History of Psychology

Session 10: Psychological testing

Loreen Tisdall, Center for Cognitive and Decision Sciences
November 25, 2024

Session information

Sessions take place Mondays, 8.15-9.45, Chemie, Organische, Grosser Hörsaal OC.

#	Date	Topic	Instructor
1	23.09.2024	Session 1: Introduction	Tisdall
2	30.09.2024	Session 2: Pre-psychology	Mata
3	07.10.2024	Session 3: The birth of psychology	Mata
4	14.10.2024	Session 4: Psychoanalysis	Mata
5	21.10.2024	Session 5: Behaviorism	Mata
6	28.10.2024	Session 6: Gestalt psychology	Mata
7	04.11.2024	Session 7: Cognitive psychology	Mata
8	11.11.2024	Session 8: Psychology today	Tisdall
9	18.11.2024	Session 9: Psychotherapy research	Tisdall
10	25.11.2024	Session 10: Psychological testing	Tisdall
11	02.12.2024	Session 11: Decision science	Tisdall
12	09.12.2024	Session 12: What kind of science is psychology?	Mata

Learning Objectives for Today

- Identify the main events concerning the development of psychological testing and “psychometrics”, particularly intelligence testing.
- Identify and describe the main statistical methods that allowed the flourishing of psychological testing, such as correlation and factor analysis
- Identify main issues of controversy (e.g., nature-nurture, eugenics) and areas of application of psychological testing (e.g., prediction of occupational performance)

Psychological testing

- **Psychological testing** refers to the administration of psychological tests.
- A **psychological test** is "an objective and **standardized** measure of a sample of behavior". The term **sample of behavior** refers to an individual's performance on tasks that have usually been prescribed beforehand. The samples of behavior that make up a paper-and-pencil test, the most common type of test, are a series of items.
- Performance on these items produces a **test score or a category membership**. A score on a well-constructed test is believed to reflect a **psychological construct** such as achievement in a school subject, cognitive ability, aptitude, emotional functioning, or personality. Differences in test scores are thought to reflect **individual differences** in the construct the test is supposed to measure.
- Most tests have **norms or standards** (e.g., norm- or criterion-referenced tests) by which the results can be used to **predict other, more important (non-test) behaviors**
- The technical term for the science behind psychological testing is **psychometrics**, which focuses on the development, validation and application of psychological tests.

https://en.wikipedia.org/wiki/Psychological_testing; https://en.wikipedia.org/wiki/Psychometrics
Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook

Psychological testing

TABLE 1.1 The Main Types of Psychological Tests

Intelligence Tests: Measure an individual's ability in relatively global areas such as verbal comprehension, perceptual organization, or reasoning and thereby help determine potential for scholastic work or certain occupations.
Aptitude Tests: Measure the capability for a relatively specific task or type of skill; aptitude tests are, in effect, a narrow form of ability testing.
Achievement Tests: Measure a person's degree of learning, success, or accomplishment in a subject or task.
Creativity Tests: Assess novel, original thinking and the capacity to find unusual or unexpected solutions, especially for vaguely defined problems.
Personality Tests: Measure the traits, qualities, or behaviors that determine a person's individuality; such tests include checklists, inventories, and projective techniques.
Interest Inventories: Measure an individual's preference for certain activities or topics and thereby help determine occupational choice.
Behavioral Procedures: Objectively describe and count the frequency of a behavior, identifying the antecedents and consequences of the behavior.
Neuropsychological Tests: Measure cognitive, sensory, perceptual, and motor performance to determine the extent, locus, and behavioral consequences of brain damage.

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook

From mental tests to latent constructs

Three major developments had to happen / take place for the rise of psychological testing as we know it today:

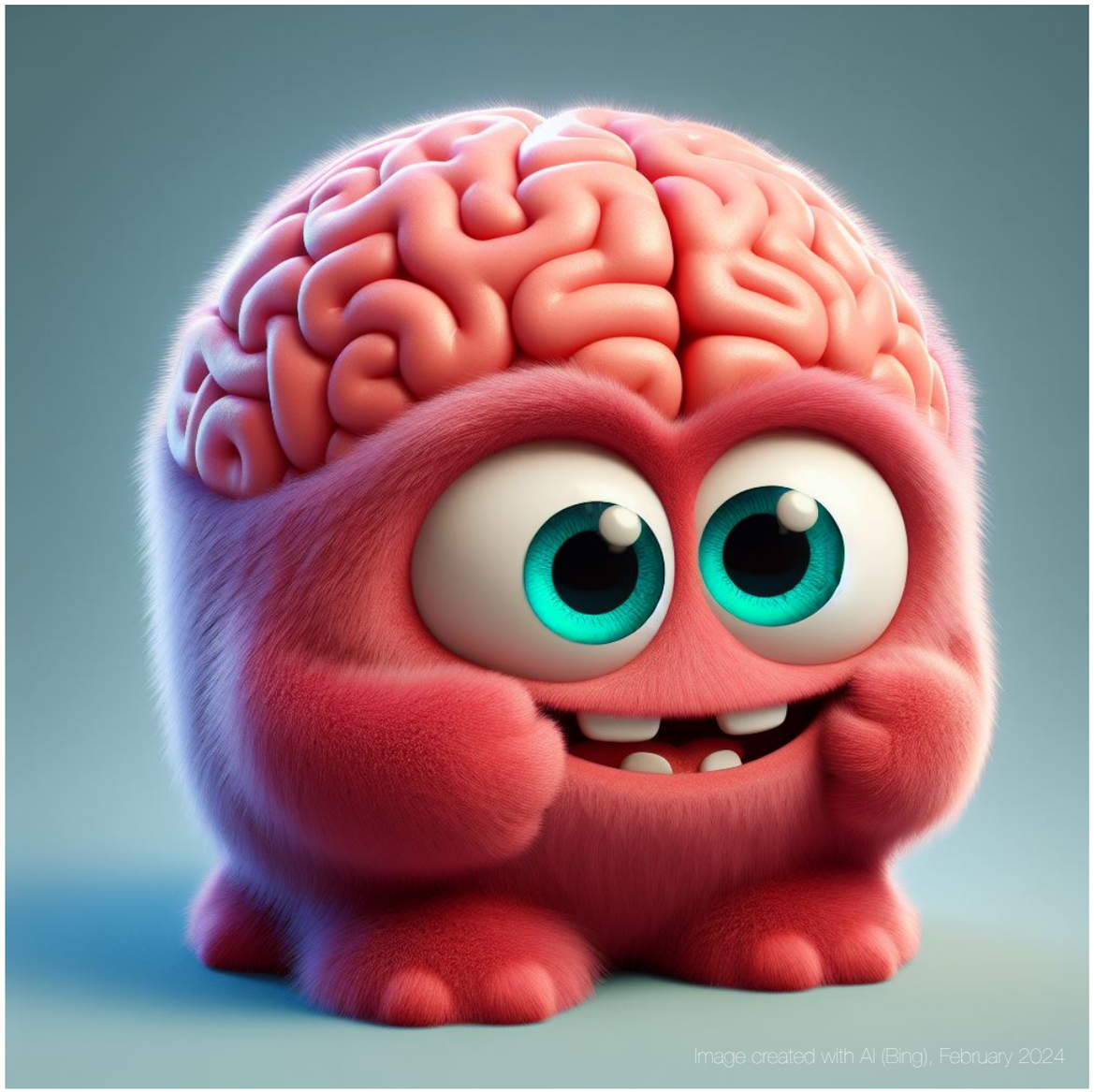
1) Psychologists had to ***devise a set of plausible measure*** to capture mental capacity

2) Psychologists had to find ***new forms of comparing individuals*** (e.g., norms, scales, scoring rules)

3) Psychologists had to develop a ***new methodological toolbox*** to assess the links between test scores and the criteria the new measures aimed to predict

➡ By developing new measures and discovering new statistical procedures to treat mental products and behavior as measurable quantities, psychological testing created a long-lasting symbiosis between statistical and psychological theory.

Your turn!



In which context / for which purpose do you think we find the most frequent use of psychological testing?

What are common aspects of (past) definitions of intelligence?

What were the main structural models of intelligence?

Which ethical concerns surround intellectual assessment?

