*Article*

# An Evaluation (and Vindication?) of Lewis Terman: What the Father of Gifted Education Can Teach the 21st Century

Gifted Child Quarterly
1–19

DOI: 10.1177/0016986218799433
journals.sagepub.com/home/gcq
SAGE

**Russell T. Warne**[1]

**Abstract**

Lewis Terman is widely seen as the "father of gifted education," yet his work is controversial. Terman's "mixed legacy" includes the pioneering work in the creation of intelligence tests, the first large-scale longitudinal study, and the earliest discussions of gifted identification, curriculum, ability grouping, acceleration, and more. However, since the 1950s, Terman has been viewed as a sloppy thinker at best and a racist, sexist, and/or classist at worst. This article explores the most common criticisms of Terman's legacy: an overemphasis on IQ, support for the meritocracy, and emphasizing genetic explanations for the origin of intelligence differences over environmental ones. Each of these criticisms is justified to some extent by the historical record, and each is relevant today. Frequently overlooked, however, is Terman's willingness to form a strong opinion based on weak data. The article concludes with a discussion of the important lessons that Terman's work has for modern educators and psychologists, including his contributions to psychometrics and gifted education, his willingness to modify his opinions in the face of new evidence, and his inventiveness and inclination to experiment. Terman's legacy is complex, but one that provides insights that can enrich modern researchers and practitioners in these areas.

**Keywords**

philosophical/theoretical, history of gifted education, Lewis Terman

In some circles, it is fashionable to bash Lewis Terman. Sometimes his work is seen in almost malevolent terms, with Terman as a champion of an IQ-stratified social order, with White men at the top of the hierarchy (e.g., Mercer, 1979; Minton, 1988; Winkler & Jolly, 2014). Terman's worldview, in these accounts, was that ability and its consequences, social status and wealth, were determined by heredity. Racist and prejudiced beliefs and support for the eugenics movement followed logically from such a view (Leslie, 2000). I encountered this viewpoint firsthand when I submitted an analysis of data from Terman's longitudinal study to a scholarly journal. In response, one blind peer reviewer for the journal *Exceptional Children* stated, "I'm not sure it is useful to continue looking at this data [sic] that was really grounded in racist attitudes toward ability that the gifted community needs to abandon . . ." (The study was later published as Warne & Liu, 2017.)

However, this perspective is an oversimplification. This article provides an overview of the most frequent criticisms of Lewis Terman's work, with a special focus on his famous longitudinal study, the Genetic Studies of Genius. The overview includes a discussion of the reasons behind these criticisms and provide evidence from Terman's own writings about the degree to which these criticisms are justified. Following is a section on a more serious criticism of Terman's work, which most prior writers have overlooked: his willingness to form a strong opinion based on weak data. The final section of this article is a discussion of the status of Terman's work today and the lessons that it can still provide for modern scholars and practitioners in education and psychology.

## Frequent Criticisms of Terman's Work

During the course of his career, Lewis Terman was—at one time or another—president of the American Psychological Association, a prominent Stanford University psychology professor, a public intellectual figure, the world's foremost intelligence test creator, and the principal investigator of a famous longitudinal study. His work in giftedness and psychometrics is well-known to 21st-century psychologists, but in his 50-year scholarly career, he also produced studies on topics as diverse as sexual behavior, suicide, and personality assessment. With such a long and productive career, it would be impractical in a single article to discuss every flaw that

[1]Utah Valley University, Orem, UT, USA

**Corresponding Author:**
Russell T. Warne, Department of Behavioral Science, Utah Valley University, 800 West University Parkway MC 115, Orem, UT 84058, USA.
Email: rwarne@uvu.edu

critics have mentioned about his work. However, his most prominent work—regarding giftedness, intelligence, and intelligence testing—has most frequently attracted the following criticisms: (a) an overemphasis on IQ and intelligence; (b) a preference for a meritocracy that undervalued non-Whites, women, and low-income individuals; and (c) a blind preference for genetic explanations of individuals' differences over environmental explanations.

## *Overemphasis on IQ*

One of the most common criticisms of Terman is his strong emphasis on intelligence as a concept and IQ as a measure of it. According to Terman (1921a), "An individual is intelligent in proportion as he is able to carry on abstract thinking" (p. 128, emphasis removed from original). The criticism that Terman viewed intelligence as being of paramount importance is not unjustified. For example, in an article written for the popular press, he stated, "There is nothing about an individual as important as his IQ, except possibly his morals . . ." (Terman, 1922e, p. 657). Likewise, in Volume IV of the Genetic Studies of Genius, he wrote, ". . . the IQ level is one of the most important facts that can be learned about any child" (Terman & Oden, 1947, p. 358; see also Terman, 1911, p. 203; 1916, p. 20). Given sentiments like these, it is easy to draw the conclusion that Terman had a single-minded devotion to intelligence and IQ. Critics of Terman who question this emphasis on IQ (e.g., Morelock & Feldman, 1997; Vialle, 1994) are correct that Terman valued intelligence to a much greater degree than most educators do today.

Yet this viewpoint is an oversimplification. Unquestionably, Terman did think that intelligence was highly important. Indeed, much of his research and writing would not make sense without this basic underlying assumption. But for Terman, intelligence/IQ was not the be-all end-all of human psychology. In his book describing the 1916 Stanford–Binet, he stated,

> The scale does not pretend to measure the entire mentality of the subject, but only *general intelligence*. There is no pretense of testing the emotions or the will . . . The scale was not designed as a tool for the analysis of those emotional or volitional aberrations which are concerned in such mental disorders as hysteria, insanity, etc. (Terman, 1916, p. 48).

And later in the same work,

> A second misunderstanding can be avoided by remembering that the Binet scale does not pretend to bring to light the idiosyncrasies of special talent, but only to measure the general level of intelligence. It cannot be used for the discovery of exceptional ability in drawing, painting, music, mathematics, oratory, salesmanship, etc., because no effort is made to explore the processes underlying these abilities. It can, therefore, never serve as a detailed chart for the vocational guidance of children . . . . (p. 49)

This intuition—that a global IQ score was only suitable for learning about the general, global intellectual capacity of examinees—is supported by modern research. Likewise, the claim that "special talent" is difficult to detect in an overall IQ score is a mainstream viewpoint today (e.g., Gottfredson, 1997a; Warne, 2016a).

Indeed, Terman frequently warned that intelligence tests—including his own Stanford–Binet—could not measure every mental ability (e.g., Terman, 1921a). For example, in a discussion of sex differences, Terman stated not only that there were no global IQ differences between males and females but also that there could be relevant mean differences in "*special* intellectual functions or in emotional and vocational traits" (Terman, 1922e, p. 660) that would not be apparent in a global IQ score. Moreover, Terman was aware of the imperfections inherent in psychometric test scores: "An obtained IQ, as I have many times pointed out, should never be taken as a final verdict, but only as a point of departure for further investigation of a subject" (Terman, 1940, p. 466).

In much of his applied work, Terman recognized the importance of psychological variables besides intelligence (e.g., Terman, 1919, chap. 12). This is apparent in the Genetic Studies of Genius, where the first volume (Terman, 1926) includes data on the gifted individuals' physical health, play and academic interests, personality traits, social skills, and more. Indeed, in later volumes of Genetic Studies of Genius, the focus is less on the subjects' IQ scores and more on other variables that would relate to adult success in college, work, and their personal lives. Many of these variables were probably collected to fulfill Terman's goal of disproving the then-prevalent folk belief of "early ripe, early rot," meaning that bright children would grow up to be dysfunctional adults (Terman & Oden, 1947). Today many modern educators would agree with Terman when he wrote, "The over-enthusiastic will gradually learn that not even the universal use of intelligence tests will bring us to an educational millennium. The child is more than intelligence, and education is more than the cultivation of intellectual faculties" (Terman, 1922c, p. 3).

Regardless of the exact level of importance Terman ascribed to intelligence, he always affirmed its relevance. Terman's emphasis on IQ has been borne out in modern research. In many studies, global IQ score is a powerful predictor of academic achievement, and this positive correlation does not disappear when controlling for parental socioeconomic status or student demographics (e.g., Cucina, Peyton, Su, & Byle, 2016; Jensen, 1998; Roth et al., 2015). Terman's opinions would not be out of place today, as modern theorists also place great importance on intelligence. For example, Plomin (2001) stated that intelligence ". . . predicts important social outcomes such as educational and occupational levels far better than any other trait . . ." (pp. 137-138). Indeed, intelligence positively correlates with many more beneficial life outcomes than Terman likely suspected, including

income (Zagorsky, 2007), job performance (Schmidt & Hunter, 2004), life expectancy (Deary, Whiteman, Starr, Whalley, & Fox, 2004), and many others (Strenze, 2015). Ironically, many gifted education scholars and practitioners have diverged from Terman's perspective on the importance of intelligence, which has hampered the field's efforts to make connections with the wider psychological community (Warne, 2016a).

## Support for the Meritocracy

Terman was an unashamed advocate for a meritocratic society where economic success, power, and the other benefits of life would be distributed to individuals according to merit. For Terman, "merit" included intelligence, though he never defined the term. Many of Terman's critics have seen his support for a meritocracy as being highly problematic because they see it as elitist, antidemocratic, or misguided (e.g., Gould, 1981; Jolly, 2008b; Leslie, 2000; Minton, 1988). Terman's support for the meritocracy becomes problematic in light of the lower mean performance of low-income individuals and most non-White groups on intelligence tests. Taken to its logical result, Terman's support for an IQ-based meritocracy appears to justify existing economic and power inequalities in society.

Again, there is a good reason to see Terman as supporting an IQ-based meritocracy: He said as much himself. Terman openly advocated using intelligence tests to help guide general vocational placement, with higher scoring children being suggested for more prestigious jobs and lower scoring children being suggested for remedial factory work (Terman, 1916, p. 17). To Terman, distributing jobs or educational opportunities according to merit was a far better alternative than using nepotism, social connections, or random chance. Indeed, he reported case studies of individuals who were given jobs through political patronage or nepotism (Terman, 1919, pp. 271-274) or admitted to college because of a wealthy parent's influence (Terman, 1922a), with negative consequences in both cases.

More important, Terman did not see a meritocracy as perpetuating inequalities. Rather, it was a way of opening doors to worthy people (often with high IQ scores) from unprivileged backgrounds (Terman, 1922a). Early in his research, Terman saw that teachers were poor judges of intellectual capacity (Terman, 1911, 1916), and he worried about bright children languishing unnoticed without an opportunity to develop their intellectual gifts. Terman saw that teachers and other individuals were overlooking children who would benefit from gifted education interventions (e.g., acceleration or ability grouping). If using an intelligence test to select people for educational opportunities resulted in inequalities, Terman viewed this as merely a reflection of preexisting differences in ability. His views were mainstream among differential psychologists at the time (Deary et al., 2004). In fact, it was this meritocratic willingness to reduce barriers to opportunity for qualified individuals that was one of the impetuses for the creation of the SAT as a college admissions test in the 1920s (Calvin, 2000).

With regard to employment testing, one vocal critic argued that "Terman virtually closed professions of prestige and monetary reward to people with IQ below 100 . . . and argued that 'substantial success' probably required an IQ above 115 or 120" (Gould, 1981, p. 182). The truth is more benign. Terman's opinions regarding the meritocracy in employment were the result of observing the average IQ for workers in various occupations (e.g., Terman, 1919). One of his sources of this information was likely the results of the Army Alpha and Army Beta examinations in World War I, which showed a moderately strong positive relationship between intelligence and job prestige (Yoakum & Yerkes, 1920, pp. 25, 198). The same data also showed a positive correlation between supervisors' ratings of job performance and soldiers' IQ scores (Yoakum & Yerkes, 1920, p. 25). Terman astutely observed that there was a wide range of IQ levels within many jobs, but that more prestigious jobs seemed to naturally have a minimum IQ level needed for entry and/or job success (Terman, 1919), a finding supported by modern research (e.g., Gottfredson, 1997b). In other words, Terman did not try to invent an intelligence-based meritocracy in the employment sphere. Rather, he merely reported on the preexisting trend of higher IQ individuals in more prestigious, better paying jobs.

