The Measurement of Human Abilities

ing systems, mate-selection strategies, child-rearing strategies, xenophobia, aggressive behavior, nepotism, and indoctrinability. We will have more to say about sociobiology in chapter 18.

As we will see in the remainder of this chapter, Darwin's ideas ultimately gave birth to a uniquely U.S. type of psychology—a psychology that emphasized individual differences and their measurement, the adaptive value of thoughts and behavior, and the study of animal behavior. Before discussing U.S. psychology, however, we must first review the works of a man who was an important link between Darwinian theory and U.S. psychology.

Sir Francis Galton

Erasmus Darwin, the physician, philosopher, poet, and early evolutionary theorist, was the grandfather of both Charles Darwin and **Francis Galton** (1822–1911), Darwin's cousin. Galton was born near Birmingham, England, on February 16, the youngest of seven children. His father was a wealthy banker, and his mother was a half-sister of Charles Darwin's father. Receiving his early education at home, Galton could read and write by the age of 2½. At age 5, he could read any book written in English, and by age 7, he was reading such authors as Shakespeare for pleasure. But things changed when Galton was sent to a boarding school where his experiences included being flogged, hell-raising, enduring sermons from the teachers, and fighting with his fellow students. At age 16, he was taken out of boarding school and sent to Birmingham General Hospital to study medicine; after this practical experience, he transferred to King's College in London. He then moved to Cambridge University where he obtained his degree in 1843. Galton planned to return to King's College to obtain his medical degree; but when his father died, he decided not to, so his formal education ended.

Because Galton was independently wealthy, he could work on what he wanted, when he wanted. After graduation he traveled in Egypt, the Sudan, and the Middle East. Then he came home and socialized with his rich friends for a few years—riding,

ARCHIVES OF THE HISTORY OF AMERICAN PSYCHOLOGY

Francis Galton

shooting, ballooning, and experimenting with electricity. After consulting with a phrenologist who recommended an active life, Galton decided to join the Royal Geographical Society on a trip to southwest Africa. The trip lasted two years, and for Galton's creation of a map of previously unexplored territories in Africa (now called Namibia), the Royal Geographical Society honored him in 1853 with their highest medal. Galton was 32 at the time. We can see in Galton's map-making ability a passion that Galton had all his adult life: the passion to measure things.

In 1853, Galton published his first book, *Narrative of an Explorer in Tropical South Africa*. He became a recognized expert on travel in the wild, and the British government commissioned him to teach camping procedures to soldiers. In 1855, he published his second book, *The Art of Travel*, which included information on how to deal with wild animals and savages. For his inventiveness, Galton was elected president of the Royal Geographical Society in 1856.

To further illustrate Galton's passion for measurement, here are a few of his other endeavors:

- In his effort to measure and predict the weather, he invented the weather map and was the first to use the terms *highs*, *lows*, and *fronts*.
- He was the first to suggest that fingerprints could be used for personal identification—a procedure later adopted by Scotland Yard.
- He attempted to determine the effectiveness of prayer (he found it ineffective).
- He tried to determine which country had the most beautiful women.
- He measured the degree of boredom at scientific lectures.

One can imagine Galton's delight when he became aware of his cousin's evolutionary theory with its emphasis on individual differences. Galton believed that if there were important individual differences among people, clearly they should be measured and cataloged. This became Galton's mission in life.

The Measurement of Intelligence

Galton assumed that intelligence was a matter of sensory acuity because humans could know the world only through the senses. Thus, the more acute the senses, the more intelligent a person was presumed to be. Furthermore, because sensory acuity was mainly a function of natural endowment, intelligence was inherited. And if intelligence was inherited, as Galton assumed, one would expect to see extremes in intelligence run in families. Assuming that high reputation or eminence was an accurate indicator of high intellectual ability, Galton set out to measure the frequency of eminence among the offspring of illustrious parents as compared to the frequency of eminence among the offspring of the general population. For comparison with the general population, Galton studied the offspring of judges, statesmen, commanders, literary men, scientists, poets, musicians, painters, and divines. The results published in *Hereditary Genius: An Inquiry into its Laws and Consequence* (1869) were clear: The offspring of illustrious individuals were far more likely to be illustrious than were the offspring of nonillustrious individuals. Galton also observed, however, that zeal and vigor must be coupled with inherited capacity before eminence could be attained.

Eugenics. Galton's conclusion raised a fascinating possibility: *selective breeding*. If intelligence was inherited, could not the general intelligence of a people be improved by encouraging the mating of bright people and discouraging the mating of people who were less bright? Galton's answer was yes. He called the improvement of living organisms through selective breeding **eugenics,** and Galton (1869) advocated its practice:

> I propose to show in this book that a man's natural abilities are derived by inheritance, under exactly the same limitations as are the form and physical features of the whole organic world. Consequently, as it is easy, notwithstanding those limitations, to obtain by careful selection a permanent breed of dogs or horses gifted with peculiar powers of running, or of doing anything else, so it would be quite practicable to produce a highly-gifted race of men by judicious marriages during several consecutive generations. I shall show that social agencies of an ordinary character, whose influences are little suspected, are at this moment working towards the degradation of human nature, and that others are working towards its improvement. I conclude that each generation has enormous power over the natural gifts of those that follow, and maintain that it is a duty we owe to humanity to investigate the range of that power, and to exercise it in a way that, without being unwise towards ourselves, shall be most advantageous to future inhabitants of the earth. (p. 45)

In 1865, Galton proposed that couples be scientifically paired and that the government pay those possessing desirable characteristics to marry. The government was also to take care of the educational expenses of any offspring. After reading *Hereditary Genius*, Darwin wrote to his cousin, "You have made a convert of an opponent in one sense, for I have always maintained that excepting fools, men did not differ much in intellect only in zeal and hard work" (quoted in Pearson, 1914, p. 6).

Mental Imagery

Galton was also among the first, if not the first, to study imagery. In *Inquiries* he reported the results of asking people to imagine the scene as they had sat down to breakfast. He found that the ability to imagine was essentially normally distributed, with some individuals almost totally incapable of imagery and others having the ability to imagine the breakfast scene flawlessly. Galton was amazed to find that many of his scientist friends had virtually no ability to form images. If sensations and their remnants (images) were the stuff of all thinking, as the empiricists had assumed, why was it that many scientists seemed unable to form and use images? Galton also found, not so surprisingly, that whatever a person's imagery ability was, he or she assumed that everyone else had the same ability.

Anthropometry

Galton's desire to measure individual differences among humans inspired him to create what he called an "anthropometric laboratory" at London's International Health Exhibition in 1884. Here, in about one year, Galton measured 9,337 humans in just about every way he could imagine. For example, he measured head size, arm span, standing height, sitting height, length of the middle finger, weight, strength of hand squeeze (measured by a dynamometer), breathing capacity, visual acuity, auditory acuity, reaction time to visual and auditory stimuli, the highest detectable auditory tone, and speed of blow (the time it takes for a person to punch a pad). Some of these measures were included because Galton believed sensory acuity to be related to intelligence, and for that reason, Galton's "anthropometric laboratory" can be viewed as an effort to measure intelligence. The nonacuity measures were included because Galton was interested in a number of issues related to individual differences. In 1888, Galton set up a similar laboratory in the science galleries of the South Kensington Museum, and it operated for several years. A handout described the purpose of the laboratory to potential participants:

1. For the use of those who desire to be accurately measured in many ways, either to obtain timely warning of remediable faults in development, or to learn their powers.
2. For keeping a methodological register of the principal measurements of each person, of which he may at any future time obtain a copy under reasonable restrictions. His initials and date of birth will be entered in the register, but not his name. The names are indexed in a separate book.
3. For supplying information on the methods, practice, and uses of human measurement.
4. For anthropometric experiment and research, and for obtaining data for statistical discussion. (Pearson, 1924, p. 358)

For a small fee (3 pence), a person would be measured in all ways described above; and for a smaller fee (2 pence), a person could be measured again at another time. Each participant was given a copy of his or her results, and Galton kept a copy for his files. Among the many things that Galton was interested in examining were test-retest relationships, gender differences on various measurements, intercorrelations among various measurements, relationships of various measurements to socioeconomic status, and family resemblances among various measurements. Because Galton's incredible amount of data existed long before there were computers or even calculators, much of it went unanalyzed at the time. Since then, however, other researchers have analyzed portions of the previously unanalyzed data. Recently, Johnson, McClearn, Yuen, Nagoshi, Ahern, and Cole (1985) reported the results of Galton's own analyses, the results of analyses of Galton's data done by researchers after him, and their own analyses of Galton's data that had not been previously analyzed.

Although intelligence is no longer believed to be related to sensory acuity, Galton's early efforts can be seen as the beginning of the mental testing movement in psychology. Following our review of Galton, we will have more to say about how intelligence testing changed after Galton's efforts.