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

PHASE III (1920s - 1940s)

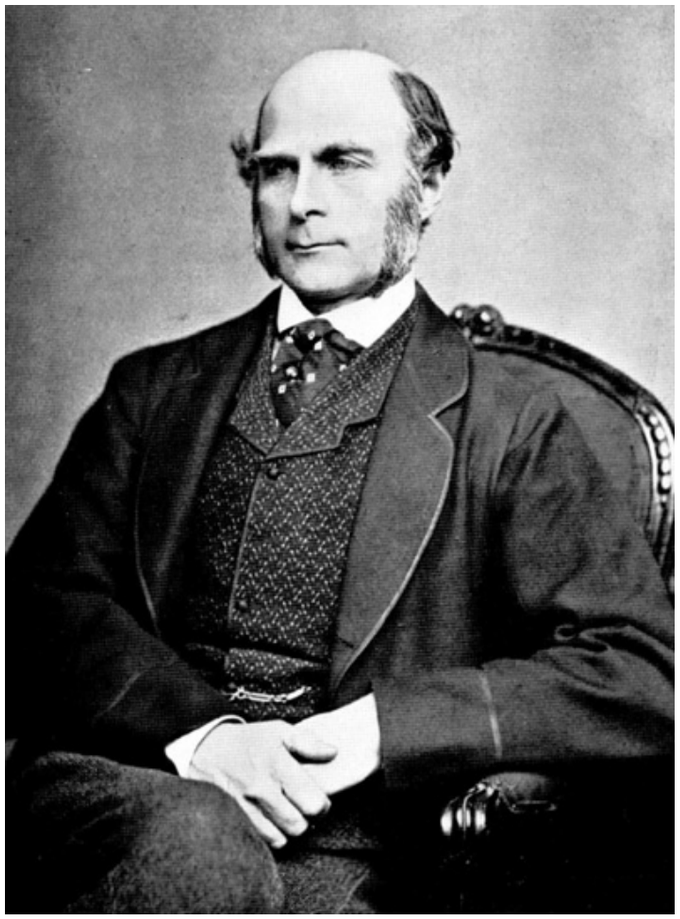
Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

Francis Galton



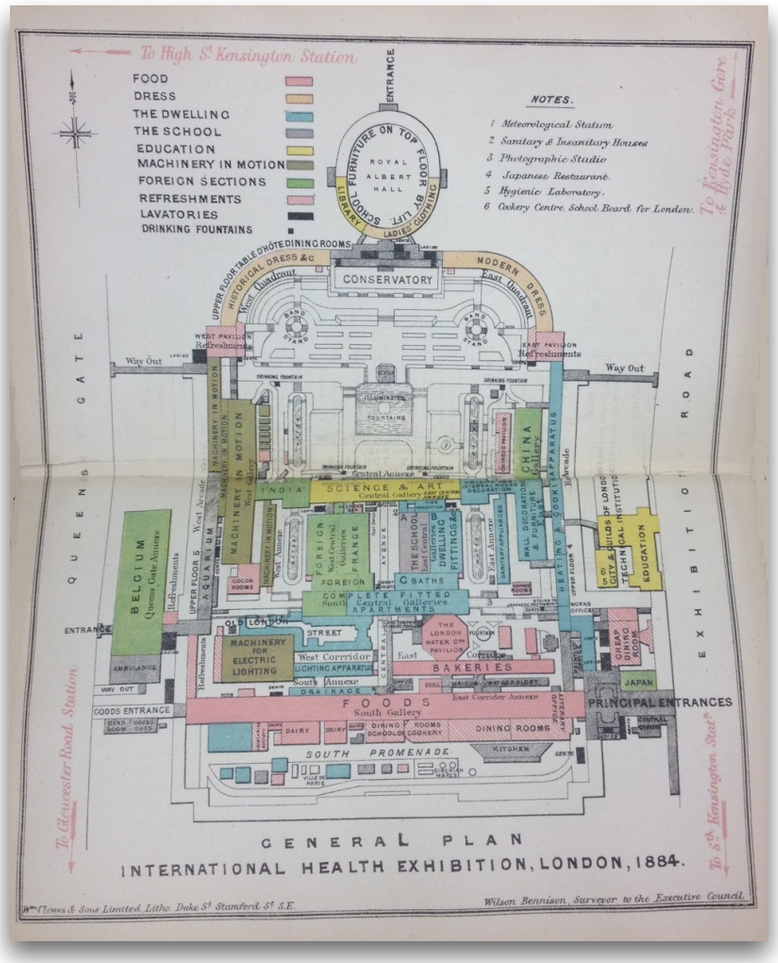
By not stated - Scanned from Karl Pearson's The Life, Letters, and Labors of Francis Galton., Public Domain, https://commons.wikimedia.org/w/index.php?curid=90718

- 1822 Birth (Charles Darwin's half-cousin, same grandfather Erasmus Darwin)
- Studies Medicine in London
- 1844 Travels through Europe and Africa
- 1869 *Hereditary Genius (*10 years after the publication of Darwin's *Origins)*
- 1875 *The history of twins*
- 1882 Center for Mental Tests in London
  - Founded psychometrics
  - Tests for visual acuity, reaction times, etc.
- 1883 *Inquiries into Human Faculty and Its Development*
- 1888 independently invents correlation
- 1909 Knighted
- 1911 Death

Fancher, R. E. (2009). Scientific cousins: the relationship between Charles Darwin and Francis Galton. *American Psychologist*, *64*(2), 84.https://psycnet.apa.org/record/2009-01602-015

Francis Galton: Brass, glass, & big data

Early measures of the new science of the mind were heavily influenced by the brass and glass era of psychology (i.e., these were geared towards physiological measures)



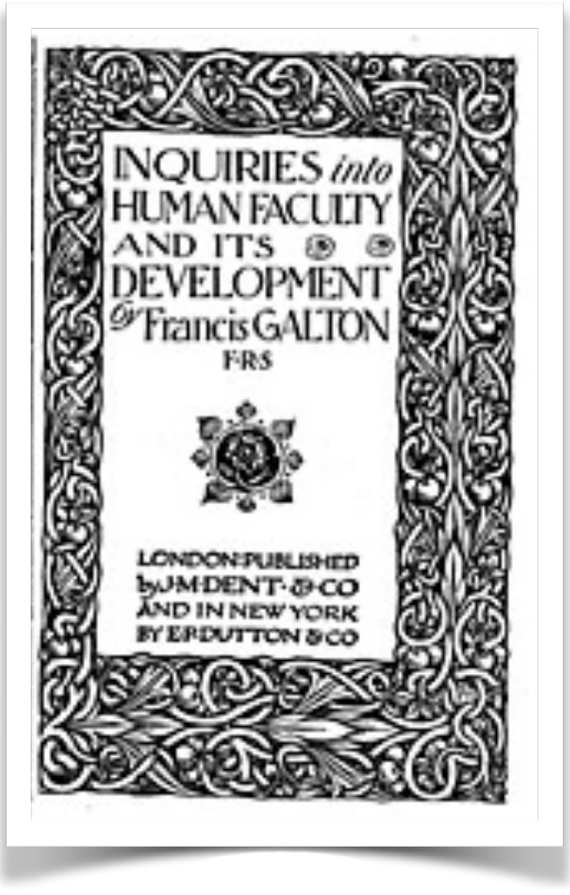
https://blogs.lshtm.ac.uk/library/2016/07/07/international-health-exhibition-1884/#:~:text=In 1884 London hosted an,displays and official education programme

General plan of the International Health Exhibition 1884

- Organised by Francis Galton, the 1884 International Health Exhibition in London "covered everything that promoted the preservation of good health in the populace through displays and official education programme" —> *"anthropometric laboratory"*
- Included a large set of measures of physical attributes, including eye and hair colour, sight, hearing, height and weight, grip strength, lung capacity, arm span, etc.
- 4 million visitors between 8 May and 30 October 1884, and within one year of its opening, Galton had collected data on close to 10'000 individuals
- Used the data to construct normative distributions and was concerned with validation