Terman believed that a meritocracy was an inevitable result of (a) individual differences, and (b) a society that allowed those differences to be expressed (Boring, 1959). As a result, he believed that meritocracy ". . . is least noticeable among the newly arrived immigrants of an oppressed race. Freed from oppression these also rapidly form into a social hierarchy based largely upon native capacity, the intelligent, the average, and the incompetent finding their own levels" (Terman, 1922e, p. 658). For Terman, a society without a meritocracy was an oppressed, unfree society.[1] Additionally, Terman viewed the disconnect between women's intelligence scores—which were on average equal to men's—and their social and economic position as evidence of societal oppression of women (Terman, 1922e), though he did little to attempt to mitigate women's oppression (Rogers, 1999).

Terman's views on the meritocracy are still relevant today. Concerns about the societal implications of an IQ-based meritocracy are common among intelligence researchers. Most notably, Herrnstein (1971) observed that some jobs are more complex than others, which makes the former open to a smaller proportion of the workforce. The natural economic forces of supply and demand mean that jobs with a higher minimum IQ for entry are often better paying jobs in a capitalistic society (see also Herrnstein, 1973; Herrnstein & Murray, 1996). The ingredients that Terman saw as necessary for an IQ-based meritocracy continued to exist well past his death and into the modern era (Gottfredson, 1997b).

Nevertheless, Terman's support for a meritocracy went beyond the views of modern researchers. For example, one of his biographers reported that Terman gave guest lectures in the mid-1920s about intelligence in a general education class at Stanford University, one component of which discussed the desirability of concentrating power among high-IQ individuals. This concentration of power included an argument in favor of denying suffrage to individuals with low intelligence and implementing social engineering policies that would benefit society as a whole (Minton, 1988). The existence of individual differences in any important trait has the potential to create a tension between a meritocracy and democracy (Gottfredson, 2005a). Terman—being part of the Progressive Era zeitgeist that valued expertise and societal planning—believed that investing power into the hands of an elite identified through intelligence tests could resolve this tension (Chapman, 1988). Much to the surprise of a 21st-century reader, Terman did not see this as undermining democracy, but as a bulwark against the societal chaos that could come from the poor decisions of large numbers of individuals with low intelligence, a common belief among social progressives at the time (Chapman, 1988; Kevles, 1995). Still, in this regard, Terman's critics are correct that his support for an IQ meritocracy was elitist and antidemocratic.

But Terman was not content to merely deny individuals with low intelligence the right to vote. He was a proponent of the early-20th-century eugenics movement (Leslie, 2000; Minton, 1988; Seagoe, 1975). Eugenics was a social movement that originated in the late 19th century by Sir Francis Galton, who saw it as a scientific means to guide human evolution to benefit the species (Gillham, 2001; Kevles, 1995; Koch, 2006). Eugenics took many forms as it spread internationally (Broberg & Roll-Hansen, 2005; Kevles, 1995). Galton, for example, was mostly concerned about the higher birthrate among lower classes in British society and advocated policies that would encourage better educated people to have more children (Gillham, 2001). This was the prevailing shape of eugenics in Great Britain and was a popular attitude among British scientists (Kevles, 1995), such as Galton's protégé Karl Pearson (1903). In the United States, eugenics often took the form of state-level forced sterilization laws that were part of a wider Progressive Era concern about public health (Stern, 2005). Terman supported these laws, including the one in his own state of California (Minton, 1988; Seagoe, 1975), and he was a member of multiple eugenics societies.

Terman viewed his work on intelligence as being an important component of the eugenics movement. This is unsurprising because the birth of the intelligence testing movement had been intertwined with eugenics. The founder of both was Galton, and the reason Galton attempted (though failed because of an emphasis on physiology and simple tasks) to measure intelligence was because he believed that high intelligence would be a useful trait to select for in humans (Fancher, 1985; Gillham, 2001). When he introduced the Stanford–Binet intelligence test to the world, Terman stated,

> It is safe to predict that in the near future intelligence tests will bring tens of thousands of these high-grade defectives under the surveillance and protection of society. This will ultimately result in curtailing the reproduction of feeble-mindedness and in the elimination of an enormous amount of crime, pauperism, and industrial inefficiency. It is hardly necessary to emphasize that the high-grade cases, of the type now so frequently overlooked, are precisely the ones whose guardianship it is most important for the State to assume. (Terman, 1916, pp. 6-7)

Terman would reiterate his view of intelligence tests as serving the eugenics movement in later years (Terman, 1922b, 1922e), and he even called intelligence testing "the beacon light of the eugenics movement" (Terman, 1924b, p. 106) in his presidential address to the American Psychological Association (see also Terman, 1924c, p. 340).[2]

Though this advocacy for eugenics is abhorrent today (and rightly so), it is also important to recognize the context for Terman's advocacy and his change in opinions later in life. In the early 20th century, eugenics was popular among educated individuals in Europe and the United States (Broberg & Roll-Hansen, 2005; Kevles, 1995; Koch, 2006). In 1912, the First International Congress of Eugenics (held in the United Kingdom) could count among its attendees Lord Alverstone (Lord Chief Justice at the time), two bishops of the Church of England, Winston Churchill, Alexander Graham Bell, the Lord Mayor of London, and other medical, political, and social luminaries from several countries (Gillham, 2001, pp. 345-346). Other eugenicists included Sigmund Freud (Richardson, 2011), George Bernard Shaw, Margaret Sanger, Oliver Wendell Holmes Jr., American presidents Theodore Roosevelt and Calvin Coolidge (Leslie, 2000), H. G. Wells, John Maynard Keynes, John D. Rockefeller Jr., George Eastman, Emma Goldman, and many then-prominent individuals whose names are less recognizable today (Kevles, 1995). Scientists who advocated the philosophy included statisticians Karl Pearson (E. S. Pearson, 1936) and Sir Ronald Fisher, many geneticists and biologists (Kevles, 1995), and Terman's colleagues Henry Goddard, Robert Yerkes, Charles Spearman, (Fancher, 1985), Carl Brigham, G. Stanley Hall (Kevles, 1995), and Truman Kelley, who also contributed statistical analyses to the Genetic Studies of Genius (Bellhouse, 2009). These lists are not given to excuse Terman's advocacy of eugenics but rather to help readers understand that Terman was part of a very large, respected international social movement that viewed itself as applying Darwin's ideas to the "betterment" of future generations.

Although not a component of every eugenicist's thought, many prominent individuals in the early 20th century viewed the decision of who should (or should not) pass on their genes to future generations in racial terms (Kevles, 1995). Terman was among these. In a frequently quoted

passage, he discussed two "Portuguese" brothers who were the sons of a "skilled laborer" with IQ scores of 77 and 78 and stated that they

> . . . represent the level of intelligence which is very, very common among Spanish-Indian and Mexican families of the Southwest and also among negroes. Their dullness seems to be racial, or at least inherent in the family stocks from which they come. The fact that one meets this type with such extraordinary frequency among Indians, Mexicans, and negroes suggests quite forcibly that the whole question of racial differences in mental traits will have to be taken up anew and by experimental methods. The writer predicts that when this is done there will be discovered enormously significant racial differences in general intelligence, differences which cannot be wiped out by any scheme of mental culture. (Terman, 1916, pp. 91-92)

In Terman's day, East Asians in America scored as well as Whites (or slightly better) on intelligence tests, who in turn outscored Hispanics and African Americans (Goodenough, 1926; Terman, 1916; Yoakum & Yerkes, 1920). Terman recognized this, along with the fact that there was a large degree of overlap among these racial and ethnic groups' distributions (Terman, 1922e). The fact that these mean differences would result in unequal group-level outcomes was not a major concern for him. If taken to its logical conclusion, Terman's advocacy for an IQ meritocracy would mean that groups that had lower scores would be less likely to be selected as members of the elite or for favorable jobs. His advocacy for eugenics would also mean that these groups would also disproportionately be victims of forced sterilization—as indeed they were in California (Stern, 2005).

On the other hand, it is important to recognize what Terman's eugenics was not. Terman never advocated for genocide or Nazi-style concentration camps. He seems to have advocated for forced sterilization on a case-by-case basis and for general policies that would encourage bright people to have more children. Also on an individual basis, Terman favored setting ". . . a minimum mental standard for our immigrants from every source" (Terman, 1922e, p. 660) but did not advocate for systematically closing immigration to any racial group or nationality.

Terman's enthusiasm for eugenics did not continue past the mid-1930s. He resigned his membership from eugenics organizations by 1935 and did not support them privately beyond 1938 (Minton, 1988). In his public writings, Terman continued to discuss eugenic ideas in the 1920s, but his interest waned as the decade wore on. As an illustration, the first volume of Genetic Studies of Genius devotes three chapters to the eugenic concerns of racial/national heritage, intellectually superior relatives, and parental fertility (Terman, 1926, chaps. 4-6). The follow-up report in 1930 mentions eugenics only twice in passing (Burks, Jensen, & Terman, 1930, pp. 239, 482), even though he and his coauthors reported plenty of relevant data. The last volume published before his death does give fertility statistics and other data that would be relevant for a eugenic argument, yet it completely lacks any eugenic commentary (Terman & Oden, 1947). I cannot find any mention of eugenic ideas in his writings after 1930, and after 1925 mentions are sparse. In his 1932 autobiography, Terman makes no mention of eugenics. Likewise, the manual for the 1937 Stanford–Binet test (Terman & Merrill, 1937) lacks any references to eugenics, a major shift compared with the frequent mentions in the 1916 version's manual regarding the practical uses of intelligence testing to further the eugenics agenda (Terman, 1916). It is unclear why Terman's interest turned away from eugenics, but historians of science (e.g., Broberg & Tydén, 2005; Glass, 1986; Hansen, 2005; Kevles, 1995) have shown that the social policies of the eugenics movement began to fall out of favor among scientists at this time. Terman may have been aware of the empirical data that failed to support the necessary assumptions underpinning the policies of early-20th-century eugenics and—like many other scholars and lay people of the 1920s and 1930s—may ultimately have found at least some of the social and legal activities of the eugenicists unsupportable.

## *Nature Over Nurture*

Another frequent criticism of Terman is his strong adherence to genetic causes of individual differences over environmental causes (e.g., Leslie, 2000). For example, Jolly (2008b) stated that "Terman's unyielding belief in science, perhaps led him to ignore obvious discrepancies in his findings. Terman held steadfast to the idea of heritability rather than environment (or a combination of the two) to explain superior intelligence or giftedness" (pp. 32-33).

Just as with the previous criticisms, Terman's own words provide ample ammunition to those who believe that Terman overemphasized the importance of genetic influences on intelligence. When discussing the correlation between social class and IQ scores, he wrote, "That the children of the superior social classes make a better showing in the tests is probably due, for the most part, to a superiority in original endowment" (Terman, 1916, p. 72). Terman then provided five lines of "supporting evidence," only one of which (studying within-family differences in IQ scores) could support his argument—and even that evidence is not presented with enough detail to support his claim. A few years later, Terman (1922e) wrote,

> Intelligence is chiefly a matter of native endowment. . . . In fact, the mathematical coefficient of family resemblances in mental traits, particularly intelligence, has been found to be almost exactly the same as for such physical traits as height, weight, cephalic index, etc. (p. 659)

Thus, critics have a basis on which to state that Terman had "remarkable insensitivity to the influence of environment" (Gould, 1981, p. 190; see also Lagemann, 2000, p. 90).

Yet this, too, is an oversimplification. Throughout Terman's career, he recognized that environmental characteristics could have an impact on an individual's intelligence test score. Indeed, in the second article he ever published, Terman wrote that ". . . as training counts for more, heredity counts for less" (Terman, 1905, p. 152; see also Terman & Childs, 1912a, p. 62). And in the same book in which he ascribed the positive correlation between socioeconomic status and intelligence to heredity, he stated several times that environmental variables can influence children's IQ scores (Terman, 1916, pp. 12, 19, 116, 135, 165, 174, 184, 264). He also recognized the detrimental impact that an extremely disadvantageous (e.g., neglectful or abusive) environment could have on a child's intellectual development and IQ score (Terman, 1919, 1928a) and the fact that schooling could increase a child's IQ score (Terman, 1922e).