The Concept of Correlation

The last of Galton's many contributions to psychology we will consider is his notion of correlation, which has become one of psychology's most widely used statistical methods. In 1888, Galton published an article entitled "Co-Relations and Their Measurement, Chiefly from Anthropometric Data," and in 1889 he published a book entitled *Natural Inheritance*. Both works describe the concepts of correlation and regression. Galton (1888) defined **correlation,** or co-relation, as follows:

> Two variable organs are said to be co-related when the variation on one is accompanied on the average by more or less variation of the other, and in the same direction. Thus the length, of the arm is said to be co-related with that of the leg, because a person with a long arm has usually a long leg, and conversely. (p. 135)

In a definition of correlation, the word *tend* is very important. Even in the above quotation, Galton said that those with long arms *usually* have long legs. After planting peas of varying sizes and measuring the size of their offspring, Galton observed that very large peas tended to have offspring not quite as large as they were and that very small peas tended to have offspring not quite as small as themselves. He called this phenomenon **regression toward the mean,** something he also found when he correlated heights of children with heights of their parents. In fact, Galton found regression whenever he correlated inherited characteristics. Earlier, Galton had observed that eminent individuals tended to have only eminent offspring.

By visually displaying his correlational data in the form of scatterplots, Galton found that he could visually determine the strength of a relationship. It was **Karl Pearson** (1857–1936) who devised a formula that produced a mathematical expression of the strength of a relationship. Pearson's formula produces the now familiar **coefficient of correlation (r).**

In addition to introducing the concept of correlation, Galton also introduced the *median* as a measure of central tendency. He found the *mean* to be overly influenced by extreme scores in a distribution and preferred to use the middle-most score (the median) in a distribution instead.

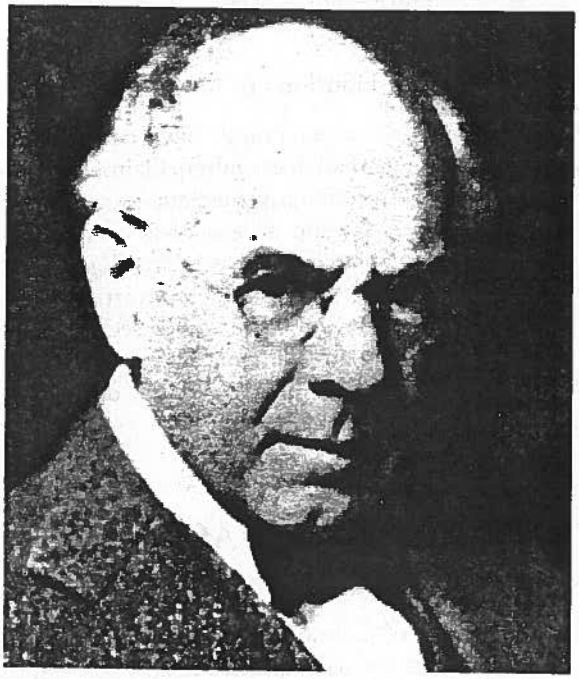
Galton's Contributions to Psychology

Few individuals in psychology have more firsts attributed to them than does Galton. Galton's firsts include study of the nature-nurture question, the use of questionnaires, the use of a word-association test, twin studies, the study of imagery, intelligence testing, and the development of the correlational technique. Everywhere in his work, we see a concern with individual differences and their measurements, a concern that was a direct reflection of the influence of Darwin's theory of evolution.

Intelligence Testing After Galton

James McKeen Cattell

The transfer of Galton's testing procedures to the United States was accomplished mainly through the efforts of **James McKeen Cattell** (1860–1944), who had studied with both Wundt and Galton in Europe but had been much more influenced by Galton. Cattell, born on May 25 in Easton, Pennsylvania, was a son of a Presbyterian clergyman who was also a professor of Latin and Greek at Lafayette College and later its president. Cattell entered Lafayette College before his 16th birthday and stood first in his class without much effort. Among his favorite subjects were mathematics and physics. After graduation from Lafayette in 1880, he traveled to Germany to study with the Kantian physiologist R. H. Lotze (1817–1881). Cattell was very impressed by Lotze, and it came as quite a blow when Lotze died a year after Cattell's arrival. The following year, Cattell returned home and wrote a paper on philosophy that won him a fellowship at Johns Hopkins University. While at Johns Hopkins (1882–1883), he did research in G. Stanley Hall's new psychology laboratory (see chapter 11) and decided to become a psychologist. In 1883, Cattell returned to Germany, this time to study with Wundt. Cattell was not only



ARCHIVES OF THE HISTORY OF AMERICAN PSYCHOLOGY

James McKeen Cattell

Wundt's first experimental assistant but was also the first student from the United States to earn a doctorate under Wundt's supervision. Cattell received his degree in 1886. While with Wundt, Cattell and a fellow student did numerous reaction-time studies. Among other things, Cattell noticed that his own reaction times differed systematically from those of his fellow researcher and proposed to Wundt that individual differences in reaction time be explored. The proposal was rejected because Wundt was more interested in the nature of the mind in general than with individual differences.

After attaining his doctorate, Cattell returned to the United States where he taught at Bryn Mawr College and the University of Pennsylvania. About this time, Cattell became aware of Galton's anthropometric laboratory in London and began a correspondence with Galton, mainly concerning the measurement of reaction time. Soon Cattell applied for and received a two-year research fellowship at Cambridge University where he worked with Galton. In Galton, Cattell finally found someone who shared his intense interest in individual differences. Galton confirmed Cattell's conviction that individual differences were important and that they could be objectively measured. Under Galton's influence, Cattell came to believe that intelligence was related to sensory acuity and was therefore largely inherited.

> As a self-proclaimed disciple of Francis Galton, Cattell's interest in eugenics is clear. . . . He proposed that incentives be given "the best elements of all the people" to intermarry and have large families [Cattell and his wife had seven children] and in fact offered each of his children $1,000 if they would marry the child of a college professor. (Sokal, 1971, p. 630)

On his return to the United States in 1888, Cattell was first affiliated with the University of Pennsylvania, where he founded the first psychology laboratory designed for undergraduate students in 1889. It was also at the University of Pennsylvania that Cattell administered Galtonian-type measures to his students. In 1890, Cattell published his techniques and results in an article that used the term *mental test* for the first time:

> Psychology cannot attain the certainty and exactness of the physical sciences, unless it rests on a foundation of experiment and measurement. A step in this direction could be made by applying a series of mental tests and measurements to a large number of individuals. The results would be of considerable scientific value in discovering the constancy of mental processes, their interdependence, and their variation under different circumstances. (p. 373)

It was also in this article that Cattell described ten mental tests that he believed could be administered to the general public and a total of 50 tests that he believed should be administered to university students. The ten mental tests were mainly Galtonian, but Cattell also added a few measurements he learned in Wundt's laboratory. Among the ten tests were hand strength, two-point threshold, amount of pressure required to cause pain, ability to discriminate between weights, reaction time, accuracy of

bisecting a 50-centimeter line, accuracy in judging a 10-second interval, and ability to remember a series of letters. The more comprehensive series of 50 tests was essentially more of the same; the vast majority of them measured some form of sensory acuity or reaction time.

In 1891, Cattell moved to Columbia University where he began administering his tests to entering freshmen. Implicit in Cattell's testing program was the assumption that if a number of his tests were measuring the same thing (intelligence), performance on those tests should be highly correlated. Also implicit was the assumption that if tests were measuring intelligence, they should correlate highly with academic success in college. That is, for a test of intelligence to be valid, it must make differential predictions about how individuals will perform on tasks requiring intelligence.

In 1901, Clark Wissler, one of Cattell's graduate students, tested Cattell's assumptions. Armed with Pearson's newly perfected correlation coefficient, Wissler measured the relationships among Cattell's tests and between performance on various tests and academic performance. Wissler's results were disastrous for Cattell's testing program. He found that intercorrelations among the tests were very low and that the correlation between various tests and success in college was nearly zero (Guilford, 1967). Thus, the tests were not measuring the same thing because if they were, they would be highly correlated; and they were not valid because if they were, scores would correlate highly with academic achievement.