Francis Galton: Founder of Eugenics

- In his 1883 book *Inquiries into Human Faculty and Its Development*, Galton coined the term eugenics, but he had already written extensively about this idea in earlier works
- The idea of eugenics to produce better human beings has existed at least since Plato, who suggested selective mating to produce a guardian class. Galton, however, coined the term to label the discipline aimed at improving the genetic quality of the human population.
- Today it is used to label a social philosophy advocating the improvement of human traits through the promotion of higher rates of sexual reproduction for people with desired traits (positive eugenics), or reduced rates of sexual reproduction or sterilization of people with less-desired or undesired traits (negative eugenics), or both.



https://en.wikipedia.org/wiki/Inquiries_into_Human_Faculty_and_Its_Development

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

PHASE III (1920s - 1940s)

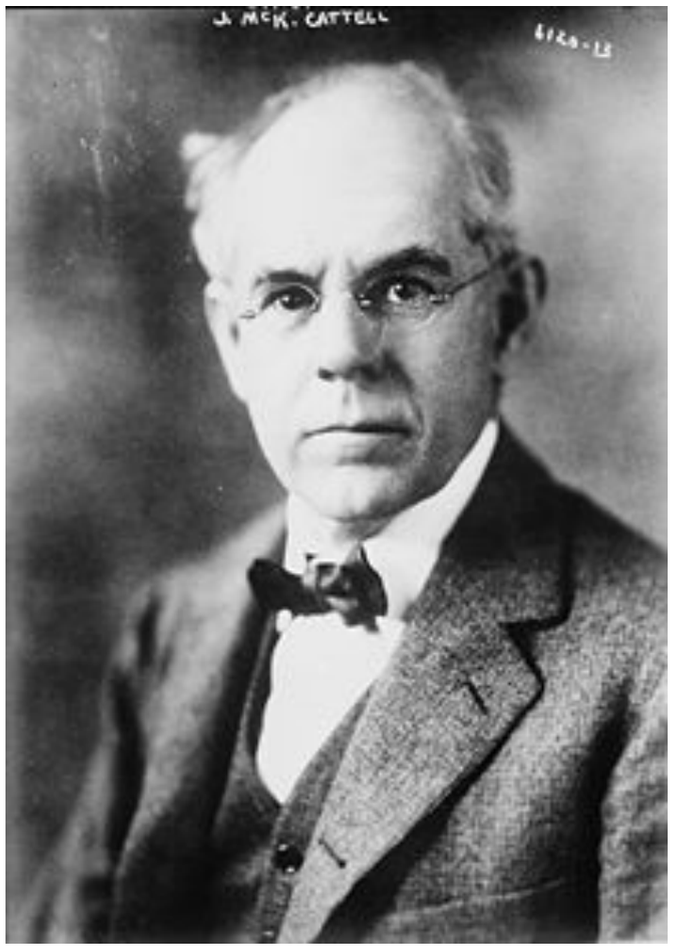
Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

James McKeen Cattell



https://en.wikipedia.org/wiki/James_McKeen_Cattell#/media/File:James_McKeen_Cattell.jpg

- 1860 - 1944
- First professor of psychology in the US and “the dean of American science”
- Having spent time during his doctoral studies with Galton in London and also with Wilhelm Wundt in Leipiz, he was inspired by such anthropometric efforts
- On his return to the US (late 1880s), he established his own psychological laboratory
- His task battery mixed Galton’s physical attribute measures with Wundt’s psychophysical approaches (e.g., reaction times for auditory stimuli, memory recall of random letter sequences, discrimination between weights) —> used in large-scale testing (e.g., university students)

Pillsbury, W. B. (1947). Biographical Memoir of James Mckeen Cattell (1860-1944). *National Academy of Sciences of the United States of America. Biographical Memoirs*, *25*(1). https://www.nasonline.org/wp-content/uploads/2024/06/cattell-james-m.pdf

James McKeen Cattell

V.—MENTAL TESTS AND MEASUREMENTS.

By Prof. J. McK. Cattell.

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. Individuals, besides, would find their tests interesting, and, perhaps, useful in regard to training, mode of life or indication of disease. The scientific and practical value of such tests would be much increased should a uniform system be adopted, so that determinations made at different times and places could be compared and combined. With a view to obtaining agreement among those interested, I venture to suggest the following series of tests and measurements, together with methods of making them.[1]

[1] Mr. Francis Galton, in his Anthropometric Laboratory at South Kensington Museum, already uses some of these tests, and I hope the series here suggested will meet with his approval. It is convenient to follow Mr. Galton in combining tests of body, such as weight, size, colour of eyes, &c., with psychophysical and mental determinations, but these latter alone are the subject of the present discussion. The name

The first series of ten tests is made in the Psychological Laboratory of the University of Pennsylvania on all who present themselves, and the complete series on students of Experimental Psychology. The results will be published when sufficient data have been collected. Meanwhile, I should be glad to have the tests, and the methods of making them, thoroughly discussed.

The following ten tests are proposed:

I. Dynamometer Pressure.
II. Rate of Movement.
III. Sensation-areas.
IV. Pressure causing Pain.
V. Least noticeable difference in Weight.
VI. Reaction-time for Sound.
VII. Time for naming Colours.
VIII. Bi-section of a 50 cm. line.
IX. Judgment of 10 seconds time.
X. Number of Letters remembered on once Hearing.

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

Correlation:
Pearson's product-moment correlation

Karl Pearson (1857 – 1936)
Influential English mathematician and biostatistician. Credited with establishing the discipline of mathematical statistics, theories of social Darwinism and eugenics. Founded the first university statistics department at University College London. Pearson was a protégé and biographer of Sir Francis Galton.

https://en.wikipedia.org/wiki/Karl_Pearson

Correlation is the standardized covariance between two variables.

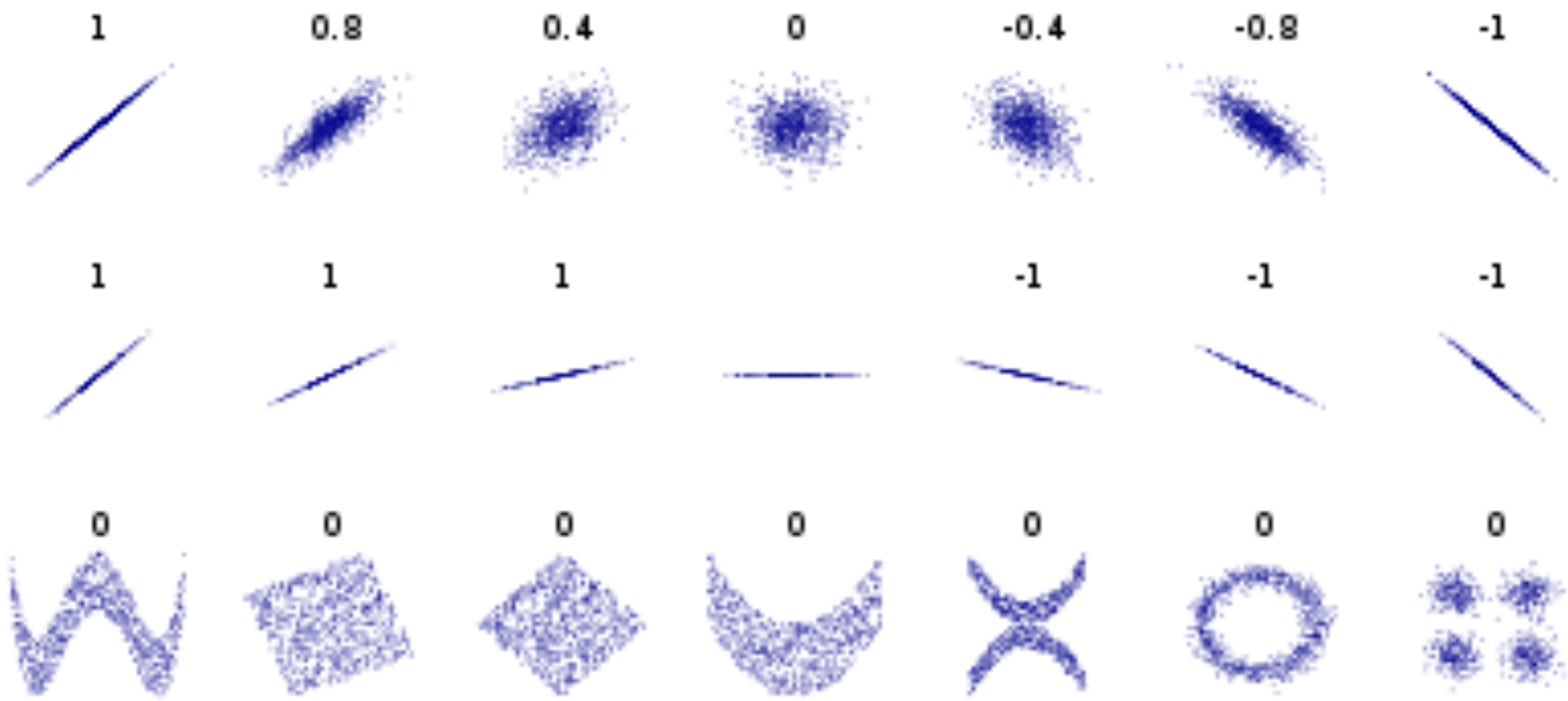
Why is this useful, particularly in the context of psychological testing?