Although Terman always held that environment could influence intelligence, his faith in the importance of heredity faded as his career progressed (Gould, 1981; Hilgard, 1957; Minton, 1988; Seagoe, 1975). By the time he gave his American Psychological Association presidential address, he stated, "Whether these 'chronic' traits [like intelligence] reflect primarily the influence of endowment or of environment, is a question to which no certain answer can at present be given" (Terman, 1924b, p. 102), though he did in the next sentence state that if a set of children have "the ordinary advantages offered by public school and by other common social contacts" that heredity mattered more (p. 102). It is apparent that by the 1920s Terman saw the importance of favorable environments in providing a setting in which individual genetic differences could manifest themselves (see also Terman, 1928a). He even acknowledged the possibility that an environmental intervention could raise intelligence (Terman, 1926, pp. 635-636).

By the 1930s, even this level of confidence in genetics had decreased (Minton, 1988, pp. 187, 199; Seagoe, 1975, p. 122). In the manual for the revision of the Stanford–Binet, Terman recognized that positive correlations between IQ scores and socioeconomic status did not provide strong evidence about the importance of genetics or environment in determining intelligence (Terman & Merrill, 1937, p. 48). Even Gould—one of Terman's harshest critics—admitted this (Gould, 1981, pp. 191-192). Terman's views on environmental influences on intelligence were modified further a few years later when, with regard to new evidence on the topic, he stated, "One who has followed the age-old controversy between hereditarians and environmentalists will hardly be surprised to find that the new evidence does not all point in one direction" (Terman, 1940, p. 460, see also, p. 466; Terman & Oden, 1947, pp. 15, 17-18). Still, even at this late point, he did believe that methodologically stronger studies showed smaller environmental effects, and he was not shy about criticizing studies that showed strong environmental impacts on intelligence (Terman, 1940).

Additional testimony to Terman's recognition of the importance of environment on cognitive performance is apparent in several aspects of his work. Much of his research in the early 1910s concerned improving school environments and student health to produce better educational outcomes (Seagoe, 1975). Likewise, Terman advocated for educational interventions for gifted children, especially ability grouping and academic acceleration (Terman, 1922c, 1954, 1958). The value that he placed on educational interventions is also apparent in his editorship of a book that includes some of the first descriptions of specialized curricula for gifted children (Stedman, 1924). Thus, the arguments that Terman did not care about interventions (e.g., Callahan, 1996) or ignored environmental influences (Jolly, 2008b) are not apparent in the evidence.

It is probably misleading to consider the title of Terman's landmark Genetic Studies of Genius as signaling his belief in nature over nurture in the development of giftedness. In fact, a common meaning of the word "genetic" at the time was "of or relating to origin or development," (in the words of the *Oxford English Dictionary*). Stanley (1974) recognized this, stating, "In retrospect the adjective 'genetic' . . . in that title may seem unfortunate. By 'genetic,' however, he probably meant 'longitudinal' more than 'Mendelian'" (p. 5; see also Lubinski, 2016, p. 935). Indeed, Terman had used the word "genetic" in this sense previously (Terman, 1905, p. 148), as had his colleague Yerkes (1917, pp. 111, 113) at a time when the two men worked closely together on the development of the Army Alpha and Army Beta tests.

Terman never denied the importance of heredity, though. This may be because he *did* have data supporting a hereditarian view of the origins of intelligence. Terman was aware of Galton's study *Hereditary Genius*, which demonstrated that in British history "genius" (today one would likely prefer the term *eminence*) disproportionately ran in families, with the multiple male "geniuses" in such families being more likely to be closely related to one another than to be more distantly related (see Gillham, 2001, for a summary). Terman's (1926) own data from the first volume of Genetic Studies of Genius supported Galton's findings. Additionally, data from the creation and norming of the Stanford–Binet showed that the correlation between IQ scores and socioeconomic status decreased in successively older samples of children—the exact opposite of what one would expect if environment were a powerful influence on intelligence (Terman et al., 1915, p. 561). In fact, modern behavioral genetics studies often replicate this latter finding (with far more sophisticated designs), and it is one reason why many modern researchers believe that genetics is an important influence on intelligence in adolescence and adulthood (Bouchard, 2014; Plomin, DeFries, Knopik, & Neiderhiser, 2016). During Terman's career, the pendulum of opinion swung from a mostly genetic explanation of the origins of intelligence to favoring mostly environmental

explanations. In the decades after his death, the pendulum has swung back to valuing genetic explanations for within-group differences in intelligence (Gottfredson, 1997a), though neither extreme (i.e., purely environmental explanations or purely genetic explanations) will likely ever be a mainstream position again. Indeed, some experts may agree with Terman's opinions from the 1910s more than with the opinions he held at the end of his life (e.g., Gottfredson, 2005b). Regardless of where the consensus of modern expert opinion lies, the caricature of Terman as a blind follower of a genetically dominated origin of intelligence is inaccurate. Throughout his entire career, he clearly recognized the value of the environment in influencing intelligence, and this recognition increased as he aged.

## But Was He a Racist?

Racial bigotry can take many forms, and the fact that Terman did not advocate for policies that were a priori discriminatory does not demonstrate that he was free of racial bias; indeed, he believed that constitutional rights of racial minorities should be protected, and he lamented racial discrimination (Hilgard, 1957). With regard to intelligence, Terman believed fervently in mean racial differences in intelligence, with most racial and ethnic minority groups scoring, on average, lower than the White mean.

Still, Terman was public about exceptions to this general trend, and he could distinguish between individual intelligence test performance and the mean score of a person's demographic group. He reported in great detail cases of minority children who scored high on intelligence tests (e.g., Burks et al., 1930, pp. 286-297; Terman, 1916, pp. 117-118; Terman, 1926, pp. 107-108). As a result, when Terman discovered low-scoring minority examinees he stated so clearly (e.g., Terman, 1916, pp. 83, 117), and concluded that, "No race or nationality has any monopoly on brains" (Terman & Oden, 1947, p. 14). This ability to distinguish between individuals and the group they belonged to was characteristic of Terman in public and in private throughout his career. A notable example of this was his decision to write a letter vouching for the loyalty of a Japanese American family that had multiple children in his longitudinal study. This influence may have kept the family out of the internment camps that many Americans of Japanese descent were confined to during World War II (Shurkin, 1992).

Terman also believed that his research could lead to a more just world for nongifted individuals of all racial groups. He served as an expert witness in the defense of a Hispanic man who was accused of sexual assault and murder of a young child. Terman's demonstration that the defendant had a mental age of 7½ likely saved the man from the death penalty (Terman, 1918b), though the laws at the time did not recognize "feeblemindedness" as a valid insanity defense, so the defendant could not be found not guilty by reason of insanity (Dahlstrom, 1985). Later that year, the law in California was amended to allow an insanity defense for similar cases in the future (Terman, 1918b). More central to his typical work was his advocacy for universal testing so that struggling students could also get the educational adjustments they needed (e.g., Terman, 1922c, 1922d). In this manner, he followed in the footsteps of Binet, whose test was designed to identify struggling students who should receive special education services. Indeed, there is evidence that California schools in the 1910s and 1920s that used intelligence and academic tests for homogeneous grouping and adjusted the curriculum accordingly experienced a reduced dropout rate as the curriculum became more accessible to nonelite students (Chapman, 1988).

## Terman's Major Scientific Sin: Going Beyond the Data

For all the previously mentioned flaws, one criticism of Terman seems absent from the previous literature: Terman's willingness to form a forceful opinion when the data were not strong enough to support this degree of confidence. For example, Terman stated, "All the available facts that science has to offer support the Galtonian theory that mental abilities are chiefly a matter of original endowment" (Terman, 1922e, p. 659). In response to sentiments like this, Minton (1988) stated, "Terman never provided unequivocal evidence that IQs reflected native ability. Based on his weddedness to Hall and Galton's biological determinism, he simply assumed that IQs were genetically determined" (p. 200). Indeed, no psychologist at the time had data that could separate the influences of heredity and environment on intelligence test scores, so anyone who had a strong opinion about the topic—including Terman—lacked the empirical evidence to support their views. The first behavioral genetics studies that could estimate the impacts of genetics and environment would be published a few years later (Burks, 1928/1973; Hildreth, 1925; Tallman, 1928; Wingfield, 1928).[3] In 1922, there were correlational data that supported Galton's and Terman's views, but these same data just as easily supported theories of purely environmental causes of interindividual differences for intelligence. Terman seems to have downplayed this possibility in the first half of his career when discussing the relative importance of nature and nurture (e.g., Terman, 1906, p. 372, 1922e, 1928a; Terman et al., 1915). He had a penchant for interpreting correlational data as being causal in nature—a tendency that extended beyond his research on intelligence (Hollingworth, 1939).

Studies published during Terman's lifetime could only provide information about the degree of heritability of intelligence within White American populations (e.g., Burks, 1928/1973). He was on even shakier ground when discussing the causes of mean racial/ethnic group differences in intelligence. An example of this is clear in the first volume of the Genetic Studies of Genius when he stated,

> How much of this inferiority [in intelligence test scores of Hispanic groups] is due to the language handicap and to other environmental factors it is impossible to say, but the relatively good showing made by certain other immigrant groups similarly handicapped would suggest that the true causes lie deeper than environment. (Terman, 1926, p. 57)

One would never know from this confident statement and others like it (e.g., Terman, 1916, pp. 91-92) that the first American transracial adoption studies that could provide evidence informing the influence of heredity on interracial group IQ differences were decades away (Moore, 1986; Scarr & Weinberg, 1976; Tizard, 1974; Weinberg, Scarr, & Waldman, 1992). Even now, Terman's early career claims of a predominantly genetic cause of intergroup mean differences are not supported by data.

Later in life, Terman ameliorated these views; in the 1930s, he stated, "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, 1932, p. 329). Although this is a less strong statement than his views on the topic in the 1910s and 1920s, the data still simply did not exist at the time to rule out the possibility of completely environmental causes of interracial mean differences in IQ. Only long after his death would circumstantial evidence begin to emerge that group mean differences among American[4] racial/ethnic groups may not have completely environmental causes.[5]

Less controversial—but no less problematic—was Terman's use of cross-sectional data to make longitudinal predictions. This is especially noticeable when he predicted children's future vocational complexity or prestige based on the average IQ of adults working in those occupations at the time. For example, he wrote,

> At every step in the child's progress the school should take account of his vocational possibilities. Preliminary investigations indicate that an IQ below 70 rarely permits anything better than unskilled labor; that the range from 70 to 80 is preeminently that of semi-skilled labor, from 80 to 100 that of the skilled or ordinary clerical labor, from 100 to 110 or 115 that of semi-professional pursuits; and that above all these are the grades of intelligence which permit one to enter the professions or the larger fields of business. (Terman, 1922c, p. 27)

Terman was likely basing these projections on data from American men drafted to serve in World War I who took the Army Alpha intelligence test (Yerkes, 1921; Yoakum & Yerkes, 1920) and from some of his own research (e.g., Terman, 1919). But these studies only reported cross-sectional data. At the time, no one knew whether the childhood IQ scores of a sample correlated with the same individuals' adult occupational complexity or prestige. These recommendations (e.g., Terman, 1916, 1922c, 1924c, 1928b, 1932) were not appropriate, given the lack of longitudinal data. The only exception in Terman's lifetime was his own longitudinal study of gifted children, but the minimum IQ cutoff score for sample members meant that his own study could not provide any vocational information for anyone with an IQ score lower than 135. Terman also made similar claims about the stability of IQ many years before there were longitudinal data showing that intelligence test scores were stable (e.g., Terman, 1924a, 1924c).