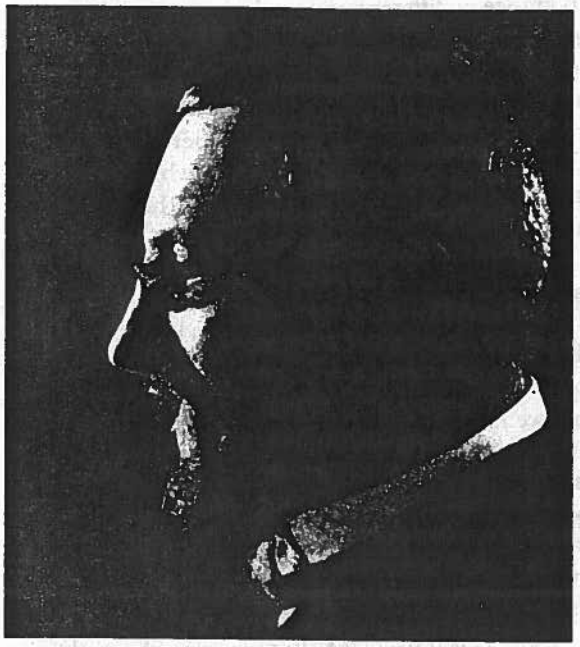
With such unambiguous, negative findings, the interest in mental testing quickly faded. Wissler switched his field to anthropology where he became an outspoken environmentalist, and Cattell turned to other aspects of applied psychology. Because Cattell was a key figure in the school of functionalism, we will consider him further in the next chapter. The emphasis in U.S. psychology was turning toward practicality, and it appeared that Galtonian measures were not very useful, at least as far as intelligence was concerned. This moratorium on mental testing was not to last long, however.

Alfred Binet

In France, a different approach to measuring intelligence was being tried, one that appeared to be more successful than Galton's. It involved *directly* measuring the complex mental operations thought to be involved in intelligence. **Alfred Binet** (1857–1911) championed this method of testing, which was more in the rationalist tradition than in the empiricist tradition.

Binet was born on July 11 in Nice, France. His father was a physician, as were both of his grandfathers. Binet's parents separated when he was a young child, and he, an only child, was reared mainly by his mother, a successful artist. Although he initially followed the family tradition by studying medicine, Binet terminated his medical studies and turned to psychology instead. Being independently wealthy allowed Binet to take the time to educate himself, and he read the works of Darwin, Galton, and the British empiricists (especially John Stuart Mill), among others. He received no formal education in psychology.

Binet began his career in psychology by working with Jean-Martin Charcot (1825–1893), the world famous psychiatrist, at La Salpêtrière. Like Charcot, Binet conducted research on hypnotism, and he claimed that in one study he had been able to manipulate the symptoms and sensations of a hypnotized subject by moving a magnet to various places around the subject's body. He also claimed that application of the magnet could convert fear of an object, such as a snake, into affection. Binet thought that such findings would have important implications for the practice of medicine in general and for psychiatry in particular, but other researchers could not reproduce Binet's findings and concluded Binet's results were due to poor experimental control. For example, it was found that Binet's subjects always knew what was expected of them and acted accordingly. When subjects were unaware of the researcher's expectations, they did not exhibit the phenomena Binet had observed. Thus, suggestion had caused Binet's results, not the magnet. After a long attempt to defend his beliefs, Binet finally admitted that his results had been due to suggestion



ARCHIVES OF THE HISTORY OF AMERICAN PSYCHOLOGY

Alfred Binet

and not to the magnet's power, and he resigned his position at La Salpêtrière in 1890. The humiliation resulting from his public admission of shoddy research procedures haunted Binet all his life. His statement "Tell me what you are looking for, and I will tell you what you will find" (Wolf, 1973, p. 347) was directed at metaphysicians, but Binet knew from personal experience that it could apply to researchers as well.

Fortunately, Binet's second career in psychology was more successful. Without a professional position, Binet directed his attention to the study of the intellectual growth of his two daughters, who were 2½ and 4½ years old at the time. The tests he created to investigate his children's mental operations were very similar to those Jean Piaget later devised. He asked, for example, which of two piles contained more objects and found that the answer was not determined by the number of objects in the piles but by the amount of space the piles took up on the table. Binet also investigated how well his daughters could remember objects that he first showed them and then removed from sight. Binet also employed a number of tests used by Galton and Cattell to measure visual acuity and reaction time. In 1890, he published three papers describing his research on his daughters, and in 1903 he published *The Experimental Study of Intelligence*, which summarized his longitudinal study of the intellectual growth of his daughters.

In 1891, Binet joined the laboratory for physiological psychology at the Sorbonne, where he performed research in such areas as memory, the nature of childhood fears, the reliability of eyewitness testimony, creativity, imageless thought, and graphology. During his years at the Sorbonne, Binet also investigated individual differences in the perception of inkblots—before the famous work of Rorschach. In her outstanding biography of Binet, Wolf (1973) says that Binet was the father of experimental psychology in France and that he had more of an impact on U.S. psychology than Wundt did. (The reader is directed to Wolf's book for more details concerning Binet's many pioneering research endeavors and for the interesting details of his life.)

Individual psychology. Rather than being interested in what people have in common, Binet was primarily interested in what made them different. In 1896, he and his assistant Victor Henri (1872–1940) wrote an article entitled "Individual Psychology," which proposed a list of variables on which individuals differ, especially intellectually. What they sought was a list of important variables and a way of determining the extent to which each variable exists in a given individual. With the variables isolated and a way of measuring them available, they hoped that it would be possible to "evaluate" any individual in a relatively short period of time. The work of Galton and Cattell was rejected because it placed too much emphasis on sensory processes and not enough on higher mental processes. In other words, Binet and Henri proposed to study cognitive abilities *directly* instead of indirectly via sensory acuity. Another reason the work of Galton and Cattell was rejected is that it minimized important differences

between a child's mind and an adult's. According to Binet and Henri, the important variables on which humans differ are complex, higher-order processes that vary according to age. The list of such variables proposed in 1896 included memory, imagery, imagination, attention, comprehension, suggestibility, aesthetic judgment, moral judgment, force of will, and judgment of visual space.

Unfortunately, Binet and Henri's goal of accessing a person's higher mental processes in a relatively short period of time failed. Administering the tests took many hours, and interpreting the results required even more hours of subjective, clinical judgment. Even more devastating, however, was the study on their tests performed by Stella Sharp, a graduate student at Cornell University. Sharp (1899) found very low intercorrelations among the Binet and Henri tests and concluded (as Wissler had concluded about Cattell's tests) that they could not be measuring the same attribute (presumably intelligence). Such findings, along with their own disappointing results, caused Binet and Henri to abandon their "individual psychology" project. The experience gained, however, would serve Binet well on his next project.

Assessing intellectual deficiency. In 1899, **Theodore Simon** (1873–1961), who worked as an intern at a large institution for children with mental retardation, asked Binet to supervise his doctoral research. Binet agreed and viewed this as an opportunity to have access to a large subject pool. Also in 1899, Binet joined the Free Society for the Psychological Study of the Child, an organization that sought scientifically valid information about children, especially about their educational problems. Binet soon became leader of the society. In 1903, Binet and Simon were appointed to the group that the French government commissioned to study the problems of children with retardation in the French schools. It was immediately clear that if children with retardation were to receive special education, it was necessary to have an adequate method of distinguishing them from normal children. At the time, variations of Galton's tests were being used to detect mental retardation, and Binet noted that because of these tests, children who were blind or deaf were erroneously being classified as having mental deficiencies.

In 1904, Binet and Simon set out to create tests that would differentiate between intellectually normal and intellectually subnormal children. Their first step was to isolate one group of children clearly diagnosed as normal and another group diagnosed as subnormal. The second step was to test both groups in a number of different ways, hoping to discover measurements that would clearly distinguish members of one group from the other. From his previous research, Binet was convinced that the best way to examine individual differences was in terms of complex, mental processes, and so many of the tests given to the normal and subnormal children were of that type. After much trial and error, Binet and Simon arrived at the first test of intelligence that measured intelligence directly instead of indirectly through measures of sensory acuity.

The 1905 Binet–Simon scale of intelligence and its revisions. Binet and Simon offered the **Binet–Simon scale of intelligence** as a valid way of distinguishing between normal children and children with mental deficiencies—a way that was to replace the less reliable physical, social, and educational signs being used at the time to identify children with mental retardation. The 1905 scale consisted of 30 tests ranging in difficulty from simple eye movements to abstract definitions. Three of the tests measured motor development, and the other 27 were designed to measure cognitive abilities. The tests were arranged in order of difficulty, so that the more tests a child passed the more fully developed his or her intelligence was assumed to be. The scale was given to normal children and to children thought to have retardation, all of them between the ages of 2 and 12. The 30 tests contained in the 1905 Binet–Simon scale were the following:

1. Demonstrating visual coordination
2. Demonstrating prehension on contact (grasping a cube after touching it)
3. Demonstrating prehension on sight

4. Recognizing food (choice between wood and chocolate)
5. Seeking food (in response to chocolate wrapped in paper)
6. Following simple orders or repeating gestures
7. Pointing to objects (head, nose, and so on)
8. Recognizing objects in a picture
9. Naming objects in a picture
10. Discriminating length of two lines
11. Repeating three digits
12. Discriminating two weights
13. Resisting suggestions
14. Defining simple words
15. Repeating sentence of 15 words
16. Giving differences between pairs of objects
17. Demonstrating visual memory
18. Drawing forms from memory
19. Demonstrating memory span for digits
20. Stating similarities between objects
21. Discriminating lines rapidly
22. Ordering five weights
23. Identifying missing weight (of the five weights in test 22)
24. Giving rhyming words
25. Completing sentences
26. Constructing sentence containing three given words
27. Answering questions (such as "What should you do when sleepy?")
28. Giving time after hands of a clock have been interchanged
29. Folding and cutting paper
30. Distinguishing between abstract terms (such as *sad* and *bored*)

Binet and Simon found that almost all normal children aged 2 years or older could easily pass tests 1 through 6. Also, children with slight or moderate retardation could pass some or all of these tests. Children with severe retardation could pass only a few or none of them. Most of tests 7 through 15 could be passed by normal children between the ages of 2 and 5. Children with slight retardation could pass several of these tests, children with moderate retardation had great difficulty, and children with severe retardation could rarely pass any of them. Tests 16 through 30 could be routinely passed by normal children between the ages of 5 and 12, but children with even slight retardation had great difficulty with them, and children with moderate and severe retardation usually could pass none.