Correlation $\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{s_X s_Y}$

Covariance $\operatorname{cov}(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - E(X))(y_i - E(Y)).$

Standard Deviation $s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}.$

Pearson's correlation: A demonstration



Several sets of (x, y) points, with the correlation coefficient of x and y for each set. Note that the correlation reflects the strength and direction of a linear relationship (top row), but not the slope of that relationship (middle), nor many aspects of nonlinear relationships (bottom). N.B.: the figure in the center has a slope of 0 but in that case the correlation coefficient is undefined because the variance of Y is zero.

https://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things* (rank order!)

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

PHASE IV (1950s - present)

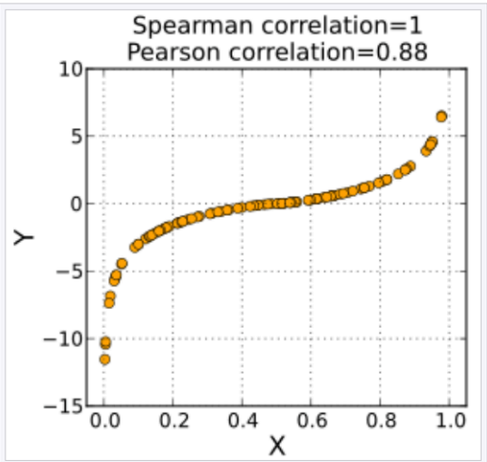
Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

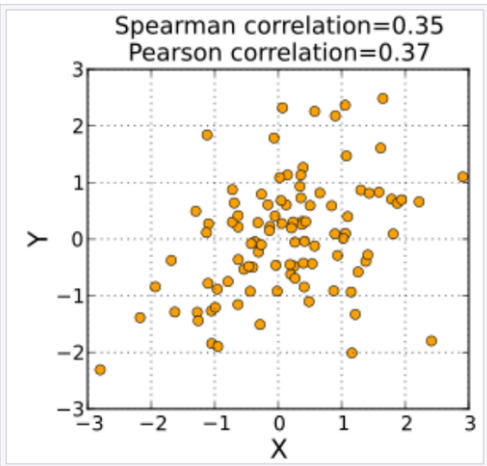
Spearman’s rank-order correlation

Feature	Pearson	Spearman
Data used	Raw values	Ranks
Suitable data	Continuous data that is linear and normally distributed	Ordinal, interval, ratio scales, non-linear monotonic data
Measures	Linear relationships (constant rate of change)	Monotonic relationships (increasing or decreasing trend)
Sensitivity	Highly sensitive to outliers (because it uses raw values)	Less sensitive to outliers because it uses ranks

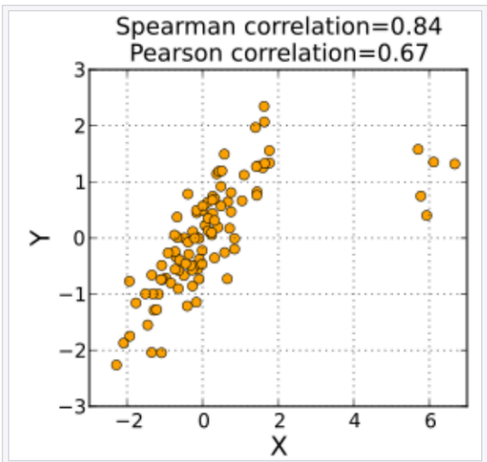
Why is the development of this kind of correlation coefficient useful, particularly in the context of psychological testing?



A Spearman correlation of 1 results when the two variables being compared are monotonically related, even if their relationship is not linear. This means that all data points with greater x values than that of a given data point will have greater y values as well. In contrast, this does not give a perfect Pearson correlation.



When the data are roughly elliptically distributed and there are no prominent outliers, the Spearman correlation and Pearson correlation give similar values.



The Spearman correlation is less sensitive than the Pearson correlation to strong outliers that are in the tails of both samples. That is because Spearman's ρ limits the outlier to the value of its rank.

https://en.wikipedia.org/wiki/Spearman's_rank_correlation_coefficient

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

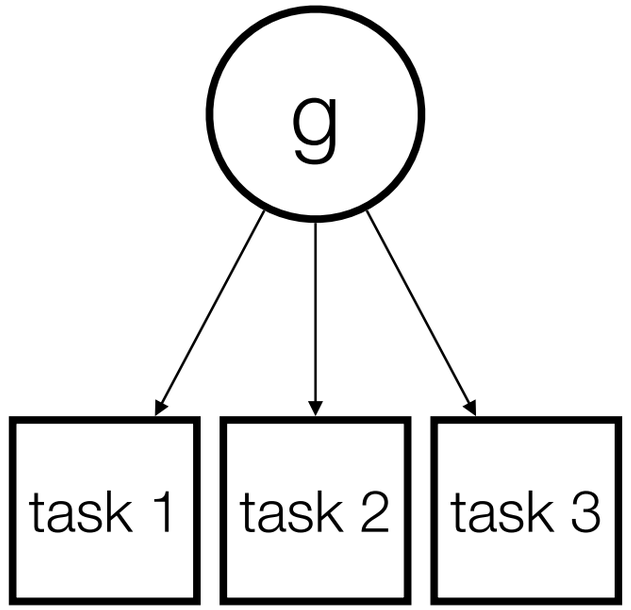
Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

Factor Analysis

A statistical method used to **describe variability among observed, correlated variables** in terms of a potentially **lower number of unobserved variables** called factors.

For example, it is possible that variations in say 3 observed tasks mainly reflect the variations in one unobserved (underlying) variable, such as intelligence. Factor analysis searches for such **joint variations in response to unobserved latent variables**. The observed variables are modelled as linear combinations of the potential factors, plus "error" terms. The information gained about the interdependencies between observed variables can be used later to reduce the set of variables in a dataset.

Factor analysis was **first proposed by Charles Spearman** and its principles are still used widely today whenever there are large numbers of observed variables that are thought to reflect a smaller number of underlying/latent variables.