## Genetic Studies of Genius

More than 60 years after his death, Terman's longitudinal study of gifted children (Genetic Studies of Genius) is still considered a monumental achievement, and even Terman's critics admire the work (e.g., Greenberg, 1955; Minton, 1988; Winkler & Jolly, 2014). Despite the fact that most (perhaps all) of its subjects are now deceased, modern researchers still find value in the study and publish analyses of the data (e.g., Schmitt-Rodermund, Schröder, & Obschonka, 2017; Warne & Liu, 2017). Because the study is unique, massive, and influential, it has attracted its own body of critical literature.

Probably the most common criticism of Terman's study is its unusual and complicated sampling procedures (described in Terman, 1926, chap. 2). The resulting racially homogeneous sample has caused concern for many readers of Genetic Studies of Genius (e.g., Jolly, 2008a; Keating, 1975; Leslie, 2000; Minton, 1988; Rinn & Bishop, 2015; Robinson, 1981; Sorokin, 1956). It is true that the sample is overwhelmingly White: Of 1,444 individuals in the study in its initial phase (84 were added after the study began—mostly younger siblings of initial study members), between 20 and 50 had at least some non-European heritage, though the exact number is difficult to ascertain because Terman's racial and ethnic group classification maps poorly onto modern categories (Terman, 1926; Warne & Liu, 2017). Therefore, the sample of Genetic Studies of Genius was likely 95% to 99% White, a shockingly low level of diversity to 21st-century educational researchers. However, in the 1920 U.S. Census of California, 96.1% of the population under age 21 was White. To not find an overwhelmingly White sample of children for any study in California at the time would have been extremely unusual. Indeed, this would have been true for most of the United States at the time; the country as a whole was 89.7% White in the 1920 census (Bureau of the Census, 1924, p. 22).[6,7]

Some writers also criticize the disproportionately middle- and upper-class makeup of the sample (e.g., Minton, 1988, Sorokin, 1956). The first volume of Genetic Studies of Genius provides the numbers clearly. Almost exactly three quarters (75.3%) of the children's fathers worked as professionals or businessmen of one type or another, which Terman acknowledged was a far greater proportion than the general population in California at the time (Terman, 1926, pp. 62-63). The fathers also worked more

prestigious jobs than the average American man at the time (Terman, 1926, pp. 66-72). The families were overwhelmingly middle class (Terman, 1926, p. 72); 60.3% lived in "very superior" or "superior" neighborhoods, with another 30.1% living in "average" neighborhoods (Terman, 1926, pp. 76-77), and the typical home had 200 to 300 books in it (Terman, 1926, pp. 81-82). The parents were also well educated by the standards of the day, with 23.5% of children having a father with a bachelor's degree and 10.3% of children having a mother with a bachelor's degree (Terman, 1926, pp. 78-81).

Although there are no good comparison data to these economic statistics, these numbers are likely not representative of the general California population at the time. However, given the positive correlation between parents' socioeconomic status and children's IQ scores (e.g., Terman, 1916), a disproportionately middle-class sample is—again—completely expected. Yet the families were not remarkably wealthy by the standards of the time (Terman, 1926, p. 72), and the fathers' occupational prestige was similar to that of fathers in a study of adoptive parents conducted in California by one of Terman's students soon after (Burks, 1928/1973). The average occupational status of the fathers was equivalent to a chef at a large first class hotel or editor of a small newspaper—respectable, but not wealthy (Terman, 1926, pp. 68, 71). Despite claims to the contrary (e.g., Sorokin, 1956), the subjects of Genetic Studies of Genius were not highly privileged children whose future success was driven entirely by their parents' socioeconomic status. By 1940 (when the average male in the study was 30 years old), 70.3% of male subjects were in occupations classified as "professional" and "semi-professional" (Terman & Oden, 1947, p. 172), and by 1955 (at an average age of 45), 96.3% of male subjects worked in one of these two high-prestige areas (Terman & Oden, 1959, p. 74). There percentages are much higher than one would expect from regression toward the mean if intelligence were completely uncorrelated with adult economic success. The Terman subjects also earned bachelor's and graduate degrees at rates that far outstripped their parents or the general population at the time (Leslie, 2000)—another fact that would not be expected based solely on their childhood economic advantage.

Another noticeable characteristic of the sample in Genetic Studies of Genius is the noticeable gender imbalance: 856 males and 672 females, which is a 1.27:1 sex ratio. Terman was never able to fully explain this imbalance, and it puzzled him (Burks et al., 1930, pp. 471-472; Terman, 1926, pp. 49-54; Terman & Oden, 1947, pp. 12-14). His previous research had indicated an absence of mean sex differences in IQ scores (e.g., Terman, 1916), so having males overrepresented in his sample by 27% would have been a surprise. He considered four different possible reasons for the sex ratio and found most promising among these the explanation that males' IQ scores were more variable than females' scores. Called the *variability hypothesis*, Terman had tested this idea previously and rejected it, both with empirical data (Terman, 1916, p. 71; Terman et al., 1915, p. 559) and in a literature review (Terman & Chase, 1920). However, in Volume I of the Genetic Studies of Genius, he favored this hypothesis again (Terman, 1926, p. 54; see also Terman & Oden, 1947, pp. 12-14, where Terman and his coauthor seem less sure of the sex difference in variability, though other—stronger—data supporting the variability hypothesis were discussed).

Once again, modern research can shed light on the issue. The variability hypothesis has been supported in large, representative samples in several countries (e.g., Deary, Thorpe, Wilson, Starr, & Whalley, 2003; Hur, te Nijenhuis, & Jeong, 2017; Lakin, 2013; Strand, Deary, & Smith, 2006), usually with males having a 5% to 10% larger standard deviation for IQ scores. Assuming equal means and normally distributed variables within each group, a 5% difference in standard deviation between two groups would produce a ratio of 1.47:1 in the top 1% of individuals—more than enough to cause Terman's imbalanced sex ratio. Therefore, suggestions of a sex bias in the sampling or selection procedures in the Genetic Studies of Genius (e.g., Minton, 1988, p. 115) are not supported by the data. In general, the cause of the greater male variability in IQ scores is unclear, though it is also apparent in physical variables, including height, weight (Centers for Disease Control and Prevention, 2012), and the size of several brain structures (Wierenga et al., 2018), and in personality test scores (Allik & Realo, 2017).

Another common criticism of the Genetic Studies of Genius is the lack of a control group (Leslie, 2000; Minton, 1988). However, this criticism is not fully accurate. It is true that Terman did not select another sample of similar children with average IQ scores with which he could compare the gifted children. But that was not the purpose of the study. Rather, the intent of the study was to observe the development of high-IQ children and to determine their physical, educational, and social characteristics (Terman, 1922b, 1926). In other words, the main point of the study was to be descriptive—not comparative. Yet, where appropriate, Terman and his coauthors did compare his sample members with other groups. These were always *ad hoc* comparisons suitable for the specific variable at hand. These comparison groups could be the general population of the United States or a geographic area within the United States (e.g., Terman & Oden, 1947, pp. 111, 172), norm samples for tests (e.g., Terman, 1926, chap. 7), or previously published data (Terman, 1926, p. 310; Terman & Oden, 1947, p. 120). While these kinds of comparisons generally do not provide a satisfactory method of learning how the gifted differ from the average individual, Terman did learn enough to demonstrate in the first volume of the study that gifted children are neither social misfits nor wan, sickly individuals (Terman, 1926, chap. 21). Later volumes supported his hypotheses that these

individuals would be more likely to be successful in their professional and educational endeavors.

Far more crippling to the integrity of the study was the way Terman meddled in his subjects' lives. He wrote letters of recommendation, pulled strings to get them admitted to college (especially Stanford), and gave vocational and education advice to his "children" (Feldman, 1984; Janos, 1987; Leslie, 2000; Shurkin, 1992). In the 1960 survey, subjects were asked about the impact of the study on their life and the age when they learned that they were a part of it. More than two thirds of subjects (68.2% of men and 70.8% of women) learned that they were in the study by age 14, and 41.1% of men and 52.1% of women said that their participation had favorable impacts on their life (Oden, 1968, pp. 36-39). Seagoe (1975) stated that "the gifted thought of Terman as their godfather" (p. 94). Similarly, Minton (1988) said that "Terman was a benevolent father figure and psychological counselor" (p. 233).

Terman himself discussed his affection for his gifted "children" before the study officially began at a time when some of the older sample members were first identified:

> Nearly every child we have found with IQ above 140 is the kind one feels, before the test is over, one would like to adopt. If the crime of kidnapping could ever be forgiven it would be in the case of a child like one of these. (Terman, 1916, p. 101)

Perhaps—as Minton (1988) believed—this was because Terman identified with these gifted children (pp. 223-224). Regardless of the reason, Terman's influence on some of his subjects' lives damaged the scientific integrity of the Genetic Studies of Genius to a degree that posterity will never fully be able to ascertain.

Additionally, the Genetic Studies of Genius suffers other problems that have received much less attention. The data are sometimes chaotic and with a high rate of missingness (Shurkin, 1992; Subotnik, Karp, & Morgan, 1989; Warne & Liu, 2017), and even Terman (1922b, p. 315) admitted that the data were sometimes "faulty." (Though, to be fair, his data collection procedures were the state of the art at the time; Cravens, 1992.) The study is also "locked in time" (Cravens, 1992, p. 183), with data that are limited by the then-current theory and practice of the 1920s. For example, because multifactor models of cognitive abilities were in their infancy when the study began, there are no measures of specific cognitive abilities beyond global IQ (Stanley, 1974). Additionally, the only personality data from childhood were collected using the Woodworth–Cady Questionnaire, a test of "emotional stability" that was adapted from an instrument used to identify World War I soldiers who were at risk for mental health problems (Warne & Liu, 2017). Terman used this instrument because personality assessment at the time was in an even more primitive state of development than cognitive assessment, and there were no other suitable alternatives available. Another way in which the Genetic Studies of Genius is bound by its historical milieu is natural cohort effects. The subjects in the study were all born between 1901 and 1923, which means they lived very different lives from modern gifted children (Holahan & Sears, 1995; Rinn & Bishop, 2015; Warne & Liu, 2017). Minton (1988) suggested that adding new cohorts of children would have improved the generalizability of the study (p. 311).

Despite the problems of the study, the results of the Genetic Studies of Genius hold up remarkably well. Many have been replicated with more modern samples using more sophisticated statistical methods and research designs (Herrnstein, 1973; Keating, 1975; Minton, 1988; Robinson, 1981). Terman's longitudinal study was a direct inspiration for Julian Stanley as he planned the Study of Mathematically Precocious Youth (Stanley, 1974). Many of its findings have replicated findings from Genetic Studies of Genius. Both samples were generally healthier (Benbow, 1988), had low levels of religiosity as adults (Lubinski & Benbow, 1994), worked disproportionately in the sciences, earned graduate degrees at higher rates than the general population (Lubinski & Benbow, 2006), had high levels of professional accomplishment (Wai, Lubinski, & Benbow, 2005; Wai, Lubinski, Benbow, & Steiger, 2010), and earned higher incomes than average (Lubinski, Benbow, & Kell, 2014). Similar results are also apparent in longitudinal follow-ups of other high-IQ children (Gross, 2004; Harris, 1990; Makel, Kell, Lubinski, Putallaz, & Benbow, 2016). Other research on intelligence has supported Terman's results, such as the negative correlation between IQ and most mental and physical health conditions (Gale, Batty, Tynelius, Deary, & Rasmussen, 2010; Gottfredson & Deary, 2004; Wrulich et al., 2014). Later researchers have even replicated one of Terman's minor exceptions to this general trend: the higher rates of myopia among high-IQ individuals (Lubinski & Humphreys, 1992; Pickrell et al., 2016; Sörberg, Allebeck, & Hemmingsson, 2014).

## Terman's Lessons for Modern Researchers

Though he has been dead for more than 60 years, Terman's life and career have lessons for today's education scholars and practitioners. One important lesson is the value of changing one's opinions as new evidence appears. Renzulli (1978) noted that one reason for Terman's "mixed legacy" (p. 183) was that many people focus on the controversial statements of his early career and not on his final beliefs. Contrasting his early writings with later writings is illuminating, and Terman was willing to modify his views on many topics. Some of these changes were major (e.g., opinions about environmental influences on intelligence), while others were more subtle (e.g., his support for the variability hypothesis). Still, Terman could have done more to make his changes of opinion more publicly known (Gould, 1981, pp. 191-192). Making his

changes in opinion explicit could have done much to improve his reputation among future generations, many of whom only know Terman's opinions through secondary sources quoting his early writings.