We see in the Binet–Simon scale a reflection of Binet's belief that intelligence is not a single ability but several. With this belief, Binet reflects the faculty psychology of several rationalistic philosophers. He did not, however, accept the nativism that often accompanies rationalistic viewpoints. He did believe that inheritance may place an upper limit on one's intellectual ability, but he also believed that almost everyone functions below their potential. Therefore, he believed strongly that *everyone* could grow intellectually, and that fact should be of prime importance to educators.

In 1908, Binet and Simon revised their scale. Their goal now was to go beyond simply distinguishing normal children from children with retardation, to distinguishing among levels of intelligence for normal children. The tests were administered to a large number of normal children from ages 3 to 13. If 75% or more of the children of a certain age passed a particular test, the test was assigned to that age level. For example, most 4-year-old children could copy a square but not a diamond. More specifically, it was found that only a minority of 3-year-olds could copy a square, a majority of 4-year-olds (75% or more) could copy a square, and essentially all 5-year-olds could do so. In this way, it could be determined whether a given child was performing at, above, or below average. A 5-year-old passing the tests that most other 5-year-olds also passed was considered to have normal intelligence. But if that child passed only the tests typically passed by 4-year-olds, he or she was thought to have below-average intelligence. And if the 5-year-old passed tests normally passed

by 6-year-olds, he or she was thought to have above-average intelligence. In other words, a child's intelligence level was determined by how much higher or lower than the norm the child performed. The 1908 revision of the Binet–Simon scale consisted of 58 tests, each showing the age at which 75% or more of the children taking it perform correctly.

The 1911 revision of the scale included normative data on adults (15-year-olds) and provided exactly five tests for each age level. The latter allowed for a more refined measure of intelligence. For example, if an 8-year-old child passed all the tests corresponding to his or her age, he or she would be considered normal. It is possible, however, that an 8-year-old will also pass some tests typically passed only by 9-year-olds. The new procedure allowed one-fifth of a year to be added to a child's score for each test the child passed beyond those that were the norm for his or her age. Thus, a child's "intellectual level" could be expressed in terms of intellectual age—that is, the age corresponding to the most difficult tests the child could pass.

Binet warned that extreme caution should be taken in interpreting a child's "intellectual age." For one thing, he observed that it was quite common for children to have an intellectual age that was only one year behind their chronological age and these children probably would have little trouble in school. Children whose intellectual age was two or more years behind their chronological age would probably have trouble in a standard school program and would need special attention. But even in the latter case, poor test performance did not necessarily mean the child had mental deficiencies. Before such a label was applied, the test administrator had to ensure that the child was healthy and motivated when he or she took the test and that he or she was knowledgeable enough about French culture to understand the reflections of that culture on the test.

Intelligence quotient. In 1911, **William Stern** (1871–1938), a German psychologist, introduced the term **mental age.** For Stern, a child's mental age was determined by his or her performance on the Binet–Simon tests. Stern also suggested that mental age be divided by chronological age, yielding an **intelligence quotient.** For example, if a particular 7-year-old passed all tests typically passed by 7-year-olds, his or her intelligence quotient would be 7/7, or 1.00. If another 7-year-old passed only those tests typically passed by 5-year-olds, his or her intelligence quotient would be 5/7, or about .71. In 1916, Lewis Terman suggested that the intelligence quotient be multiplied by 100 to remove the decimal point. It was also Terman who abbreviated intelligence quotient as IQ. Thus, combining the suggestions made by Stern and Terman, we have the familiar formula for IQ:

$$IQ = \frac{\text{Mental Age (MA)}}{\text{Chronological Age (CA)}} \times 100$$

Binet was opposed to the use of the intelligence quotient. He believed that intelligence was too complex to be represented by a simple term or number. History shows, however, that Stern's simplifications won out over Binet's opposition. In any case, Binet and Simon had developed a relatively brief, easy-to-administer measure of intelligence, and it became extremely popular. By the beginning of World War I, the Binet–Simon test was being used throughout most of the world.

Binet's view of his intelligence scale. Before reviewing what happened to the Binet–Simon scale in the United States, it is important to review how Binet viewed his scale. First and foremost, Binet saw the scale as a device for identifying children who needed some sort of special education. Binet strongly believed that children with low test scores could benefit considerably if given special attention. Although Binet believed that inheritance may set an upper limit on intellectual potential, he also believed that everyone could grow a great deal intellectually if properly stimulated. He worried very much about students in classrooms where teachers believed that students' intellectual performance was innately determined. This, of course, was especially regretful for students believed to have low intelligence.

> I have often observed, to my regret, that a widespread prejudice exists with regard to the educability of intelligence. The familiar proverb, "When one is stupid, it is for a long time," seems to be

> accepted indiscriminately by teachers with a stunted critical judgment. These teachers lose interest in students with low intelligence. Their lack of sympathy and respect is illustrated by their unrestrained comments in the presence of the children: "This child will never achieve anything. . . . He is poorly endowed. . . . He is not intelligent at all." I have heard such rash statements too often. They are repeated daily in primary schools, nor are secondary schools exempt from the charge. (Binet, 1909/1975, p. 105)

In Binet's (1909/1975) reaction to those who maintained that some children would *never* accomplish certain things, he indicates clearly that he did not accept an extreme nativist view of intelligence:

> "Never!" What a strong word! A few modern philosophers seem to lend their moral support to these deplorable verdicts when they assert that an individual's intelligence is a fixed quantity, a quantity which cannot be increased. We must protest and react against this brutal pessimism. We shall attempt to prove that it is without foundation. (pp. 105–106)

Mental orthopedics. Binet believed that mental orthopedics could prepare disadvantaged children for school. **Mental orthopedics** consisted of exercises that would improve a child's will, attention, and discipline—all abilities that Binet thought were necessary for effective classroom education. Binet (1909/1975) believed that by engaging in mental orthopedics, children learned how to learn.

> If we consider that intelligence is not a single function, indivisible and of a particular essence, but rather that it is formed by the chorus of all the little functions of discrimination, observation, retention, etc., the plasticity and extensibility of which have been determined, it will appear undeniable that the same law governs the whole and its parts, and that consequently anyone's intelligence is susceptible to being developed. With practice, training, and above all, method, we manage to increase our attention, our memory, our judgment and literally to become more intelligent than we were before. Improvement goes on in this way until the time when we reach our limit. (p. 107)

Both Binet and Galton died in 1911. Galton was an old man of 89 who had a long, highly productive life; Binet was 54 and at the height of his career.

Charles Spearman and the Concept of General Intelligence

After a military career in the English army that lasted until he was 34, **Charles Spearman** (1863–1945) turned to a career in psychology, studying with both Wundt and Külpe in Germany. During a break in his studies with Wundt, during which he returned to England to serve in the army during the Boer War (1900–1902), Spearman began reading the works of Galton. Thoroughly impressed, he performed a number of experiments on village schoolchildren, and the results tended to confirm Galton's belief concerning the relationship between sensory acuity and intelligence. He found that not only did measures of sensory acuity correlate highly among themselves but, more important, they also correlated highly (+.38) with "cleverness in school." In 1904, he published his results in an article entitled "'General Intelligence,' Objectively Determined and Measured." In part because of this controversial article, Spearman was offered a position at University College, London, where he began a career that included attacks on sensationalism, associationism, hedonism, and most other accepted philosophical and psychological beliefs. Among the more specific things he attacked were the results of studies, such as Wissler's, that showed little intercorrelation among Galton's and Cattell's measures of sensory acuity and almost no correlation between measures of sensory acuity and academic performance. Because his own results were almost the opposite, he concluded that the results contrary to his were statistical artifacts. He also concluded that because the measures of sensory acuity were intercorrelated, they must be measuring a common ability or faculty, which he named **general intelligence (g).** Furthermore, following in the Galtonian tradition, he claimed that *g* was determined almost exclusively by inheritance.