Reflective model:

- Latent variable (e.g., *g*) drives the correlations among observed variables (as opposed to a formative model)
- Changes in the latent variable should result in changes in all the observed indicators
- The observed indicators are expected to correlate (positive manifold) because they share the same underlying cause (the latent variable)

https://en.wikipedia.org/wiki/Factor_analysis

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)
- **1905:** Binet-Simon scale, 30 ordered items

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

Binet-Simon Scale

- Developed in France by psychologist Alfred Binet and psychiatrist Théodore Simon, first version published in 1905, with revisions later on
- Binet-Simon scale was aimed at assessing a child's general mental development and practical judgment, and was intended to identify children with learning difficulties (spurred on by the introduction of compulsory education)
- In its original form, it was a relatively brief test composed of a heterogeneous group of tasks that took less than an hour to administer

TABLE 1.1 The 1905 Binet-Simon Scale

1. Follows a moving object with the eyes.
2. Grasps a small object which is touched.
3. Grasps a small object which is seen.
4. Recognizes the difference between a square of chocolate and a square of wood.
5. Finds and eats a square of chocolate wrapped in paper.
6. Executes simple commands and imitates simple gestures.
7. Points to familiar named objects, e.g., "Show me the cup."
8. Points to objects represented in pictures, e.g., "Put your finger on the window."
9. Names objects in pictures, e.g., "What is this?" [examiner points to a picture of a sign].
10. Compares two lines of markedly unequal length.
11. Repeats three spoken digits.
12. Compares two weights.
13. Shows susceptibility to suggestion.
14. Defines common words by function.
15. Repeats a sentence of 15 words.
16. Tells how two common objects are different, e.g., "paper and cardboard."
17. Names from memory as many as possible of 13 objects displayed on a board for 30 seconds. [This test was later dropped because it permitted too many possibilities for distraction.]
18. Reproduces from memory two designs shown for 10 seconds.
19. Repeats a longer series of digits than in item 11 to test immediate memory.
20. Tells how two common objects are alike, e.g., "butterfly and flea."
21. Compares two lines of slightly unequal length.
22. Compares five blocks to put them in order of weight.
23. Indicates which of the previous five weights the examiner has removed.
24. Produces rhymes, e.g., "What rhymes with 'school'?"
25. A word completion test based on those proposed by Ebbinghaus.
26. Puts three nouns, e.g., "Paris, river, fortune" (or three verbs) in a sentence.
27. Responds to 25 abstract (comprehension) questions, e.g., "When a person has offended you, and comes to offer his apologies, what should you do?"
28. Reverses the hands of a clock.
29. After paper folding and cutting, draws the form of the resulting holes.
30. Defines abstract words by designating the difference between, e.g., "boredom and weariness."

"There is in intelligence, it seems to us, a fundamental agency the lack or alteration of which has the greatest importance for practical life; that is judgement, otherwise known as good sense, practical sense, initiative, or the faculty of adapting oneself. To judge well, to understand well, to reason well—these are the essential wellsprings of intelligence."

Binet & Simon (1905)

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)
- **1905:** Binet-Simon scale, 30 ordered items
- **1914:** William Stern publishes *The psychological methods of testing intelligence*
- **1914/1916:** Stanford-Binet Scale through translation of Binet Scale by Terman, introduction of IQ.

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

The Stanford-Binet Scales

Alfred Binet
(1857-1911)

French psychologist Binet and physician Simon. The first version of the Binet-Simon scale did not offer a specific way to derive a total score! In later editions, whichever items were passed by 80 to 90 percent of an age group were placed in the appropriate mental level.

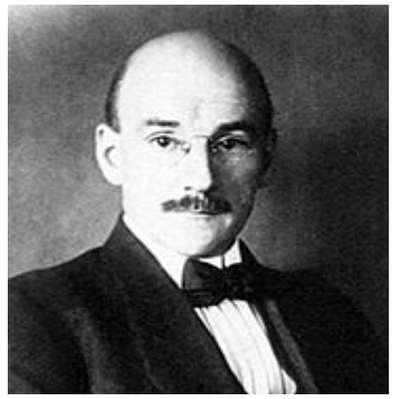
In 1911, a third revision of the Binet-Simon scale appeared which extended the scale into the adult range. Each age level now had exactly five tests. Binet introduced a new scoring that allowed for one-fifth of a year for each subtest passed beyond the basal level.

Théodore Simon
(1872-1961)

William Stern
(1871-1938)

German psychologist, Professor at Duke University. Stern coined the term intelligence quotient and suggested to use a ratio between mental and chronological age as a measure of intelligence.

IQ = Mental Age/
Chronological Age



Henry H. Goddard
(1866-1957)

American psychologist, Director of Research at the Vineland Training School for Feeble-Minded Girls and Boys in Vineland, New Jersey. Goddard was the first to translate the Binet intelligence test into English in 1908, and distributed an estimated 22,000 copies of the translated test across the United States.

Lewis Terman
(1877-1956)

American psychologist, Professor at Stanford University. Created and popularised a version of the scale, naming it Stanford–Binet Intelligence Scale. Terman used the test not only to help identify children with learning difficulties but also identify individuals who had above average levels of intelligence. Terman multiplied the intelligence quotient by 100 to remove fractions.

IQ = 100 x Mental Age/
Chronological Age

https://en.wikipedia.org/wiki/Stanford–Binet_Intelligence_Scales

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)
- **1905:** Binet-Simon scale, 30 ordered items
- **1914:** William Stern publishes *The psychological methods of testing intelligence*
- **1914/1916:** Stanford-Binet Scale through translation of Binet Scale by Terman, introduction of IQ.
- **1917:** Woodworth Personal Data Sheet (WPDS)
- **1917:** Army Alpha & Beta, group tests for military recruits, Committee led by Yerkes (advisors, Terman, Cattell, Goddard)

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

The Army Alpha and Beta

Army Alpha

TEST 3

This is a test of common sense. Below are sixteen questions. Three answers are given to each question. You are to look at the answers carefully; then make a cross in the square before the best answer to each question, as in the sample:

SAMPLE — Why do we use stoves? Because
- ☐ they look well
- ☒ they keep us warm
- ☐ they are black

Here the second answer is the best one and is marked with a cross. Begin with No. 1 and keep on until time is called.

1 If plants are dying for lack of rain, you should
- ☐ water them
- ☐ ask a florist's advice
- ☐ put fertilizer around them

2 A house is better than a tent, because
- ☐ it costs more
- ☐ it is more comfortable
- ☐ it is made of wood

3 Why does it pay to get a good education? Because
- ☐ it makes a man more useful and happy
- ☐ it makes work for teachers
- ☐ it makes demand for buildings for schools and colleges

4 If the grocer should give you too much money in making change, what is the right thing to do?
- ☐ buy some candy of him with it
- ☐ give it to the first poor man you meet
- ☐ tell him of his mistake

5 Why should food be chewed before swallowing?
- ☐ it is better for the health
- ☐ it is bad manners to swallow without chewing
- ☐ chewing keeps the teeth in condition

6 If you saw a train approaching a broken track you should
- ☐ telephone for an ambulance
- ☐ signal the engineer to stop the train
- ☐ look for a piece of rail to fit in

7 If you are lost in a forest in the daytime, what is the thing to do?
- ☐ hurry to the nearest house you know of
- ☐ look for something to eat
- ☐ use the sun or a compass for a guide

8 It is better to fight than to run, because
- ☐ cowards are shot
- ☐ it is more honorable
- ☐ if you run you may get shot in the back

☞ Go to No. 9 above

9 Why are warships painted gray? Because gray paint
- ☐ is cheaper than other colors
- ☐ is more durable than other colors
- ☐ makes the ships harder to see

10 Why should all parents be made to send their children to school? Because
- ☐ it prepares them for adult life
- ☐ it keeps them out of mischief
- ☐ they are too young to work

11 The reason that many birds sing in the spring is
- ☐ to let us know spring is here
- ☐ to attract their mates
- ☐ to exercise their voices

12 Gold is more suitable than iron for making money because
- ☐ gold is pretty
- ☐ iron rusts easily
- ☐ gold is scarcer and more valuable