Another lesson Terman can teach modern education scholars is the value of description. Terman was not adept at statistical procedures, and he rarely calculated anything more complicated than a correlation coefficient or a probable error. Instead, his forte was in description, perhaps because of his literary interest in biographies (Hilgard, 1957; Seagoe, 1975; Terman, 1932). Terman had an eye for detail, whether he was writing a case study (e.g., Terman, 1918a; Terman & Fenton, 1921), or describing the content of a test (e.g., Terman, 1906, 1916). The various volumes of Genetic Studies of Genius are packed with case studies, some of which make for the most interesting reading of the series. Additionally, the data collected on the subjects of the Genetic Studies of Genius are so exhaustive that almost any trait of interest to the psychologist, educator, or sociologist has a few pages in at least one volume devoted to it. The breadth and detail of the descriptive data may be a contributing factor to the ongoing influence of the study.

Modern researchers can also learn from Terman's trial-and-error approach to science. Terman was not a great theorist (Boring, 1959; Greenberg, 1955; Hilgard, 1957), and he recognized this in himself (Terman, 1932). Where he made his contributions to science is in his intuitive, experimental approach to gathering data. For example, the 1916 Stanford–Binet required three revisions before it was of high enough quality for Terman to publish (Terman et al., 1915). With each revision, Terman and his students tried out every test item on a pilot sample of children, analyzed the data, and determined which items functioned well enough to be retained. Unsuitable items were discarded, and new items were created. The entire process took 5 years (though some items from his 1906 dissertation and from Binet's tests appeared on the 1916 Stanford–Binet). The intermediate reports indicated that each step taught Terman much about psychological testing and cognitive development (e.g., Terman, 1911, 1915; Terman & Childs, 1912b, 1912c, 1912d). Terman had little guiding theory, so his only option was to dive in and pilot different tasks with hundreds of children individually. Modern educators and scholars could benefit from a trial-and-error approach to investigating a thorny problem, such as planning for the educational needs of a twice exceptional child.

A similar spirit of inventiveness is apparent in his work in the Genetic Studies of Genius. No one had ever conducted a longitudinal study before (Minton, 1988), and Terman had to create methods of maintaining contact with large numbers of individuals, archiving the huge data set, and developing questionnaires. Additionally, in the years leading up to the study he used the Stanford–Binet pilot groups to help him learn how to identify bright children quickly. This led to one of the quirks of his sampling method: testing the youngest children in a classroom (Terman, 1916). As odd as this method seems, it was his most efficient way of identifying high-IQ individuals (Terman, 1926, pp. 32-33).

Another lesson of Terman's career is the importance of accumulating evidence to test theories or hypotheses. Although Terman was often hasty in voicing a scientific opinion before strong evidence was available, many of his beliefs have been vindicated by scholarly evidence in the decades since his death. Even during his lifetime, Terman was willing to search for evidence to test his hypotheses. The Genetic Studies of Genius is a prime example of this, as Terman created it to search for evidence to support his beliefs that gifted children were healthy, well-adjusted individuals who would not experience a disproportionate share of negative adult outcomes. While Terman did not explicitly seek to subject his ideas to a strong disconfirmation test (such investigations were not common in the pre-Popper era), Terman did set an example of collecting data on many variables from large samples to support or refine his beliefs.

One final lesson of Terman's career that is valuable for modern researchers is his realization of how intertwined cognitive, developmental, educational, and socioemotional issues are for gifted children. While he began his study of giftedness by focusing on cognitive development (e.g., Terman, 1906, 1915), Terman soon moved on to studying socioemotional development, a trend which continued as the Genetic Studies of Genius progressed. In fact, the later volumes of the study contain more information about socioemotional variables than cognitive or educational variables (e.g., Holahan & Sears, 1995; Oden, 1968; Terman & Oden, 1959). Terman was prescient in understanding the importance of socioemotional issues in the field of gifted education. He noted, for example, the value of extracurricular activities for grade skippers and the difficulty that some had as they entered dating age (Terman, 1954). Terman's wide interests in many aspects of human development are an admirable example of a scientist's well-roundedness in contrast to the 21st-century age of specialization.

## How Should Modern Scholars Respond to Terman's Legacy?

Simple understandings of Terman's legacy are not forthcoming. He supported a meritocracy but was not concerned with how it would perpetuate social inequalities. He advocated eugenics in the early part of his career. Yet he recognized talent in racial minorities when it appeared—though he did not look very hard to find it. Many of his ideas are supported by modern data, but the data did not come until after his death, and he often overstated the strength of the contemporary evidence supporting his opinions.

It is likely that many workers in gifted education—and educators and psychologists in general—would benefit from grappling with the details of Terman's legacy. Decisions to either ignore his contribution to science or completely

embrace his career are too crude a solution. Instead, a proper solution is to investigate Terman's writings and to choose the ideas from his work that are useful today and to discard those that are not. Adopting one of Terman's ideas (e.g., placing importance on socioemotional issues in gifted children) does not mean that one must adopt all of his ideas (e.g., his belief in a genetic influence on intelligence). The claim—as the anonymous reviewer made to me—that because Terman was "grounded in racist attitudes" his data or any of his ideas are verboten—is an example of the *genetic fallacy* (where the word "genetic" is used in its original Latin sense of "related to origins"). The genetic fallacy is the belief that an idea must be rejected because it comes from the same source as other factually incorrect or morally repugnant ideas. Using Terman's data or arguing that he was correct to place a high value on intelligence in his study of giftedness does not mean that one also embraces eugenics or has a lack of concern about societal inequality.

Individuals who analyze Terman's work should seek to understand the historical time period and context of his career (see Chapman, 1988, for an excellent resource for this). This is a key step to avoiding *presentism*, which is a form of ethnocentrism that demands that individuals from the past be judged by present-day moral and ethical standards. Avoiding presentism means that one must withhold moral judgment in an effort to understand *how* figures of the past thought and *why* they behaved the way they did. Understanding the history of the eugenics movement, the influence of Galton and Binet on the early study of individual differences, and the genesis of the Genetic Studies of Genius is far more productive than the moral preening and virtue signaling of condemning Terman as a racist. Insights into context and why Terman and his colleagues believed what they did can help modern researchers and practitioners better understand which contributions of Terman's are worth honoring today—and which can be safely discarded.

This more complex view of Terman's work would avoid the oversimplifications and hasty condemnations of one of the pioneers of gifted education. Many researches can benefit as they delve into Terman's writings and discover statements and articles that are relevant to discussions in their fields today. For example, gifted education experts may find it interesting that Terman at one point or another

- implemented above-level testing (Terman & Fenton, 1921; see Warne, 2012),
- advocated for universal screening for giftedness (Terman, 1922c),
- recognized regression toward the mean in gifted individuals (Burks et al., 1930),
- advocated for academic acceleration for gifted children (Terman, 1954),
- observed that gifted children often have uneven cognitive ability profiles (Terman, 1926),
- showed the flaws of teacher nominations of giftedness (Terman, 1911, 1916; Terman et al., 1915),
- discussed the problems of strict cutoff score for identifying giftedness (Terman, 1916, 1921b), and
- explained the benefits of flexible ability grouping over strict academic tracking (Terman, 1922e).

Building on Terman's discoveries can provide modern researchers with insight into these issues and demonstrate that they are long-standing phenomena in gifted education. Likewise, Terman's contributions to the history of testing include

- the insight that different cognitive tasks vary in their quality as measures of intelligence (Terman, 1916, pp. 208),
- his finding that different individuals can tap into different abilities to solve the same problem (Terman, 1916, pp. 321-322),
- the poor quality of face validity evaluations of test items (Terman, 1916, pp. 39, 76-77, 114-115),
- the importance of predictive validity (Terman et al., 1915, p. 559), and
- the potential impact that testing could have on society (Terman, 1932, pp. 329-330).

## Conclusion

Lewis Terman has been called "the father of gifted education" (e.g., Winkler & Jolly, 2014, p. 70)—and for good reason. He almost singlehandedly established the field of gifted education as a scholarly discipline and was the first to raise many issues in the field that are still pressing today. His Genetic Studies of Genius is still the longest longitudinal study in the history of psychology, and it provided thousands of variables that modern researchers still fruitfully draw from.

However, Terman's legacy has been controversial, with the most common claims being that he overemphasized the importance of intelligence and IQ scores, supported a socially unjust meritocracy, and put too much faith in the power of "nature" over "nurture." This article shows that these criticisms are somewhat justified, but that the reality of Terman's thoughts on these issues is complex. Critics often overlook Terman's changes of opinion about many of these issues, or they oversimplify his views. The purpose of this article is to show that fairly judging Terman and his legacy is a multifaceted endeavor but one that provides benefits to practitioners and scholars in education, psychometrics, and psychology.

This article, however, is not a comprehensive view of Terman's work and legacy. His behavior and views regarding women, for example, are contradictory and also defy a simple label of "sexist" (Cravens, 1992; Gould, 1981; Rogers, 1999; Seagoe, 1975). For example, he had many female research assistants, students, and coauthors whom he treated as valued colleagues. But—unlike his male students and collaborators—he rarely helped them with their careers once they left Stanford University (Rogers, 1999). Likewise, some

readers may be uncomfortable with his social opinions because they are sometimes very alien from mainstream modern views. In evaluating Terman's social and political views, Cravens (1992) stated,

> It should be noted, however, that Terman was a consistent and insistent liberal reformer in American politics all his life. If he believed in innate differences, he was not, strictly speaking, a racist, nor did he believe that women were inherently inferior to men. He was probably as politically correct as he could be for his age, for those who care about such matters as retrospective ideological assessment . . . (p. 187)

Terman's biographers agree with this assessment (Minton, 1988; Seagoe, 1975), and given the evidence and the context,[8] it seems a fair one to me.

Nevertheless, Terman's legacy is complicated, and there will likely never be any consensus regarding every aspect of it. I believe, though, that it is a disservice to the man and his many contributions to dismiss him as a bigot or to label his work as irrelevant. To those who listen, Lewis Terman still has much to say to educators and scholars in several fields. Readers who take the time to become acquainted with his research can find a nuanced, fascinating, and intellectually rich body of work that still provides insight into intelligence, testing, giftedness, and education.

## Author's Note

This paper was presented at the annual conference of the National Association for Gifted Children on November 11, 2017, in Charlotte, NC.

## Acknowledgments

The author appreciates the feedback that he received from Dr. Karen Rogers, Dr. Jennifer Jolly, and other attendees at his presentation.

## Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author received no financial support for the research, authorship, and/or publication of this article.