Spearman's conclusions about the nature of intelligence are important for three reasons: (1) He

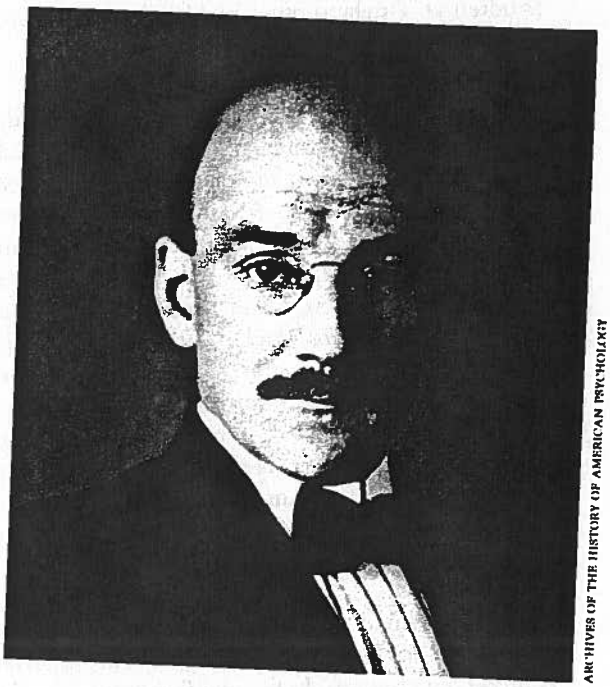
viewed intelligence as a unitary faculty, whereas Binet viewed it as comprising many different faculties; (2) he viewed intelligence as largely inherited, whereas Binet viewed it as modifiable by experience; and (3) it was largely Spearman's conception of intelligence that was embraced by the new testing movement in the United States, not Binet's. That is, IQ was viewed as measuring something like Spearman's *g* rather than Binet's multifarious "intellectual level."

The Binet–Simon Scale in the United States

Henry Herbert Goddard

Henry Herbert Goddard (1866–1957) was born into a New England Quaker family and obtained his bachelor's and master's degrees from Haverford College. After being a high school teacher and then principal for six years, he enrolled in the doctoral program in psychology at Clark University to pursue his interests in education and psychology. Goddard did his doctoral dissertation, which investigated the psychological factors involved in faith healing, under the supervision of G. Stanley Hall (see chapter 11). After completing his degree in 1899, Goddard first accepted a teaching position at Pennsylvania's West Chester State Teacher's College, and then in 1906 he became director of research at the Training School for the Feebleminded in Vineland, New Jersey.

It was Goddard who translated the Binet–Simon scale into English. Although initially skeptical of the scale, he found it to be very effective in classifying children in terms of their degree of retardation. Goddard then translated all of Binet and Simon's works into English and, following Binet's death in 1911, became the world's leading proponent of Binet's approach to measuring intelligence. However, although accepting Binet's testing procedures, Goddard accepted the Galton–Cattell–Spearman view of the nature of intelligence rather than Binet's. The fact that Goddard believed that intelligence was a single faculty whose strength is determined by heredity is clearly seen in the following quotation:



ARCHIVES OF THE HISTORY OF AMERICAN PSYCHOLOGY

Henry Herbert Goddard

> Stated in its boldest form, our thesis is that the chief determiner of human conduct is a unitary mental process which we call intelligence: that this process is conditioned by a nervous mechanism which is inborn: that the degree of efficiency to be attained by that nervous mechanism and the consequent grade of intelligence or mental level for each individual is determined by the kind of chromosomes that come together with the union of the germ cells: that it is but little affected by any later influences except such serious accidents as may destroy part of the mechanism. (Goddard, 1920, p. 1)

In addition to administering the translated Binet–Simon scale to the children at the Vineland School, Goddard also administered it to 2,000 public school students in New Jersey. He was shocked to find that many of the public school students performed below the norms for their ages. This especially disturbed Goddard because of his belief that intelligence was largely inherited—a belief he thought was supported by the observation that the

children at Vineland often had brothers and sisters who were "feebleminded" (Goddard's term).

Study of the "Kallikak" family. Goddard decided to investigate the relationship between family background and intelligence more carefully. In 1911, he administered the Binet–Simon scale to Deborah Kallikak, who had been living at the Vineland School since 1897. "Kallikak" was a fictitious name that Goddard created out of the Greek words *kalos* (good) and *kakos* (bad). Although Deborah's chronological age was 22, her test performance yielded a mental age of 9, producing an IQ of about 41. Goddard coined the term *moron* to denote Deborah's intellectual level. He then traced Deborah's ancestry back to the American Revolution, when Martin Kallikak, Sr., had had a relationship with a "feebleminded" barmaid that resulted in the birth of Martin Kallikak, Jr. After leaving the army, the elder Martin married a "worthy girl," and they had seven children. The younger Martin eventually married and had ten children. In Goddard's analysis, the descendants of the elder Martin and the "worthy girl" represented the "good" side of Deborah's ancestry, and the descendants of the younger Martin represented the "bad" side.

Goddard found that of the elder Martin's children, none were "feebleminded," whereas five of the younger Martin's children were. In subsequent generations on the younger Martin's side, Goddard found an abundance of individuals with mental deficiencies. In Goddard's time, people believed that "feeblemindedness" was the cause of most criminal, immoral, and antisocial behavior; and Goddard supported this belief by showing that many descendants of the younger Martin had been horse thieves, prostitutes, convicts, alcoholics, parents of illegitimate children, or sexual deviates. Of the hundreds of descendants from the elder Martin's marriage, only three had had mental deficiencies, and one had been considered "sexually loose." Among the elder Martin's descendants had been doctors, lawyers, educators, and other prestigious individuals.

Goddard reported his findings in *The Kallikak Family, a Study in the Heredity of Feeble-Mindedness* (1912). His research was taken as support for the Galtonian belief that intelligence was genetically determined. Along with Goddard, several leading scientists of the day urged that those with mental deficiencies be sterilized or segregated from the rest of society. They contended that because the feebleminded could not be expected to control their own reproduction, the intelligent members of society must control it for them.

> If both parents are feeble-minded all the children will be feeble-minded. It is obvious that such matings should not be allowed. It is perfectly clear that no feeble-minded person should ever be allowed to marry or to become a parent. It is obvious that if this rule is to be carried out, the intelligent part of society must enforce it. (Goddard, 1914, p. 561)

No fewer than 20 states passed sterilization laws, and thousands of "undesirables" were sterilized. In some states, the sterilization law was enforced until the 1970s. Galton would have been pleased.

Mental testing and immigration. In the years from 1905 to 1913, millions of individuals migrated from Europe to the United States, and there was growing concern that many of these immigrants might have mental deficiencies. The question was how to know for certain. In 1912, the commissioner of immigration invited Goddard to Ellis Island to observe the immigrants. Goddard claimed he could tell that many of the immigrants had mental deficiencies simply by observing their physical characteristics, but to be sure he administered the Binet–Simon scale. On the basis of the test results, many immigrants were labeled "mentally defective," and thousands were deported. Goddard even went so far as to specify the European countries for which the percentage of immigrants with mental deficiencies was the highest. In general, Goddard concluded that between 40 and 50% of the immigrants were "morons."

As with his earlier work, Goddard assumed that the immigrants' test performance was due mainly to inherited intelligence and not to educational, cultural, or personal experience—all factors that were later found to profoundly influence test performance. But the immigrants were also taking the test under special circumstances.

> For the evident reason, consider a group of frightened men and women who speak no English and who have just endured an oceanic voyage in steerage. Most are poor and have never gone to school; many have never held a pencil or pen in their hand. They march off the boat: one of Goddard's [assistants] takes them aside shortly thereafter, sits them down, hands them a pencil, and asks them to reproduce on paper a figure shown to them a moment ago, but now withdrawn from their sight. Could their failure be a result of testing conditions, of weakness, fear, or confusion, rather than of innate stupidity? Goddard considered the possibility, but rejected it. (Gould, 1981, p. 166)

Furthermore, the tests were administered by a translator whose accuracy in translating the test into the immigrant's native tongue was taken on faith.

Because of Goddard's efforts, the rate of deportation increased 350% in 1913 and 570% in 1914. Except for all the common, inexpensive laborers the United States was losing, Goddard was pleased. In his later years, Goddard radically changed his beliefs by embracing many of Binet's views. For example, he finally agreed that the proper treatment for individuals who scored low on intelligence tests was special education, not segregation or sterilization. But he had already done much damage.