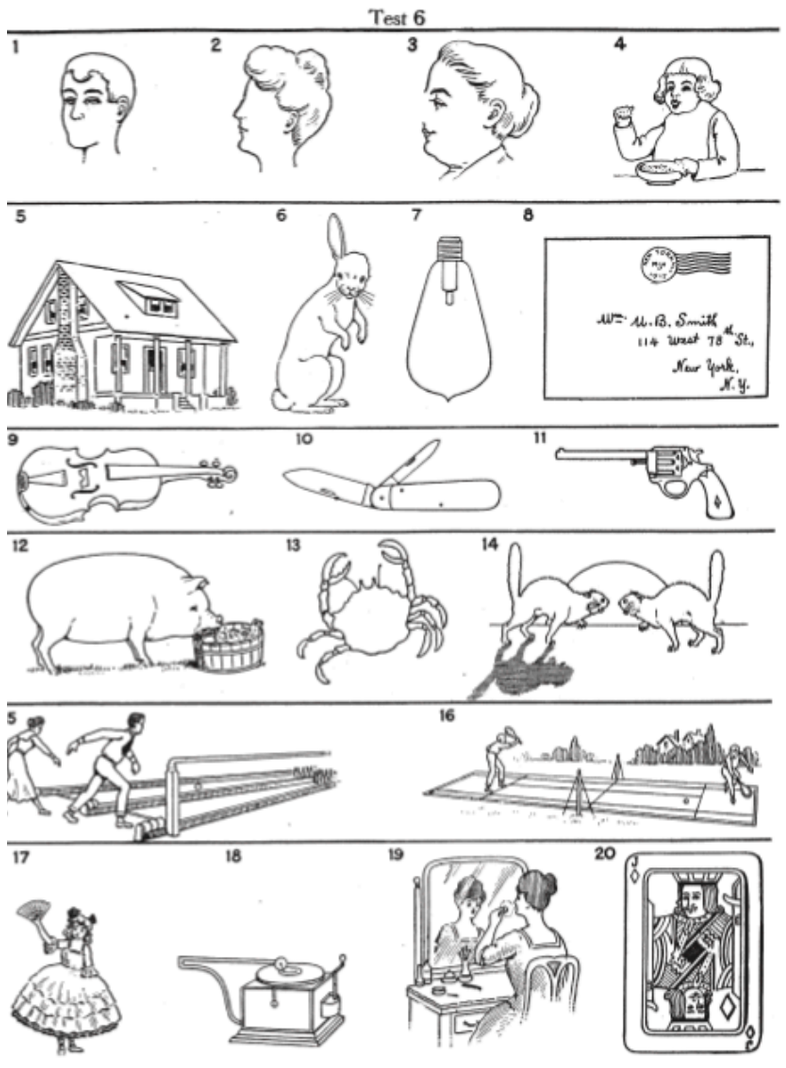
13 The cause of echoes is
- ☐ the reflection of sound waves
- ☐ the presence of electricity in the air
- ☐ the presence of moisture in the air

14 We see no stars at noon because
- ☐ they have moved around to the other side of the earth
- ☐ they are so much fainter than the sun
- ☐ they are hidden behind the sky

15 Some men lose their breath on high mountains because
- ☐ the wind blows their breath away
- ☐ the air is too rare
- ☐ it is always cold there

16 Why do some men who could afford to own a house live in a rented one? Because
- ☐ they don't have to pay taxes
- ☐ they don't have to buy a rented house
- ☐ they can make more by investing the money the house would cost

Army Beta



The **Army Alpha** consisted of eight verbally loaded tests for average and high-functioning recruits.

The **Army Beta** was a nonverbal group test designed for use with illiterates and recruits whose first language was not English. The examiners used largely pictorial and gestural methods for explaining the tests to the prospective Army recruits.

Test scores yielded grade ratings from A (very superior) to E (very inferior).

Over 1.7 million people were tested but it is unclear how/whether results were used in personnel allocation decisions.

Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)
- **1905:** Binet-Simon scale, 30 ordered items
- **1914:** William Stern publishes *The psychological methods of testing intelligence*
- **1914/1916:** Stanford-Binet Scale through translation of Binet Scale by Terman, introduction of IQ.
- **1917:** Woodworth Personal Data Sheet (WPDS)
- **1917:** Army Alpha & Beta, group tests for military recruits, Committee led by Yerkes (advisors, Terman, Cattell, Goddard)

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

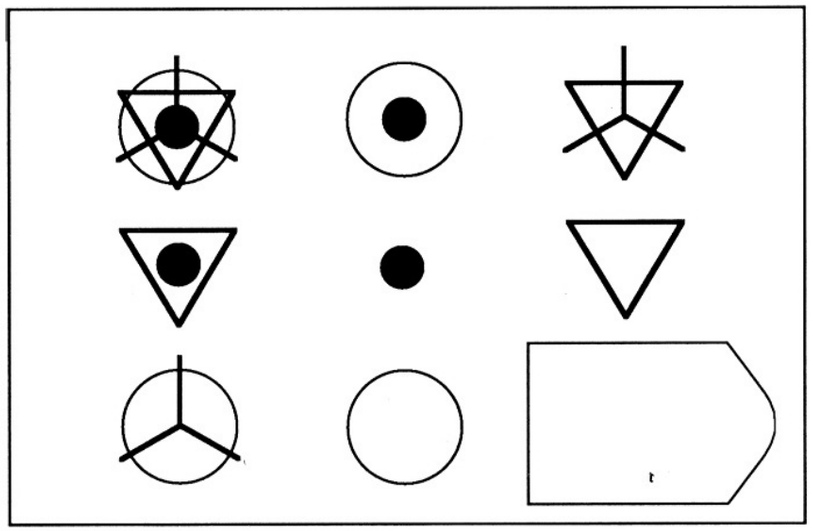
- **1921:** Psychological Corporation founded by J.M. Cattell, R.B. Woodworth, & E.L. Thorndike
- **1926:** Scholastic Aptitude Test (SAT) to predict college success (but only normed later in 1941)
- **1933:** Publication of *Vectors of Mind* by L.L. Thurstone suggestion reformulation of Spearman's approach
- **1935:** Foundation of the Psychometric Society
- **1938:** Ravens' progressive matrices

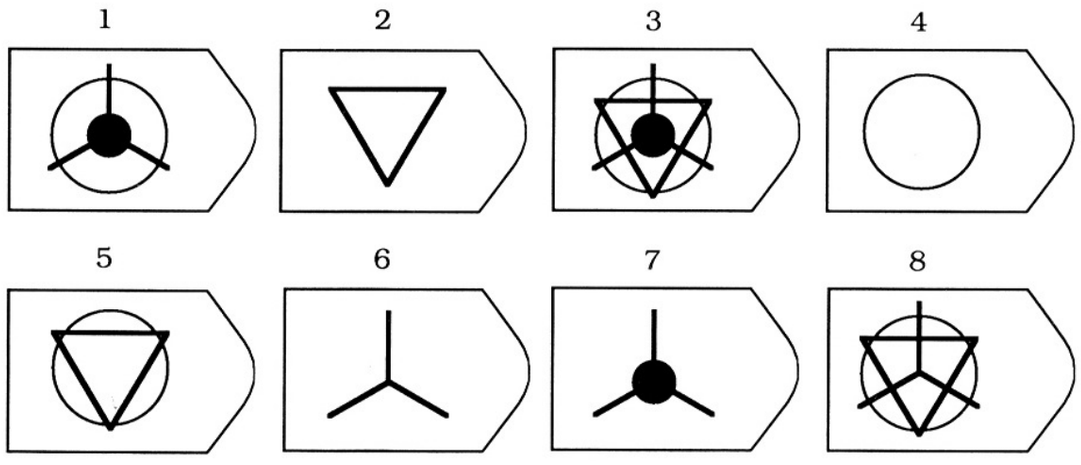
PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

Progressive matrices





Raven's Progressive Matrices:

- Developed by John Raven (1902-1970), British psychologist and student of Charles Spearman
- Non-verbal test typically used to measure general human intelligence and abstract reasoning and is regarded as a non-verbal estimate of fluid intelligence.
- One of the most common tests administered to both groups and individuals ranging from 5-year-olds to the elderly
- Made of 60 multiple choice questions, listed in order of difficulty. In each test item, the subject is asked to identify the missing element that completes a pattern. Many patterns are presented in the form of a 6x6, 4x4, 3x3, or 2x2 matrix, giving the test its name.

https://en.wikipedia.org/wiki/Raven%27s_Progressive_Matrices

Raven, J., Raven, J.C., & Court, J.H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)
- **1905:** Binet-Simon scale, 30 ordered items
- **1914:** William Stern publishes *The psychological methods of testing intelligence*
- **1914/1916:** Stanford-Binet Scale through translation of Binet Scale by Terman, introduction of IQ.
- **1917:** Woodworth Personal Data Sheet (WPDS)
- **1917:** Army Alpha & Beta, group tests for military recruits, Committee led by Yerkes (advisors, Terman, Cattell, Goddard)

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

- **1921:** Psychological Corporation founded by J.M. Cattell, R.B. Woodworth, & E.L. Thorndike
- **1926:** Scholastic Aptitude Test (SAT) to predict college success (but only normed later in 1941)
- **1933:** Publication of *Vectors of Mind* by L.L. Thurstone suggestion reformulation of Spearman's approach
- **1935:** Foundation of the Psychometric Society
- **1938:** Ravens' progressive matrices
- **1938:** Wechsler-Bellevue Intelligence Test (published by Psych Corp.)