## Notes

1. As I was revising this article, a study was published (Rimfeld et al., 2018) that showed that in Estonia both educational attainment and occupational prestige had lower levels of genetic influence before the fall of the Soviet regime than after. Although this finding needs to be replicated, it does provide evidence supporting Terman's belief that a meritocracy flourishes more in free nations than under an oppressive government. This is one of Terman's many ideas that would not find empirical support until decades after his death.
2. Notwithstanding Terman's exuberance about the importance of intelligence testing for the eugenics movement (Terman, 1924b, p. 106; 1924c, p. 340), historical studies show that the prime movers of the eugenics movement were mostly theorists, statisticians, biologists, and social crusaders. Psychologists seem to have only been influential in the eugenics of "feeblemindedness," and even then only with regard to identifying who should have this label. Psychologists seem to have had little influence on actual laws or policies, and few had major leadership positions in eugenics advocacy organizations (Broberg & Roll-Hansen, 2005; Kevles, 1995; Snyderman & Herrnstein, 1983).
3. One of Terman's students published an article (Merriman, 1924) based on his 1922 dissertation on the correlation of twins' IQ scores. But Merriman did not make the necessary comparisons of correlations between monozygotic and dizygotic twins' scores to estimate heritability (see Rende, Plomin, & Vandenberg, 1990).
4. Because heritability can only apply to populations and environments being studied, there is no guarantee that comparisons of American racial and ethnic groups would apply to international or worldwide comparisons of racial and ethnic groups. The variation in environments—and possibly genes—is much larger worldwide than it is within the United States. This has been recognized since the beginning of behavioral genetics studies of the heritability of intelligence (Burks, 1928/1973).
5. In response to a request from two peer reviewers, I briefly describe four types of circumstantial evidence that between-group differences in IQ are likely not 100% environmental. First, research on Spearman's hypothesis (e.g., Hartmann, Kruuse, & Nyborg, 2007; Jensen, 1980, 1998) shows that—for most racial groups—the size of racial group differences on cognitive tasks is positively correlated with the degree to which tasks measure general intelligence. An implication of this fact is that the same causes of within-group differences in intelligence (i.e., genes and environment) likely also contribute to between-group differences in intelligence. (However, Spearman's hypothesis does sometimes break down when comparing White and Asian groups within the United States; see Nagoshi, Johnson, DeFries, Wilson, & Vandenberg, 1984; Warne, 2016b, for examples. This may tentatively indicate that group differences on cognitive tasks for these two racial groups may not be strongly linked to differences in overall intelligence.) Second, multigroup confirmatory factor analyses show that for native-born American examinees, intelligence test scores for Black and White individuals show measurement invariance, which—like Spearman's hypothesis—indicates that between-group differences are likely to have very similar causes as within-group differences (Dolan, 2000; Lubke, Dolan, Kelderman, & Mellenbergh, 2003). In other words, both types of differences are probably due to genes and environment. Measurement invariance also eliminates the possibility that an outside variable affects one racial group and not another (Dalliard, 2014; Wicherts et al., 2004). Third, within- and between-group differences are algebraically related (Jensen, 1998, pp. 447-458) and that as environments for the two groups become more similar, the strength of genetic causes of between-group differences must increase. (The only exceptions are if the two groups are formed randomly or if groups consist of identical twins raised in separate

environments. Neither of these scenarios applies to racial groups within the United States.) Although environmental differences between Black and White individuals on a worldwide scale (e.g., comparing wealthy White Americans with Black sub-Saharan Africans in extreme poverty) are plausibly large enough to cause between-group genetic differences to drop to zero, environmental differences between Black and White Americans are not large enough to make these groups' intelligence score differences entirely environmental (Warne, Astle, & Hill, 2018). Finally, genetically derived scores (called polygenic scores) that positively correlate to adult educational attainment in Europeans also positively correlate with both variables in individuals who are descended from Africans, though the correlations are weaker than those seen in European populations (Domingue, Belsky, Conley, Harris, & Boardman, 2015; Piffer, 2015). This indicates that at least some of the genes that influence educational attainment in White individuals also influence these variables in Black individuals. Although one could postulate a 100% environmental cause of one of these facts, there is no plausible completely nongenetic explanation for all four. Conversely, there is no circumstantial evidence for a purely environmental explanation for between-group racial group differences in intelligence for Americans that is as strong as any one of these lines of circumstantial evidence that I explain in this footnote (Gottfredson, 2005b). Readers should not take this evidence to indicate that between-group differences in intelligence among Americans are 100% genetic. Such a conclusion is not supported by *any* of these types of circumstantial evidence. Instead, this evidence indicates that genes have a nonzero influence on group differences in intelligence; the exact magnitude of that influence (in absolute terms or when compared with the influence of environmental factors) cannot be determined on the basis of published information.

6. Less excusable is the decision to not canvass schools that served California's Chinese schoolchildren in the segregated education system (Terman, 1926, pp. 56-57). As a result, no children of Chinese descent were selected for the study. Terman never explained why these schools were not investigated, though he did later regret not doing so (Terman & Oden, 1947, pp. 14-15). In the end, only six to nine students of East Asian descent—all Japanese and four of them from the same family—are in the sample. Census numbers from 1920 indicated that Chinese individuals were 0.4% of children under 21 in California, and Japanese children were 1.8% of the population. However, only 64.7% of Chinese children and 54.4% of Japanese children of ages 5 to 20 attended school (compared with 69.4% of White individuals of the same age). It is impossible to know how many children would have been found if these Chinese schools were investigated, but the number would probably not have changed the overall demographics of the study.
7. Some writers mention that Nobel Physics Prize winners Luis Alvarez and William Shockley were tested for Terman's study and were rejected because their IQ scores were too low (e.g., Leslie, 2000; Shurkin, 1992). On the surface, this may seem like a damning flaw for Terman's sampling method, intelligence testing, and/or a study designed to identify "genius." However, outcomes with low base rates are extremely difficult to predict (Taylor & Russell, 1939). Terman's research assistants screened and/or tested 168,000 children in their search for sample members for the Genetic Studies of Genius (Warne, Godwin, & Smith, 2013), meaning that the base rate for future Nobel winners was 0.000019. For Terman's sample (which consisted of 0.91% of children considered for the study) to include *either* future Nobelist, all of the following conditions would have to have been present: (a) Terman's IQ cutoff (140 for most sample members) would have to be equal to or lower than the minimum IQ score for Nobel Prize winners, (b) the test score would need to have a reliability value of .99 or higher, and (c) IQ scores and the skills needed to win a Nobel Prize would have to be nearly perfectly correlated. There is no evidence that all three of these requirements were present at the time Terman was selecting children for participation in Genetic Studies of Genius. Terman likely would have been thrilled to have identified a gifted child who grew up to win a Nobel Prize, but with the hindsight of current psychometric knowledge, it is now apparent that it was almost impossible for Terman's sampling procedure to select Shockley or Alvarez—let alone both—to be part of Genetic Studies of Genius.
8. Terman's political actions from the mid-1930s until the end of his life may please some modern readers. See the sections in Seagoe (1975) and Minton (1988) on his political activities, which included supporting civil liberties and opposing American pacifism toward Japan and Germany in the late 1930s and early 1940s. Hilgard (1957) also makes it clear that Terman opposed racial discrimination in the United States.

## ORCID iD

Russell T. Warne https://orcid.org/0000-0003-4763-3625

## References

Allik, J., & Realo, A. (2017). Universal and specific in the five factor model of personality. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 173-190). New York, NY: Oxford University Press.

Bellhouse, D. R. (2009). Karl Pearson's influence in the United States. *International Statistical Review*, *77*, 51-63. doi:10.1111/j.1751-5823.2009.00066.x

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral and Brain Sciences*, *11*, 169-232.

Boring, E. G. (1959). *Lewis Madison Terman*. In *Biographical memoirs of the National Academy of Sciences* (Vol. 33, pp. 414-461). New York, NY: Columbia University Press.

Bouchard, T., Jr. (2014). Genes, evolution and intelligence. *Behavior Genetics*, *44*, 549-577. doi:10.1007/s10519-014-9646-x

Broberg, G., & Roll-Hansen, N. (Eds.). (2005). *Eugenics and the welfare state: Norway, Sweden, Denmark, and Finland*. East Lansing, MI: Michigan State University Press.

Broberg, G., & Tydén, M. (2005). Eugenics in Sweden: Efficient care. In G. Broberg & N. Roll-Hansen (Eds.), *Eugenics and the welfare state: Norway, Sweden, Denmark, and Finland* (pp. 77-149). East Lansing: Michigan State University Press.

Bureau of the Census. (1924). *Fourteenth census of the United States: State compendium: California*. Washington, DC: Government Printing Office.

Burks, B. S. (1973). The relative influence of nature and nurture upon mental development: A comparative study of foster parent–foster child resemblance and true parent–true child resemblance. In H. J. Eysenck (Ed.), *The Measurement of Intelligence* (pp. 325-357). Dordrecht, Netherlands: Springer. (Original work published 1928)

Burks, B. S., Jensen, D. W., & Terman, L. M. (1930). *Genetic studies of genius: Vol. III. The promise of youth: Follow-up studies of a thousand gifted children*. Stanford, CA: Stanford University Press.

Callahan, C. M. (1996). A critical self-study of gifted education: Healthy practice, necessary evil, or sedition? *Journal for the Education of the Gifted*, *19*, 148-163.

Calvin, A. (2000). Use of standardized tests in admissions in post-secondary institutions of higher education. *Psychology, Public Policy, and Law*, *6*, 20-32. doi:10.1037/1076-8971.6.1.20

Centers for Disease Control and Prevention. (2012). *Anthropometric reference data for children and adults: United States, 2007-2010*. Retrieved from http://www.cdc.gov/nchs/data/series/sr_11/sr11_252.pdf

Chapman, P. D. (1988). *School as sorters: Lewis M. Terman, applied psychology, and the intelligence testing movement, 1890-1930*. New York: New York University Press.

Cravens, H. (1992). A scientific project locked in time: The Terman Genetic Studies of Genius, 1920s-1950s. *American Psychologist*, *47*, 183-189.

Cucina, J. M., Peyton, S. T., Su, C., & Byle, K. A. (2016). Role of mental abilities and mental tests in explaining high-school grades. *Intelligence*, *54*, 90-104. doi:10.1016/j.intell.2015.11.007

Dahlstrom, W. G. (1985). *The development of psychological testing*. In G. A. Kimble & K. Schlesinger (Eds.), *Topics in the history of psychology* (Vol. 2, pp. 63-113). New York, NY: Lawrence Erlbaum.

Dalliard, M. (2014). The elusive x-factor: A critique of J. M. Kaplan's model of race and IQ. *Open Differe ntial Psychology*. doi:10.26775/ODP.2014.08.25

Deary, I. J., Thorpe, G., Wilson, V., Starr, J. M., & Whalley, L. J. (2003). Population sex differences in IQ at age 11: The Scottish mental survey 1932. *Intelligence*, *31*, 533-542. doi:10.1016/s0160-2896(03)00053-9

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology*, *86*, 130-147. doi:10.1037/0022-3514.86.1.130

Dolan, C. V. (2000). Investigating Spearman's hypothesis by means of multi-group confirmatory factor analysis. *Multivariate Behavioral Research*, *35*, 21-50. doi:10.1207/S15327906MBR3501_2

Domingue, B. W., Belsky, D. W., Conley, D., Harris, K. M., & Boardman, J. D. (2015). Polygenic influence on educational attainment: New evidence from the National Longitudinal Study of Adolescent to Adult Health. *AERA Open*, *1*(3), 1-13. doi:10.1177/2332858415599972

Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy*. New York, NY: W. W. Norton.

Feldman, D. H. (1984). A follow-up of subjects scoring above 180 IQ in Terman's "Genetic Studies of Genius." *Exceptional Children*, *50*, 518-523. doi:10.1177/001440298405000604

Gale, C. R., Batty, G. D., Tynelius, P., Deary, I. J., & Rasmussen, F. (2010). Intelligence in early adulthood and subsequent hospitalization for mental disorders. *Epidemiology*, *21*, 70-77. doi:10.1097/EDE.0b013e3181c17da8

Gillham, N. W. (2001). *A life of Sir Francis Galton: From African exploration to the birth of eugenics*. New York, NY: Oxford University Press.

Glass, B. (1986). Geneticists embattled: Their stand against rampant eugenics and racism in America during the 1920s and 1930s. *Proceedings of the American Philosophical Society*, *130*, 130-154.

Goodenough, F. L. (1926). Racial differences in the intelligence of school children. *Journal of Experimental Psychology*, *9*, 388-397. doi:10.1037/h0073325

Gottfredson, L. S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*, 13-23. doi:10.1016/S0160-2896(97)90011-8

Gottfredson, L. S. (1997b). Why ***g*** matters: The complexity of everyday life. *Intelligence*, *24*, 79-132. doi:10.1016/S0160-2896(97)90014-3

Gottfredson, L. S. (2004). Intelligence: Is it the epidemiologists' elusive "fundamental cause" of social class inequalities in health? *Journal of Personality & Social Psychology*, *86*, 174-199. doi:10.1037/0022-3514.86.1.174

Gottfredson, L. S. (2005a). *Implications of cognitive differences for schooling within diverse societies*. In C. L. Frisby & C. R. Reynolds (Eds.), *Comprehensive handbook of multicultural school psychology* (pp. 517-554). Hoboken, NJ: Wiley.