Lewis Madison Terman

Lewis Madison Terman (1877–1956) was born on January 15, the 12th of 14 children of a farm family from central Indiana. He went to a one-room school and completed the eighth grade when he was 12 years old. At age 9, a phrenology book salesman gave each member of the Terman family a phrenological analysis. Terman's analysis indicated great promise, thus stimulating him to aspire for a life beyond the farm. At age 15, Terman left the farm to attend Central Normal College in Danville, Indiana. At age 17, he began teaching in a rural school. Within six years after leaving home, Terman had taught school and earned three undergraduate degrees: one in arts, one in sciences, and one in pedagogy. The next three years were busy ones for Terman; he became a high school principal, a husband, and a father. In 1901, he enrolled at Indiana University where he pursued a master's degree in pedagogy. Upon completing his master's degree, he was about to seek a teaching position when he received the offer of a fellowship for doctoral study at Clark University. With financial support from his family, Terman was able to accept the offer, and soon he was off to study with G. Stanley Hall, as Goddard had done.

ARCHIVES OF THE HISTORY OF AMERICAN PSYCHOLOGY

Lewis Madison Terman

Hall conducted a Monday evening seminar at Clark during which two students would present their work to about 30 psychology, philosophy, and education students. At the end of the student presentations and the discussion that ensued, Hall offered what Terman (1932) considered to be a most impressive summary.

> Hall would sum things up with an erudition and fertility of imagination that always amazed us and made us feel that his offhand insight into the problem went immeasurably beyond that of the student who had devoted months of slavish drudgery to it. . . . [At the end of each session] I always went home dazed and intoxicated, took a hot bath to quiet my nerves, then lay awake for hours rehearsing the drama and formulating the clever things I should have said and did not. (p. 316)

Terman did not write his dissertation under Hall's supervision, however. Terman became increasingly interested in mental testing, and Hall had little enthusiasm for the topic. Under the supervision of Edmund C. Sanford, Terman isolated a group of "bright" students and a group of "dull" students and then attempted to determine what types of tests could be used to differentiate between members of the two groups. (Terman was unaware that Binet and Simon had done essentially the same thing earlier.) Terman's dissertation was entitled "Genius and Stupidity: A Study of the Intellectual Processes of Seven 'Bright' and Seven 'Stupid' Boys." Terman was to say later in his life that all of his career interests were shaped during his years at Clark.

Before obtaining his doctorate from Clark University in 1905, Terman had become seriously ill with tuberculosis, and although he recovered, he thought it best that he choose a warm climate in which to work. For that reason, he accepted the position of high school principal in San Bernardino, California. A year later, he accepted a position teaching child study and pedagogy at Los Angeles State Normal School (later to become the University of California at Los Angeles). In 1910, Terman accepted an appointment to the education department at Stanford University where he spent the rest of his career. He became chair of the psychology department in 1922, a position he held until his retirement in 1942.

It was coincidental with his arrival at Stanford that Terman became aware of the Binet–Simon intelligence scale (through Goddard's translation). Terman began immediately to work with the scale and found that it could not be used accurately on U.S. children without modifications.

The Stanford-Binet tests. Terman found that when the Binet–Simon scale was administered to U.S. children, the results were uneven. That is, the average scores of children of various ages were either higher or lower than the chronological age of the age group being tested. For example, Terman observed that items from the Binet–Simon scale were too easy for 5-year-olds and too difficult for 12-year-olds. This caused the mental age of average 5-year-olds to be artificially high and that of average 12-year-olds to be artificially low. Along with his graduate student, H. G. Childs, Terman deleted existing items from the Binet–Simon scale and added new items until the average score of a sample of children was 100, no matter what their age. This meant that for each age group tested, the average mental age would equal the group's chronological age. Terman and Childs published their first revision of the Binet–Simon tests in 1912, and in 1916 Terman alone published a further revision. The 1916 revision became known simply as the Stanford–Binet. It was in 1916 that Terman adopted Stern's "intelligence ratio" and suggested that the ratio be multiplied by 100 to remove the decimal and to call the ratio *IQ*. The Stanford–Binet, which made Terman both rich and famous, was revised in 1937 and again in 1960 (after Terman's death).

Terman's position on the inheritance of intelligence. Throughout his career, Terman believed that intelligence was largely inherited. Furthermore, Terman, like Goddard, believed that low intelligence was the cause of most criminal and other forms of antisocial behavior. For Terman (1916), a stupid person could not be a moral person.

> Not all criminals are feeble-minded, but all feeble-minded persons are at least potential criminals. That every feeble-minded woman is a potential prostitute would hardly be disputed by anyone. Moral judgment, like business judgment, social judgment, or any other kind of higher thought process, is a function of intelligence. Morality cannot flower and fruit if intelligence remains infantile. (p. 11)

And in 1922 Terman said:

> There is nothing about an individual as important as his IQ, except possibly his morals . . . the great test problem of democracy is how to adjust itself to the large IQ differences which can be demonstrated to exist among the members of any race or nationality group. . . . All the available facts that science has to offer support the Galtonian theory that mental abilities are chiefly a matter of original endow-

> ment. . . . It is to the highest 25 per cent. of our population, and more especially to the top 5 per cent., that we must look for the production of leaders who will advance science, art, government, education, and social welfare generally. . . . The least intelligent 15 or 20 per cent. of our population . . . are democracy's ballast, not always useless but always a potential liability. How to make the most of their limited abilities, both for their own welfare and that of society; how to lead them without making them helpless victims of oppression; are perennial questions in any democracy. (quoted in Minton, 1988, p. 99)

Although Terman was impressed by and borrowed much from Binet, his view of intelligence was much more like that of Galton. Terman was so impressed by Galton that he published an intellectual portrait of him in which he estimated Galton's IQ to be nearly 200 (Terman, 1917).

Terman's contention that IQ was a valid measure of native intelligence did not go unchallenged. Among Terman's harshest critics was the journalist Walter Lippmann. Lippmann and Terman debated in a series of articles appearing in the *The New Republic* between 1922 and 1923. In one such article Lippmann (1923) wrote:

> I hate the impudence of a claim that in fifty minutes you [Terman] can judge and classify a human being's predestined fitness in life. I hate the pretentiousness of that claim. I hate the abuse of scientific method which it involves. I hate the sense of superiority which it creates and the sense of inferiority which it imposes. (p. 46)

Terman validated the Stanford–Binet by correlating test performance with teacher ratings of academic performance, teacher estimations of intelligence, and school grades. He found fairly high correlations in each case, but this was not surprising because the traits and abilities that schools and teachers valued highly in students were the same traits and abilities that yielded high scores on the Stanford–Binet. Nonetheless, the correlations meant that academic performance could be predicted with some success from test performance. Whether the tests were truly measuring native intelligence, however, Terman never determined.

Terman's study of genius. In Terman's day, it was widely believed that very bright children were abnormal in more than a statistical sense. One common expression describing such children was "early ripe, early rot," suggesting that if mental ability developed too fast at an early age, not enough would remain for the later years. To objectively study the experience of bright children through the years, Terman ran one of the most famous studies in psychology's history. By identifying highly intelligent children and observing them over a long period of time, Terman could evaluate his belief that children with high IQs are more successful in life than children with lower IQs.

As his first step, Terman defined genius as a score of 135 or higher on his test. Next, he and his colleagues administered the test to thousands of California schoolchildren, and he isolated 1,528 gifted children (856 boys and 672 girls). The average chronological age of the group was 11, and the average IQ of the group was 151. Learning everything he could about his subjects—including their interests, family history, health, physical characteristics, and personality—Terman wanted to study the experiences of group members as they matured through the years. He began his study in 1921 and reported the first results in *Genetic Studies of Genius* (1926). The term *genetic* can have two meanings. First, it can mean "developmental." When the term is being used in this sense, a genetic study is one that traces how something varies as a function of maturation, or time. Second, the term *genetic* can refer to the genes or chromosomes responsible for various traits. Terman used the term in the developmental sense.

Terman found that the children in his study (who referred to themselves as "Termites") had parents with above-average educational backgrounds, that the children had learned to read at an early age, that they participated in a wide range of activities, and that their schoolwork was usually excellent. All of this might have been expected; the major question was how these children would fare as they became older. Terman did follow-up studies in 1927–1928, when the average age of the group was about 16, and again in 1939–1940, when the average age was about 29. These studies indicated that test

scores were still in the upper 1% of the general population, that members of the group still participated in a wide variety of activities and excelled in most of them, and that they were still outstanding academically. Seventy percent of the men and 67% of the women had finished college, and 56% of the men and 33% of the women had gone on for at least one advanced degree. All these percentages were far higher than for the general population at the time.

In 1947, Terman appeared on the radio show "Quiz Kids." On the show, bright, healthy children were asked extremely difficult questions to which they typically knew the answers. Terman appeared on the program because he felt that it was responsible for correcting many of the misconceptions about gifted children. In fact, Terman thought the program did more in that regard than his own work had done.