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

History of subtests

Why is the development of subtests at different points in time of relevance, particularly in the context of psychological testing?

Table 1
Subtests on the current versions of the Stanford-Binet and Wechsler tests.

Subtest names[ab]	SB5	WPPSI-IV	WISC-V	WAIS-IV
Arithmetic/verbal and nonverbal quantitative reasoning	X		X	X
Block design		X	X	X
Block span	X			
Cancellation		X	X	X
Coding/animal coding		X	X	X
Comprehension		X	X	X
Delayed response	X			
Digit span/picture span/letter-number sequencing			X	X
Early reasoning	X			
Figure weights			X	X
Form board and form patterns/visual puzzles/object assembly	X	X	X	X
Information		X	X	X
Last word	X			
Matrices/object series/matrix reasoning	X	X	X	X
Memory for sentences	X			
Picture absurdities	X			
Picture completion				X
Picture concepts		X	X	
Picture memory/picture naming		X		
Position and direction	X			
Procedural knowledge	X			
Similarities		X	X	X
Symbol search/bug search		X	X	X
Verbal analogies	X			
Verbal absurdities	X			
Vocabulary/receptive vocabulary	X	X	X	X
Zoo locations		X		

[a] Some subtests are listed as having multiple names because subtests from different instruments may have different names but formats that are so similar that we considered the subtests to be the same.
[b] Subtests are listed alphabetically.

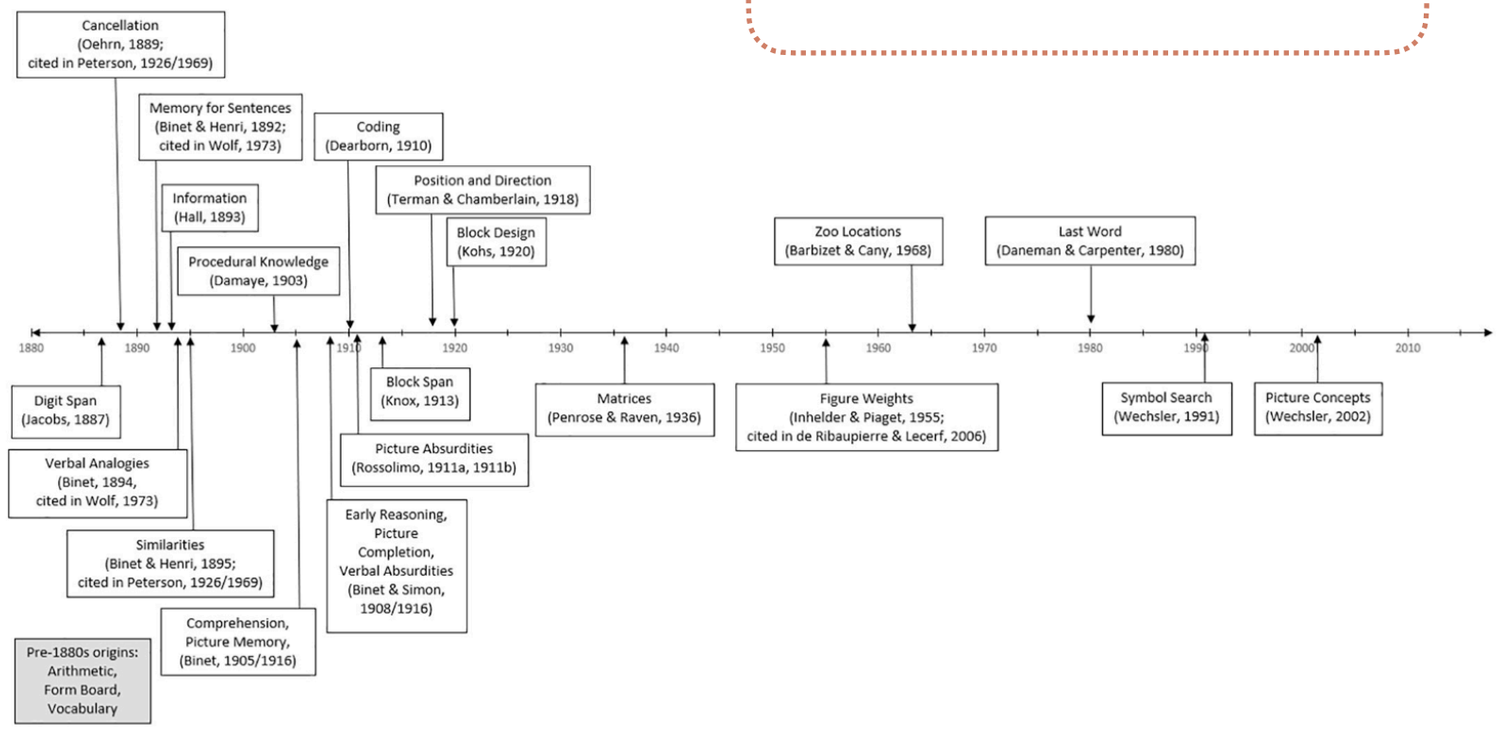


Fig. 2. Timeline of our proposed candidates for the first known publication of Stanford-Binet 5 and modern Wechsler subtests.

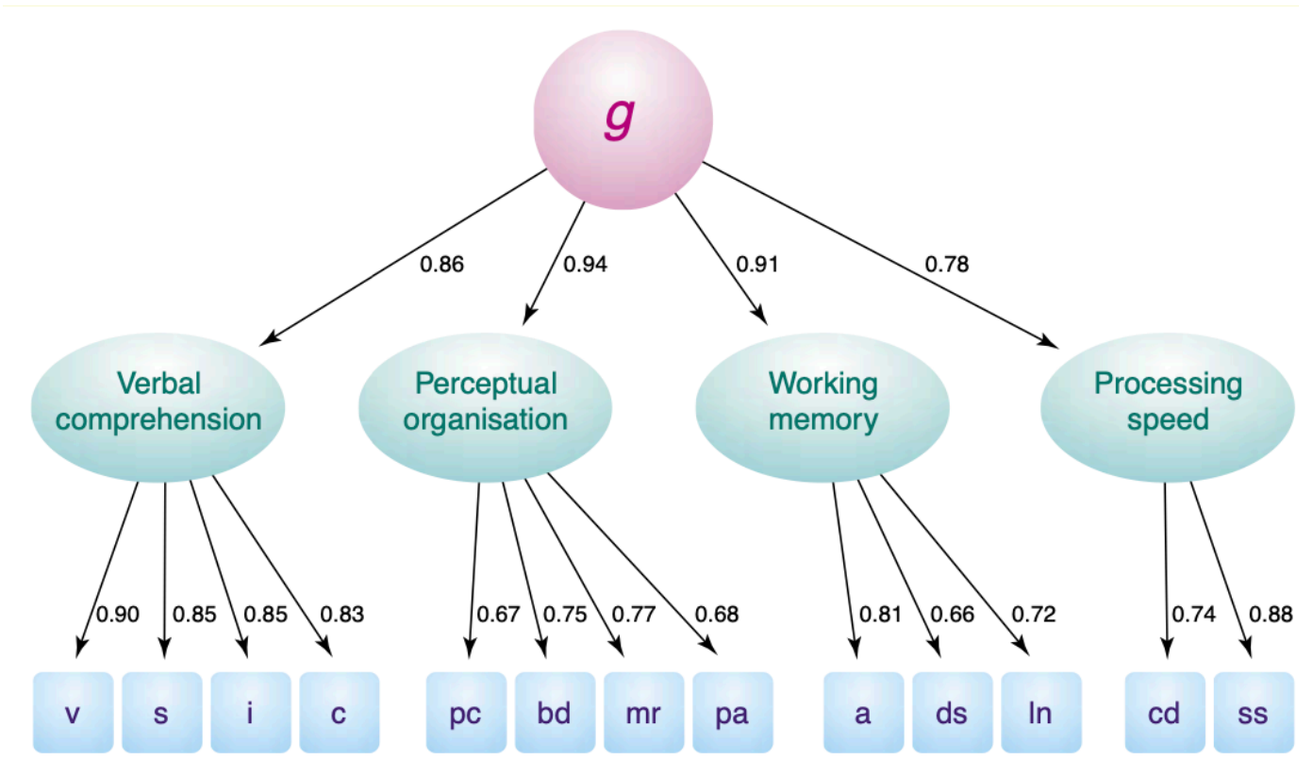
"We found that the majority of these subtest formats were first created in 1908 or earlier and that only three have been created since 1980. We discuss the implications of these findings, which are that (1) many subtests have lengthy research histories that support their use in measuring intelligence; (2) many subtests have formats that predate modern theories of test creation, cognitive psychology, and intelligence; and (3) the history of many subtests is more complex than psychologists probably realize."

