, the genomics revolution offers us a return to the past where genetecists have become the new eugenicists. Interestingly, the new heriditereans seem more cruel, at least in theory, than their forebears Goddard and Terman. Their closest idealogical cousins appear to be modern neo-conservatives who espouse a faith in heriditerean IQ with *The Bell Curve* as their bible. But while Goddard and Terman believed that society cannot in good conscience hold “inferior” people responsible for their actions since they lacked sufficient self-control, their neo-conservative cousins are far more punative, advocating strict “three strikes and you’re out” laws and harsh mandatory minimum sentencing.

Granted, under either system of belief, the result is social control on an unprecedented scale by means of a vast incarceration network, but the conceptual difference is still significant. Modern neo-conservatives advocate harsh punishments for those suffering from the same faults to which they attribute to genetic endowment. In contrast, Goddard and Terman both believed that the feebleminded should be segregated from the population for their own good, but not as a punitive method. Terman described this best when wrote that “when we have learned the lessons which intelligence tests have to teach, we shall no longer blame mentally defective workmen for their industrial inefficiency, punish weak-minded children because of their inability to learn, or imprison and hang mentally defective criminals because they lacked the intelligence to appreciate the ordinary codes of social conduct” (Terman 1916, chapter 1). The stark contrast between Goddard’s fantasy of a nation dotted with thousands of replica’s of his Vineland School for the Feebleminded and modern neo-

conservatives' largely realized vision of vast multitudes teeming under virtual slavery in the service of a relentless prison-industrial complex is chilling.

## References

Henry Herbert Goddard. *The Kallikak Family: A Study in the Heredity of Feeble-Mindedness*. Macmillan, 1913.

Stephen Jay Gould. *The Mismeasure of Man*. W. W. Norton, 1996.

Diane B. Paul. *Controlling Human Heredity*. Humanities Press International, 1995.

Lewis M. Terman. *The Measurement of Intelligence*. Houghton Mifflin, 1916.

Lewis M. Terman. Autobiography of Lewis M. Terman. In Carl Murchison, editor, *History of Psychology in Autobiography*, volume 2, pages 297–331. Clark University Press, 1930.

Clark Wissler. On the inadequacy of mental tests. In Richard J. Herrnstein and Edwin G. Boring, editors, *A Source Book in the History of Psychology*, pages 442–445. Harvard University Press, 1956.