> I have devoted a good part of my life to research on children of high I.Q. . . . But despite all my investigations, and those of others, many people continued to think of the brainy child as a freak—physically stunted, mentally lop-sided, nonsocial, and neurotic. Then came the Quiz Kid program, featuring living specimens of highly gifted youngsters who were obviously healthy, wholesome, well-adjusted, socially minded, full of fun, and versatile beyond belief. . . . Result: the program has done more to correct popular misconceptions about bright children than all the books ever written. (quoted in Minton, 1988, pp. 222–223)

It is probably best that it was not until after Terman's death that it was discovered that the "Quiz Kids" were often given their questions in advance of the show (Minton, 1988, p. 223).

The final follow-up in which Terman participated took place in 1950–1952, and it showed that members of the group continued to excel in most of the categories studied. By now, many members of the group had attained prominence as doctors, lawyers, teachers, judges, engineers, authors, actors, scientists, and businesspeople. Upon Terman's death in 1956, the directorship of the investigation was taken over by Robert R. Sears, a Stanford professor who was one of Terman's Termites. In the 1970s, two other Stanford professors were added to the investigation team, Lee J. Cronbach (another Termite) and Pauline S. Sears, Robert's wife. The most recent data collection phase of the study was completed in 1986 under the supervision of Robert Sears and Albert Hastorf.

The group of gifted individuals identified by Terman in 1921 has been studied intensely for more than 60 years, and the study continues. For example, Tomlinson-Keasey and Little (1990) examined 1,069 of the original 1,528 Termites and found that, although generally successful and well-adjusted, some were more successful and well-adjusted than others. Tomlinson-Keasey and Little (1990) isolated the variables related to differential achievement and personal adjustment levels so that they may be used to predict and enhance the achievement and adjustment of other gifted individuals. Friedman, Tucker, Schwartz, Tomlinson-Keasey, Martin, Wingard, and Criqui (1995) examined the backgrounds of a sample of Terman's Termites who were deceased as of 1991. They found that certain psychosocial and behavioral variables were significant predictors of premature mortality, such as parental divorce during childhood, unstable marriage patterns during adulthood, certain childhood personality characteristics (such as being unconscientious), psychological instability in adulthood, and unhealthy habits (such as excessive smoking and drinking).

For the researchers involved in Terman's longitudinal study, the primary results were clear: *The gifted child becomes a gifted adult*. Terman's study put to rest many mistaken beliefs about gifted children, but it left unanswered the question of whether "giftedness" was inherited or the result of experience. Terman believed strongly that it was inherited, but subsequent researchers have shown that many of Terman's results can be explained by taking into account the group members' experiences. How much of intelligence is genetically determined and how much is environmentally determined are still hotly contested questions in psychology. Most modern researchers, however, concede that both factors are important. In any case, Terman's longitudinal study of gifted individuals clearly showed that individuals who score high on so-called measures of intelligence early in life do not deteriorate later in life. In fact, his results

showed that those who fare best in youth also tend to fare best as mature adults.

Leta Stetter Hollingworth

For Terman, the primary purpose of mental testing was the identification of gifted individuals so that they could be encouraged to reach their full potential and become societal leaders. He believed that a tracking system whereby gifted students were provided educational experiences different from those provided for nongifted children was essential for the survival of democracy. Mainly through the efforts of Terman and his colleagues, intelligence testing and ability grouping were common practices in U.S. elementary schools by 1930. However, although strongly recommending a differentiated school curriculum, Terman had no specific recommendations concerning the educational methods that should be adopted in meeting the needs of intellectually superior children. It was **Leta Stetter Hollingworth** (1886–1939) who was primarily concerned with developing educational strategies that would ensure the developmental well-being of gifted students.

Born Leta A. Stetter, Hollingworth attained her bachelor's degree from the University of Nebraska. In 1908, Hollingworth, who had been teaching school in Nebraska, accompanied her husband, Harry, to New York where he had been hired as a psychology instructor at Barnard College. Harry L. Hollingworth himself went on to gain considerable prominence as a psychologist. Leta Hollingworth intended to continue teaching in New York but discovered that the city had a policy of not employing married women as teachers. She decided to enroll as a graduate student at Columbia University where she took courses from Edward L. Thorndike (see chapter 11), who became her advisor. It was through Thorndike that she developed an interest in psychological testing. However, Hollingworth was also interested in the many misconceptions about women that were prevalent at the time. To her surprise, Thorndike agreed to supervise her dissertation on "Functional Periodicity," which investigated the notion that women are psychologically impaired during menstruation. She found no evidence for such impairment (Hollingworth, 1914).

ARCHIVES OF THE HISTORY OF AMERICAN PSYCHOLOGY

Leta Stetter Hollingworth

Hollingworth also challenged the widely accepted beliefs that intelligence is largely inherited and that women are intellectually inferior to males. At the time, Thorndike was among those who shared these beliefs. Hollingworth (1940) believed that women reach positions of prominence less often than males not because of intellectual inferiority but because of the social roles assigned to them.

> Why do we not consider first the established, obvious inescapable fact that women bear and rear the race, and that this has always meant, and still means that nearly 100% of their energy is consumed in the performance and supervision of domestic and allied tasks, a field where eminence is impossible. No one knows who is the best housekeeper in America. Eminent housekeepers do not and cannot exist. If we discuss at all the matter of sex differences in achievement, we should consider first the most obvious conditioning factors. Otherwise our discussion is futile scientifically. (p. 16)

Thorndike later modified his views on intelligence to stress nurture more than nature. Hollingworth believed that she was at least partially responsible for his revised beliefs. She also discussed with Terman her belief that more men than women are classified as gifted not because of differential intellectual abilities but because of social factors. Terman did eventually modify his nativistic position concerning gender differences in intelligence allowing for social influences, but he maintained his belief that intelligence was primarily genetically determined.

After receiving her master's degree in 1913, Hollingworth worked for a while as a clinical psychologist at the New York City Clearing House for Mental Defectives, where she administered Binet tests. She then worked at Bellevue Hospital as a clinical psychologist until attaining her doctorate from Columbia University in 1916. Soon thereafter she became a professor of education at Teachers College, Columbia University. Her work at the Clearing House made her realize that there were as many myths about so-called mentally defective individuals as there were about women. For example, she found that many individuals were classified as "defective" but in reality they were manifesting social and personal adjustment problems. In a series of books, Hollingworth attempted to correct this and related problems: *The Psychology of Subnormal Children* (1920); *Special Talents and Defects: Their Significance for Education* (1923); and *The Psychology of the Adolescent* (1928). The latter replaced G. Stanley Hall's text (see chapter 11) as the standard in the field.

Hollingworth next concentrated her attention on the education of gifted children. She observed that simply classifying a child as gifted is not enough. By emphasizing abstract test scores or group characteristics, the needs of individual students are often overlooked. As an example, she described the experience of a gifted 8-year-old girl named Jean who typically finished her assignments more quickly than her classmates. The teacher's reaction to this problem was to have Jean write digits in a book over and over until her classmates could finish their assignments.

> Jean had with her the copy books in which she had been writing for the past year, one digit after another by the hour. Jean's mother said, "she can't stand the numbers any longer. Her hand gets stiff." I wish you could see the thousands of rows of digits obediently inscribed by this intelligent child, till finally she burst out crying, "I can't stand the numbers anymore." (Hollingworth, 1940, p. 127)

Correcting such mistreatment of gifted children occupied Hollingworth for the rest of her career. In 1926, she published *Gifted Children*, which became the standard text in schools of education for many years, and *Children Above 180 I.Q.* was published posthumously in 1942. (For interesting biographical sketches of Hollingworth see Benjamin, 1975; and Shields, 1975, 1991).

Intelligence Testing in the Army

Robert M. Yerkes

Robert M. Yerkes (1876–1956) was the first-born son of a rural Pennsylvania farm family. He was disillusioned by farm life, however, and dreamed of becoming a medical doctor. During his college years, Yerkes lived with an uncle for whom he did chores in return for tuition to Ursinis College. After Ursinis, Yerkes went to Harvard where he became interested in animal behavior. Obtaining his doctorate in 1902, he remained at Harvard as a faculty member. With his friend, John B. Watson (see chapter 12), who was then at Johns Hopkins University, Yerkes established comparative psychology in the United States. In recognition of his ultimate success, Yerkes was elected president of the American Psychological Association (APA) in 1917.