Gottfredson, L. S. (2005b). What if the hereditarian hypothesis is true? *Psychology, Public Policy, and Law*, *11*, 311-319. doi:10.1037/1076-8971.11.2.311

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science*, *13*, 1-4. doi:10.1111/j.0963-7214.2004.01301001.x

Gould, S. J. (1981). *The mismeasure of man*. New York, NY: W. W. Norton.

Greenberg, L. (1955). A critique of classic methods of identifying gifted children. *School Review*, *63*, 25-30. doi:10.2307/1083412

Gross, M. U. M. (2004). *Exceptionally gifted children* (2nd ed.). New York, NY: Routledge.

Hansen, B. S. (2005). Something rotten in the state of Denmark: Eugenics and the ascent of the welfare state. In G. Broberg & N. Roll-Hansen (Eds.), *Eugenics and the welfare state: Norway, Sweden, Denmark, and Finland* (pp. 9-76). East Lansing: Michigan State University Press.

Harris, C. R. (1990). The Hollingworth longitudinal study: Follow-up, findings, and implications. *Roeper Review*, *12*, 216-222. doi:10.1080/02783199009553275

Hartmann, P., Kruuse, N. H. S., & Nyborg, H. (2007). Testing the cross-racial generality of Spearman's hypothesis in two samples. *Intelligence*, *35*, 47-57. doi:10.1016/j.intell.2006.04.004

Herrnstein, R. J. (1971). I.Q. *The Atlantic*, *228*(3), 43-64.

Herrnstein, R. J. (1973). *I.Q. in the meritocracy*. Boston, MA: Atlantic Monthly Press.

Herrnstein, R. J., & Murray, C. (1996). *The bell curve: Intelligence and class structure in American life* (2nd ed.). New York, NY: Free Press.

Hildreth, G. H. (1925). *The resemblance of siblings in intelligence and achievement*. New York, NY: Columbia University, Teachers College.

Hilgard, E. R. (1957). Lewis Madison Terman: 1877-1956. *American Journal of Psychology*, *70*, 472-479.

Holahan, C. K., & Sears, R. R. (1995). *The gifted group in later maturity*. Stanford, CA: University of Stanford Press.

Hollingworth, H. L. (1939). Psychological factors in marital happiness [Book review]. *Psychological Bulletin*, *36*, 191-197. doi:10.1037/h0063047

Hur, Y.-M., te Nijenhuis, J., & Jeong, H.-U. (2017). Testing Lynn's theory of sex differences in intelligence in a large sample of Nigerian school-aged children and adolescents ($n > 11{,}000$) using Raven's Standard Progressive Matrices Plus. *Mankind Quarterly*, *57*, 428-437.

Janos, P. M. (1987). A fifty-year follow-up of Terman's youngest college students and IQ-matched agemates. *Gifted Child Quarterly*, *31*, 55-58. doi:10.1177/001698628703100202

Jensen, A. R. (1980). Précis of bias in mental testing. *Behavioral and Brain Sciences*, *3*, 325-333. doi:10.1017/S0140525X00005161

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Jolly, J. L. (2008a). Genetic Study of Genius: Elementary school students. *Gifted Child Today*, *31*(1), 27-33.

Jolly, J. L. (2008b). A paradoxical point of view: Lewis M. Terman. *Gifted Child Today*, *31*(2), 36-37.

Keating, D. P. (1975). Possible sampling bias in *Genetic Studies of Genius*. *Educational and Psychological Measurement*, *35*, 657-662. doi:10.1177/001316447503500313

Kevles, D. J. (1995). *In the name of eugenics: Genetics and the uses of human heredity*. Cambridge, MA: Harvard University Press.

Koch, L. (2006). Past futures: On the conceptual history of eugenics: A social technology of the past. *Technology Analysis & Strategic Management*, *18*, 329-344. doi:10.1080/09537320600777085

Lagemann, E. C. (2000). *An elusive science: The troubling history of education research*. Chicago, IL: University of Chicago Press.

Lakin, J. M. (2013). Sex differences in reasoning abilities: Surprising evidence that male–female ratios in the tails of the quantitative reasoning distribution have increased. *Intelligence*, *41*, 263-274. doi:10.1016/j.intell.2013.04.004

Leslie, M. (2000). The vexing legacy of Lewis Terman. *Stanford Magazine*. Retrieved from https://alumni.stanford.edu/get/page/magazine/article/?article_id=40678

Lubinski, D. (2016). From Terman to today: A century of findings on intelllectual precocity. *Review of Educational Research*, *86*, 900-944. doi:10.3102/0034654316675476

Lubinski, D., & Benbow, C. P. (1994). *The Study of Mathematically Precocious Youth: The first three decades of a planned 50-year study of intellectual talent*. In R. F. Subotnik & K. D. Arnold (Eds.), *Beyond Terman: Contemporary longitudinal studies of giftedness and talent* (pp. 255-281). Westport, CT: Ablex.

Lubinski, D., & Benbow, C. P. (2006). Study of Mathematically Precocious Youth after 35 years: Uncovering antecedents for the development of math–science expertise. *Perspectives on Psychological Science*, *1*, 316-345. doi:10.1111/j.1745-6916.2006.00019.x

Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science*, *25*, 2217-2232. doi:10.1177/0956797614551371

Lubinski, D., & Humphreys, L. G. (1992). Some bodily and medical correlates of mathematical giftedness and commensurate levels of socioeconomic status. *Intelligence*, *16*, 99-115. doi:10.1016/0160-2896(92)90027-o

Lubke, G. H., Dolan, C. V., Kelderman, H., & Mellenbergh, G. J. (2003). On the relationship between sources of within- and between-group differences and measurement invariance in the common factor model. *Intelligence*, *31*, 543-566. doi:10.1016/S0160-2896(03)00051-5

Makel, M. C., Kell, H. J., Lubinski, D., Putallaz, M., & Benbow, C. P. (2016). When lightning strikes twice: Profoundly gifted, profoundly accomplished. *Psychological Science*, *27*, 1004-1018. doi:10.1177/0956797616644735

Mercer, J. R. (1979). In defense of racially and culturally non-discriminatory assessment. *School Psychology Review*, *8*, 89-115.

Merriman, C. (1924). The intellectual resemblance of twins. *Psychological Monographs*, *33*, 1-57. doi:10.1037/h0093212

Minton, H. L. (1988). *Lewis M. Terman: Pioneer in psychological testing*. New York: New York University Press.

Moore, E. G. (1986). Family socialization and the IQ test performance of traditionally and transracially adopted Black children. *Developmental Psychology*, *22*, 317-326. doi:10.1037/0012-1649.22.3.317

Morelock, M. J., & Feldman, D. H. (1997). *High-IQ children, extreme precocity, and savant syndrome*. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 439-459). Boston, MA: Allyn & Bacon.

Nagoshi, C. T., Johnson, R. C., DeFries, J. C., Wilson, J. R., & Vandenberg, S. G. (1984). Group differences and first principal-component loadings in the Hawaii family study of cognition: A test of the generality of "Spearman's hypothesis." *Personality and Individual Differences*, *5*, 751-753. doi:10.1016/0191-8869(84)90125-9

Oden, M. H. (1968). The fulfillment of promise: 40-year follow-up of the Terman gifted group. *Genetic Psychology Monographs*, *77*, 3-93.

Pearson, K. (1903). On the inheritance of the mental and moral characters in man, and its comparison with the inheritance of the physical characters. *Journal of the Anthropological Institute of Great Britain and Ireland*, *33*, 179-237. doi:10.2307/2842809

Pearson, E. S. (1936). Karl Pearson: An appreciation of some aspects of his life and work: Part I: 1857-1906. *Biometrika*, *28*, 193-257. doi:10.1093/biomet/28.3-4.193

Pickrell, J. K., Berisa, T., Liu, J. Z., Segurel, L., Tung, J. Y., & Hinds, D. A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nature Genetics*, *48*, 709-717. doi:10.1038/ng.3570

Piffer, D. (2015). A review of intelligence GWAS hits: Their relationship to country IQ and the issue of spatial autocorrelation. *Intelligence*, *53*, 43-50. doi:10.1016/j.intell.2015.08.008

Plomin, R. (2001). The genetics of ***g*** in human and mouse. *Nature Reviews Neuroscience*, *2*, 136-141. doi:10.1038/35053584

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science*, *11*, 3-23. doi:10.1177/1745691615617439

Rende, R. D., Plomin, R., & Vandenberg, S. G. (1990). Who discovered the twin method? *Behavior Genetics*, *20*, 277-285.

Renzulli, J. S. (1978). What makes giftedness? Reexamining a definition. *Phi Delta Kappan*, *60*, 180-184, 261. doi:10.1177/003172171109200821

Richardson, J. T. E. (2011). *Howard Andrew Knox: Pioneer of intelligence testing at Ellis Island*. New York, NY: Columbia University Press.

Rimfeld, K., Krapohl, E., Trzaskowski, M., Coleman, J. R. I., Selzam, S., Dale, P. S., . . . Plomin, R. (2018). Genetic influence on social outcomes during and after the Soviet era in Estonia. *Nature Human Behaviour*, *2*, 269-275. doi:10.1038/s41562-018-0332-5

Rinn, A. N., & Bishop, J. (2015). Gifted adults: A systematic review and analysis of the literature. *Gifted Child Quarterly*, *59*, 213-235. doi:10.1177/0016986215600795

Robinson, H. B. (1981). *The uncommonly bright child*. In M. Lewis & L. A. Rosenblum (Eds.), *The uncommon child: Genesis of behavior* (pp. 57-81). New York, NY: Plenum Press.

Rogers, K. B. (1999). The lifelong productivity of the female researchers in Terman's Genetic Studies of Genius longitudinal study. *Gifted Child Quarterly*, *43*, 150-169. doi:10.1177/001698629904300303

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118-137. doi:10.1016/j.intell.2015.09.002

Scarr, S., & Weinberg, R. A. (1976). IQ test performance of Black children adopted by White families. *American Psychologist*, *31*, 726-739. doi:10.1037/0003-066x.31.10.726

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality & Social Psychology*, *86*, 162-173. doi:10.1037/0022-3514.86.1.162

Schmitt-Rodermund, E., Schröder, E., & Obschonka, M. (2017). Studying entrepreneurial occupations in the Terman women. *International Journal of Psychology*. Advance online publication. doi:10.1002/ijop.12450

Seagoe, M. V. (1975). *Terman and the gifted*. Los Altos, CA: Kaufmann.

Shurkin, J. N. (1992). *Terman's kids: The groundbreaking study of how the gifted grow up*. Boston, MA: Little, Brown.

Snyderman, M., & Herrnstein, R. J. (1983). Intelligence tests and the Immigration Act of 1924. *American Psychologist*, *38*, 986-995.

Sorokin, P. A. (1956). *Fads and foibles in modern sociology and related sciences*. Chicago, IL: Henry Regency.

Sörberg, A., Allebeck, P., & Hemmingsson, T. (2014). IQ and somatic health in late adolescence. *Intelligence*, *44*, 155-162. doi:10.1016/j.intell.2014.04.002

Stanley, J. C. (1974). *Intellectual precocity*. In J. C. Stanley, D. P. Keating, & L. H. Fox (Eds.), *Mathematical talent: Discovery, description, and development* (pp. 1-22). Baltimore, MD: Johns Hopkins University Press.

Stedman, L. M. (1924). *Education of gifted children*. Yonkers-on-Hudson, NY: World Book.

Stern, A. M. (2005). Eugenics and historical memory in America. *History Compass*, *3*, 1-11.

Strand, S., Deary, I. J., & Smith, P. (2006). Sex differences in Cognitive Abilities Test scores: A UK national picture. *British Journal of Educational Psychology*, *76*, 463-480. doi:10.1348/000709905x50906

Strenze, T. (2015). *Intelligence and success*. In S. Goldstein, D. Princiotta, & J. A. Naglieri (Eds.), *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts* (pp. 405-413). New York, NY: Springer.