As a student, Yerkes had to borrow considerable money, and his faculty post at Harvard did not pay very much. This meant that he had to take part-time jobs in order to survive financially. Thus, in 1912 he took the job as the director of psychological research at the Boston State Psychopathic Hospital; it was here that Yerkes had his first experience with intelligence testing. At the hospital, the Binet–Simon scale was being explored as an instrument to aid clinical diagnoses. One of Yerkes's Harvard professors, and now his friend and colleague, was the biologist Charles Davenport who corresponded with

ARCHIVES OF THE HISTORY OF AMERICAN PSYCHOLOGY

Robert M. Yerkes

Galton and was a leader in the U.S. eugenics movement. Yerkes, too, became a strong advocate of eugenics. Yerkes became increasingly involved in testing at the Boston Psychopathic Hospital, at the expense of his work in comparative psychology.

Yerkes's "contribution" to intelligence testing was his suggestion that all individuals be given all items on the Binet–Simon test and be given points for the items passed. Thus, a person's score would be in terms of total points earned instead of an IQ. This removes age as a factor in scoring. The traditional procedure followed in administering the Binet–Simon scale was to locate the range of tests appropriate for a given individual. For example, if a 7-year-old was being tested, the tests appropriate for that age would be given. If the child missed any of those tests, the tests appropriate for the next lowest age (6) would be administered. If, in this case, the child initially passed all tests appropriate for the 7-year-old level, tests from the 8-year-old level would be administered, and so forth until the child began to fail tests. In other words, using age as a frame of reference, the testing procedure was customized for each child. Yerkes's "point-scale" procedure rendered all of this unnecessary. Yerkes did point out, however, that point norms could be established for various ages or for any group one wanted to compare. Yerkes believed that, besides being easier to administer, point scores were more amenable to statistical analyses than IQ scores were. Also, because with point scores all individuals took the same tests without regard to their age or level, Yerkes's method was conducive to group testing, whereas the Binet–Simon test had to be given to one person at a time. Soon Yerkes would see his method tried on a level he never dreamed possible.

The army testing program. When the United States entered World War I in 1917, Yerkes was president of the APA. He called a special meeting of the association to determine how psychologists could help in the war effort. It was decided that psychologists could contribute by devising ways of selecting and evaluating recruits into the armed forces. Upon Goddard's invitation, a small group of psychologists, including Yerkes and Terman, went to the Vineland School to develop psychological tests that were then tried at various army and navy bases. Because the results were encouraging, Yerkes was made an army major and given the job of organizing a testing program for the entire army (the navy rejected the idea). The goals of the program were to identify those with mental deficiencies, to classify men in terms of their intelligence level, and to select individuals for special training—for example, to become officers. Yerkes believed that, to be effective, the test used had to be a group test rather than an individual test, had to measure "native" intelligence, and had to be easy to administer and score. Using Yerkes's point-score method of scoring, the group created a test that met these criteria but found that 40% of the recruits could not read well enough to take the test. The group solved the problem by creating two forms of the test: the *Army Alpha* for literate individuals and the *Army Beta* for illiterate individuals or for those who spoke and read a language other than English.

The war ended in 1918, and the testing program was terminated in 1919, by which time over 1.75 million individuals had been tested. Many people claimed that the army testing program had demonstrated psychology's practicality, but the evidence does not support such a contention. Samelson (1977) reports that only .005% of those tested were recommended for discharge as mentally unfit, and in many cases the army ignored the recommendations. Also, if the army had perceived the testing program as effective, it would not have terminated the program so soon after the war ended. In his evaluation of the army testing program under Yerkes's leadership, Reed (1987) reached the following conclusion:

> In retrospect, Yerkes's greatest coup as a scientific bureaucrat and promoter was not in getting the Surgeon General to find a place for psychologists in the army, although that was a notable accomplishment, nor in writing tests, recruiting several hundred officers and technicians, and administering examinations to over 1.7 million individuals, despite fierce competition for resources and status from army officers and psychiatrists, although that too was a notable accomplishment. His most remarkable achievement was the myth that the army testing program had been a great practical success and that it provided a "goldmine" of data on the heritability of intelligence. (p. 84)

The Deterioration of National Intelligence

The use of the Army Alpha and Beta tests rekindled concern about the deterioration of the nation's intelligence level. About half of the white males tested in the army had native intelligence equal to that of a 13-year-old or lower, and the situation was even worse for black soldiers. Goddard's response was that people with low mental ability should not be allowed to vote. Along with Goddard, Terman and Yerkes were very concerned about the deterioration of the nation's intelligence, which they believed was caused by immigration and the fact that intellectually inferior individuals were reproducing faster than normal or above-normal individuals.

As was common at the time, Yerkes (1923) believed that many of the nation's ills were being caused by people of low intelligence and that immigration policies were only aggravating the problem.

> By some people meagre intelligence in immigrants has been considered an industrial necessity and blessing; but when all the available facts are faced squarely, it looks more like a burden. Certainly the results of psychological examining in the United States Army establish the relation of inferior intelligence to delinquency and crime, and justify the belief that a country which encourages, or even permits, the immigration of simple-minded, uneducated, defective, diseased, or criminalistic persons, because it needs cheap labor, seeks trouble in the shape of public expense.
>
> It might almost be said that whoever desires high taxes, full almshouses, a constantly increasing number of schools for defectives, of correctional institutions, penitentiaries, hospitals, and special classes in our public schools, should by all means work for unrestricted and non-selective immigration. (p. 365)

Fortunately, the extremely nativistic position that Goddard, Terman, and Yerkes represented did not go unchallenged. More and more, people realized that performance on so-called intelligence tests could be at least partially explained by such factors as early experience and education. Rather than simply measuring native intelligence, the tests were apparently also measuring personal achievement and the influence of life's circumstances. It followed that the more privileged a person was in terms of enriching experiences and education, the higher his or her scores would be on so-called intelligence tests.

The recent book *The Bell Curve: Intelligence and Class Structure in American Life* (1994), by Richard J. Herrnstein and Charles Murray, reflects many of the earlier beliefs about intelligence accepted by Galton, Cattell, Goddard, Terman, and Yerkes. Herrnstein and Murray organize their book around six conclusions, or points, about intelligence that are "beyond dispute." By "beyond dispute" they mean

> That if you gathered the top experts on testing and cognitive ability, drawn from all points of view, to

> argue over these points, away from television cameras and reporters, it would quickly become apparent that a consensus already exists on all of the points, in some cases amounting to near unanimity. (p. 23)

Here are the six points:

1. There is such a thing as a general factor of cognitive ability on which human beings differ.
2. All standarized tests of academic aptitude or achievement measure this general factor to some degree, but IQ tests expressly designed for that purpose measure it most accurately.
3. IQ scores match, to a first degree, whatever it is that people mean when they use the word *intelligent* or *smart* in ordinary language.
4. IQ scores are stable, although not perfectly so, over much of a person's life.
5. Properly administered IQ tests are not demonstrably biased against social, economic, ethnic, or racial groups.
6. Cognitive ability is substantially heritable, apparently no less than 40 percent and no more than 80 percent. (pp. 22–23)

Not on the list, but featured in the book, is the contention that in the United States the best jobs and the highest income tend to go to the most intelligent individuals, the "cognitive elite." The less intellectually endowed are doomed to menial labor in our information-based economy, if they can find work at all. Couple this with the fact that (according to Herrnstein and Murray) intelligence is largely inherited and we have a major problem—that is, an economic class structure based on inherited intelligence. The authors do not offer a solution to the problem, but others have. Galton, Cattell, Goddard, Terman, and Yerkes all described a similar problem, and all suggested that the solution was to somehow discourage less intelligent individuals from reproducing. There is nothing new, and much that is quite old, in Herrnstein and Murray's list. In fact, each of their "indisputable points" about intelligence has been, and is, hotly disputed (see, for example, Azar, 1994, 1995a, 1995b; DeAngelis, 1995; Jacoby & Glauberman, 1995; *The New Republic*, 1994).

Currently there is little agreement even on an adequate definition of intelligence. When 24 prominent researchers in the field of intelligence were asked to define intelligence they provided 24 different definitions (Sternberg & Detterman, 1986). After reviewing which of the many notions concerning intelligence have scientific support and which do not, Neisser et al. (1996) conclude:

> In a field where so many issues are unresolved and so many questions unanswered, the confident tone that has characterized most of the debate on these topics is clearly out of place. The study of intelligence does not need politicized assertions and recriminations; it needs self-restraint, reflection, and a great deal more research. The questions that remain are socially as well as scientifically important. There is no reason to think them unanswerable, but finding the answers will require a shared and sustained effort as well as the commitment of substantial scientific resources. Just such a commitment is what we strongly recommend. (p. 97)

Summary

Evolutionary theory has existed in one form or another since the time of the early Greeks. The biblical account of the origin of species silenced evolutionary theory for many centuries, but by the 18th century there was again speculation about the evolutionary process. Lamarck claimed that traits acquired during an individual's lifetime that were conducive to survival were passed on to the individual's offspring. Spencer originally followed Lamarck by saying that frequently used associations were passed on to offspring in the form of reflexes and instincts. Later, Spencer accepted Darwin's version of evolutionary