Subotnik, R. F., Karp, D. E., & Morgan, E. R. (1989). High IQ children at midlife: An investigation into the generalizability of Terman's genetic studies of genius. *Roeper Review*, *11*, 139-144. doi:10.1080/02783198909553190

Tallman, G. G. (1928). A comparative study of identical and non-identical twins with respect to intelligence resemblances. In G. M. Whipple (Ed.), *The twenty-seventh yearbook of the National Society for the Study of Education: Nature and nurture: Part I. Their influence upon intelligence* (pp. 83-86). Bloomington, IL: Public School.

Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. *Journal of Applied Psychology*, *23*, 565-578. doi:10.1037/h0057079

Terman, L. M. (1905). A study in precocity and prematuration. *American Journal of Psychology*, *16*, 145-183. doi:10.2307/1412123

Terman, L. M. (1906). Genius and stupidity: A study of some of the intellectual processes of seven "bright" and seven "stupid" boys. *Pedagogical Seminary*, *13*, 307-373. doi:10.1080/08919402.1906.10534367

Terman, L. M. (1911). The Binet–Simon scale for measuring intelligence: Impresses gained by its application upon four hundred non-selected children. *Psychological Clinic*, *5*, 199-206.

Terman, L. M. (1915). The mental hygiene of exceptional children. *Pedogogical Seminary*, *22*, 529-537. doi:10.1080/08919402.1915.10533983

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet–Simon Intelligence Scale*. New York, NY: Houghton Mifflin.

Terman, L. M. (1918a). A experiment in infant education. *Journal of Applied Psychology*, *2*, 219-228. doi:10.1037/h0071908

Terman, L. M. (1918b). Expert testimony in the case of Alberto Flores. *Journal of Delinquency*, *4*, 145-164.

Terman, L. M. (1919). *The intelligence of school children: How children differ in ability, the use of mental tests in school grading, and the proper education of exceptional children*. Boston, MA: Houghton Mifflin.

Terman, L. M. (1921a). Intelligence and its measurement: A symposium—II. *Journal of Educational Psychology*, *12*, 127-133. doi:10.1037/h0064940

Terman, L. M. (1921b). Mental growth and the IQ. *Journal of Educational Psychology*, *12*, 401-407. doi:10.1037/h0069457

Terman, L. M. (1922a). Adventures in stupidity: A partial analysis of the intellectual inferiority of a college student. *Scientific Monthly*, *14*, 24-40.

Terman, L. M. (1922b). A new approach to the study of genius. *Psychological Review*, *29*, 310-318. doi:10.1037/h0071072

Terman, L. M. (1922c). *The problem*. In L. M. Terman (Ed.), *Intelligence tests and school reorganization* (pp. 1-31). Yonkers-on-Hudson, NY: World Book.

Terman, L. M. (1922d). The psychological determinist; or democracy and the I.Q. *Journal of Educational Research*, *6*, 57-62. doi:10.1080/00220671.1922.10879281

Terman, L. M. (1922e). Were we born that way? *World's Work*, *44*, 649-660.

Terman, L. M. (1924a). The American Psychological Association. *Science*, *59*, 546-548. doi:10.1126/science.59.1538.546

Terman, L. M. (1924b). The mental test as a psychological method. *Psychological Review*, *31*, 93-117. doi:10.1037/h0070938

Terman, L. M. (1924c). The possibilities and limitations of training. *Journal of Educational Research*, *10*, 335-343. doi:10.1080/00220671.1924.10879512

Terman, L. M. (1926). *Genetic studies of genius: Vol. I. Mental and physical traits of a thousand gifted children* (2nd ed.). Stanford, CA: Stanford University Press.

Terman, L. M. (1928a). The influence of nature and nurture upon intelligence scores: An evaluation of the evidence in Part I of the 1928 Yearbook of the National Society for the Study of Education. *Journal of Educational Psychology*, *19*, 362-373. doi:10.1037/h0071466

Terman, L. M. (1928b). The ultimate influence of standard tests. *Journal of Educational Research*, *17*, 57-59. doi:10.1080/00220671.1928.10879813

Terman, L. M. (1932). *Autobiography*. In C. Murchison (Ed.), *A history of psychology in autobiography* (Vol. 2, pp. 297-332). Worcester, MA: Clark University Press.

Terman, L. M. (1940). Personal reactions of the yearbook committee. In G. M. Whipple (Ed.), *The thirty-ninth yearbook of the National Society for the Study of Education: Intelligence: Its nature and nurture: Part II. Comparative and critical exposition* (pp. 460-467). Bloomington, IL: Public School.

Terman, L. M. (1954). The discovery and encouragement of exceptional talent. *American Psychologist*, *9*, 221-230. doi:10.1037/h0060516

Terman, L. M. (1958). What education for the gifted should accomplish. In N. B. Henry (Ed.), *The fifty-seventh yearbook of the National Society for the Study of Education: Part II. Education for the gifted* (pp. 15-19). Chicago, IL: University of Chicago Press.

Terman, L. M., & Chase, J. M. (1920). The psychology, biology and pedagogy of genius. *Psychological Bulletin*, *17*, 397-409. doi:10.1037/h0073715

Terman, L. M., & Childs, H. G. (1912a). A tentative revision and extension of the Binet-Simon measuring scale of intelligence: Part I. *Journal of Educational Psychology, 3*, 61-74. doi:10.1037/h0075624

Terman, L. M., & Childs, H. G. (1912b). A tentative revision and extension of the Binet–Simon measuring scale of intelligence: Part II. Supplementary tests. 1. Generalization test: Interpretation of fables. *Journal of Educational Psychology*, *3*, 133-143. doi:10.1037/h0073830

Terman, L. M., & Childs, H. G. (1912c). A tentative revision and extension of the Binet–Simon measuring scale of intelligence. Part II. Supplementary tests—continued. *Journal of Educational Psychology*, *3*, 198-208. doi:10.1037/h0071576

Terman, L. M., & Childs, H. G. (1912d). A tentative revision and extension of the Binet–Simon measuring scale of intelligence: Part III. Summary and criticisms. *Journal of Educational Psychology*, *3*, 277-289. doi:10.1037/h0073453

Terman, L. M., & Fenton, J. C. (1921). Preliminary report on a gifted juvenile author. *Journal of Applied Psychology*, *5*, 163-178. doi:10.1037/h0074962

Terman, L. M., Lyman, G., Ordahl, G., Ordahl, L., Galbreath, N., & Talbert, W. (1915). The Stanford revision of the Binet–Simon scale and some results from its application to 1000 non-selected children. *Journal of Educational Psychology*, *6*, 551-562. doi:10.1037/h0075455

Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence: A guide to the administration of the new Stanford–Binet tests of intelligence*. Boston, MA: Houghton Mifflin.

Terman, L. M., & Oden, M. H. (1947). *Genetic studies of genius: Vol. IV. The gifted child grows up: Twenty-five years' follow-up of a superior group*. Stanford, CA: Stanford University Press.

Terman, L. M., & Oden, M. H. (1959). *Genetic studies of genius: Vol. V. The gifted group at mid-life*. Stanford, CA: Stanford University Press.

Tizard, B. (1974). IQ and race. *Nature*, *247*, 316. doi:10.1038/247316a0

Vialle, W. (1994). "Termanal" science? The work of Lewis Terman revisited. *Roeper Review*, *17*(1), 32-38. doi:10.1080/02783199409553614

Wai, J., Lubinski, D., & Benbow, C. P. (2005). Creativity and occupational accomplishments among intellectually precocious youths: An age 13 to age 33 longitudinal study. *Journal of Educational Psychology*, *97*, 484-492. doi:10.1037/0022-0663.97.3.484

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. *Journal of Educational Psychology*, *102*, 860-871. doi:10.1037/a0019454

Warne, R. T. (2012). History and development of above-level testing of the gifted. *Roeper Review*, *34*, 183-193. doi:10.1080/02783193.2012.686425

Warne, R. T. (2016a). Five reasons to put the ***g*** back into giftedness: An argument for applying the Cattell–Horn–Carroll theory of intelligence to gifted education research and practice. *Gifted Child Quarterly*, *60*, 3-15. doi:10.1177/0016986215605360

Warne, R. T. (2016b). Testing Spearman's hypothesis with Advanced Placement examination data. *Intelligence*, *57*, 87-95. doi:10.1016/j.intell.2016.05.002

Warne, R. T., Astle, M., & Hill, J. C. (2018). What do undergraduates learn about human intelligence? An analysis of introductory psychology textbooks. *Archives of Scientific Psychology*, *6*, 32-50. doi:10.1037/arc0000038.

Warne, R. T., Godwin, L. R., & Smith, K. V. (2013). Are there more gifted people than would be expected in a normal distribution? An investigation of the overabundance hypothesis. *Journal of Advanced Academics*, *24*, 224-241. doi:10.1177/1932202x13507969

Warne, R. T., & Liu, J. K. (2017). Income differences among grade skippers and non-grade skippers across genders in the Terman sample, 1936-1976. *Learning and Instruction*, *47*, 1-12. doi:10.1016/j.learninstruc.2016.10.004

Weinberg, R. A., Scarr, S., & Waldman, I. D. (1992). The Minnesota Transracial Adoption Study: A follow-up of IQ test performance at adolescence. *Intelligence*, *16*, 117-135. doi:10.1016/0160-2896(92)90028-p

Wicherts, J. M., Dolan, C. V., Hessen, D. J., Oosterveld, P., van Baal, G. C. M., Boomsma, D. I., & Span, M. M. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, *32*, 509-537. doi:10.1016/j.intell.2004.07.002

Wierenga, L. M., Sexton, J. A., Laake, P., Giedd, J. N., Tamnes, C. K., & Pediatric Imaging, Neurocognition, and Genetics Study. (1918). A key characteristic of sex differences in the developing brain: Greater variability in brain structure of boys than girls. *Cerebral Cortex*, *28*, 2741-2751. doi:10.1093/cercor/bhx154

Wingfield, A. H. (1928). *Twins and orphans: The inheritance of intelligence*. London, England: J. M. Dent.

Winkler, D. L., & Jolly, J. L. (2014). *Lewis M. Terman: A misunderstood legacy (1877-1956)*. In A. Robinson & J. L. Jolly (Eds.), *A century of contributions to gifted education: Illuminating lives* (pp. 64-78). New York, NY: Routledge.

Wrulich, M., Brunner, M., Stadler, G., Schalke, D., Keller, U., & Martin, R. (2014). Forty years on: Childhood intelligence predicts health in middle adulthood. *Health Psychology*, *33*, 292-296. doi:10.1037/a0030727

Yerkes, R. M. (1917). The Binet versus the point scale method of measuring intelligence. *Journal of Applied Psychology*, *1*, 111-122. doi:10.1037/h0070364

Yerkes, R. M. (1921). *Psychological examining in the United States army*. Washington, DC: Government Printing Office.

Yoakum, C. S., & Yerkes, R. M. (1920). *Army mental tests*. New York, NY: Henry Holt.

Zagorsky, J. L. (2007). Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress. *Intelligence*, *35*, 489-501. doi:10.1016/j.intell.2007.02.003

## Author Biography

**Russell T. Warne**, PhD, is an associate professor of psychology in the Department of Behavioral Sciences at Utah Valley University. He earned his bachelor's degree in psychology from Brigham Young University in 2007 and his doctorate in educational psychology from Texas A&M University in 2011. He serves on the editorial boards for the *Journal of Psychoeducational Assessment*, the *Journal of School Psychology*, and *Gifted Child Quarterly* (*GCQ*). This is his fifth article in *GCQ*. His 2014 *GCQ* article, "Using above-level testing to track growth in academic achievement in gifted students" was named the journal's Paper of the Year. He is the author of *Statistics for the Social Sciences: A General Linear Model Approach*, a statistics textbook for undergraduate students recently published by Cambridge University Press.

Manuscript received January 7, 2018; Final revision received June 26, 2018; Accepted August 15, 2018