Gibbons, A., & Warne, R. T. (2019). First publication of subtests in the Stanford-Binet 5, WAIS-IV, WISC-V, and WPPSI-IV. *Intelligence*, *75*, 9-18. Link to paper

Wechsler Intelligence Scales

- Developed by David Wechsler (1896-1981), American psychologist (worked with Cattell and Spearman)
- Several versions, including Wechsler Adult Intelligence Scale (WAIS), Wechsler Intelligence Scale for Children (WISC), and Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV)
- WAISC is a revision of the Wechsler-Bellevue Intelligence Scale (originally released in 1939), contains multiple tasks that are supposed to capture different latent constructs

https://en.wikipedia.org/wiki/Wechsler_Adult_Intelligence_Scale

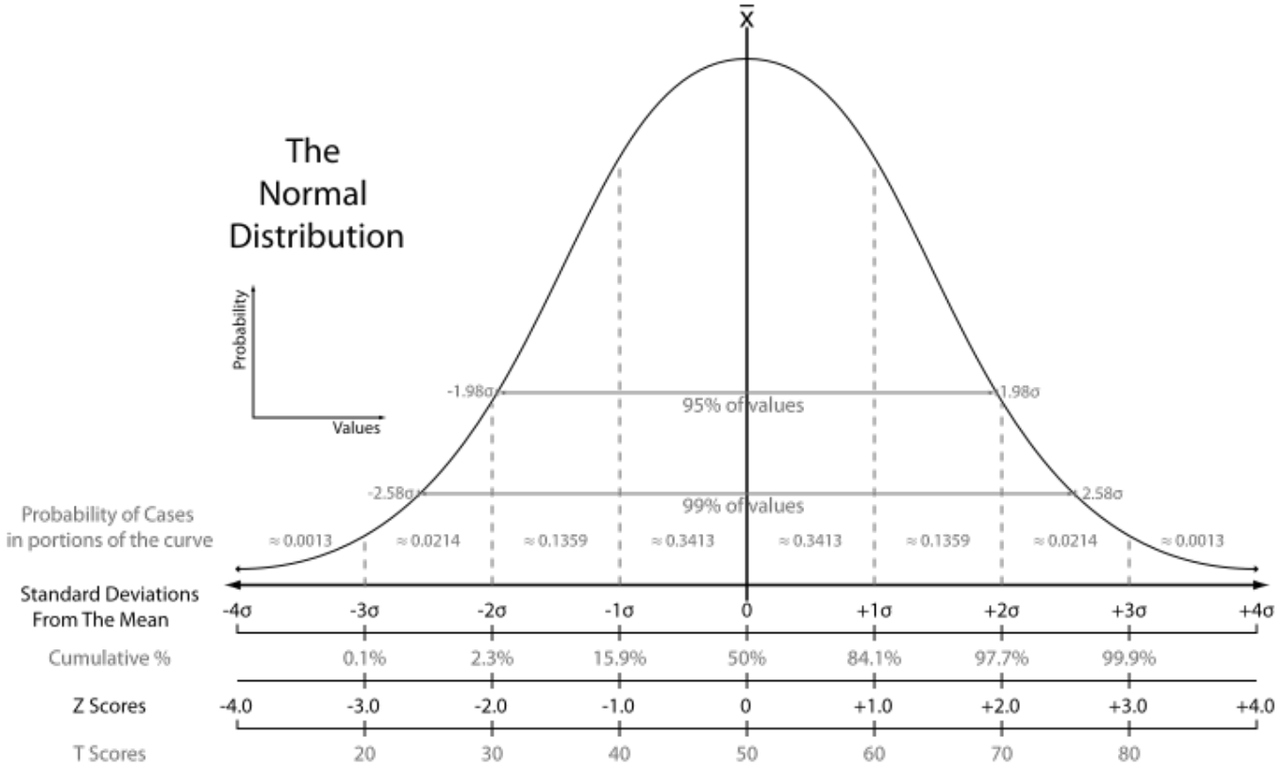


Scale	Abbreviation	Factor
Vocabulary	v	Verbal Comprehension
Similarities	s	
Information	i	
Comprehension	c	
Picture completion	pc	Perceptual Organization
Bloc design	bd	
Matrix reasoning	mr	
Picture arrangement	pa	
Arithmetic	a	Working Memory
Digit span	ds	
Letter-number sequencing	ln	
Digit-symbol coding	cd	Processing Speed
Symbol search	ss	

Deary, I. J. (2001). Human intelligence differences: a recent history. *Trends in cognitive sciences*, *5*(3), 127-130. Link to paper

Wechsler Intelligence Scales

- The Wechsler scales were not revolutionary in the tests used (a lot of them were copied from other extant scales, as we have seen already)! The main innovation was the use of the point scale (rather than a chronological age scale as in Stanford-Binet scales)
- This allowed assigning points to each item and thus items to be grouped according to content The scale thus also allowed testers to obtain multiple scores for each individual.
- The point scale transformed how intelligence was measured and interpreted, providing a multidimensional perspective on cognitive abilities.



What does this new scoring method mean for the interpretation of what IQ is or measures?

-> IQ is fundamentally relative, it quantifies the deviation of an individual's cognitive abilities from the statistical mean of a standardized norm-building sample.

$$IQ = 100 + 15 \times \frac{\text{Raw score - Mean}}{\text{Standard Deviation}}$$

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)
- **1905:** Binet-Simon scale, 30 ordered items
- **1914:** William Stern publishes *The psychological methods of testing intelligence*
- **1914/1916:** Stanford-Binet Scale through translation of Binet Scale by Terman, introduction of IQ.
- **1917:** Woodworth Personal Data Sheet (WPDS)
- **1917:** Army Alpha & Beta, group tests for military recruits, Committee led by Yerkes (advisors, Terman, Cattell, Goddard)

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

- **1921:** Psychological Corporation founded by J.M. Cattell, R.B. Woodworth, & E.L. Thorndike
- **1926:** Scholastic Aptitude Test (SAT) to predict college success (but only normed later in 1941)
- **1933:** Publication of *Vectors of Mind* by L.L. Thurstone suggestion reformulation of Spearman's approach
- **1935:** Foundation of the Psychometric Society
- **1938:** Ravens' progressive matrices
- **1938:** Wechsler-Bellevue Intelligence Test (published by Psych Corp.)
- **1943:** Minnesota Multiphasic Personality Inventory (MMPI)

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)
- **1905:** Binet-Simon scale, 30 ordered items
- **1914:** William Stern publishes *The psychological methods of testing intelligence*
- **1914/1916:** Stanford-Binet Scale through translation of Binet Scale by Terman, introduction of IQ.
- **1917:** Woodworth Personal Data Sheet (WPDS)
- **1917:** Army Alpha & Beta, group tests for military recruits, Committee led by Yerkes (advisors, Terman, Cattell, Goddard)

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

- **1921:** Psychological Corporation founded by J.M. Cattell, R.B. Woodworth, & E.L. Thorndike
- **1926:** Scholastic Aptitude Test (SAT) to predict college success (but only normed later in 1941)
- **1933:** Publication of *Vectors of Mind* by L.L. Thurstone suggestion reformulation of Spearman's approach
- **1935:** Foundation of the Psychometric Society
- **1938:** Ravens' progressive matrices
- **1938:** Wechsler-Bellevue Intelligence Test (published by Psych Corp.)
- **1943:** Minnesota Multiphasic Personality Inventory (MMPI)

PHASE IV (1950s - present)

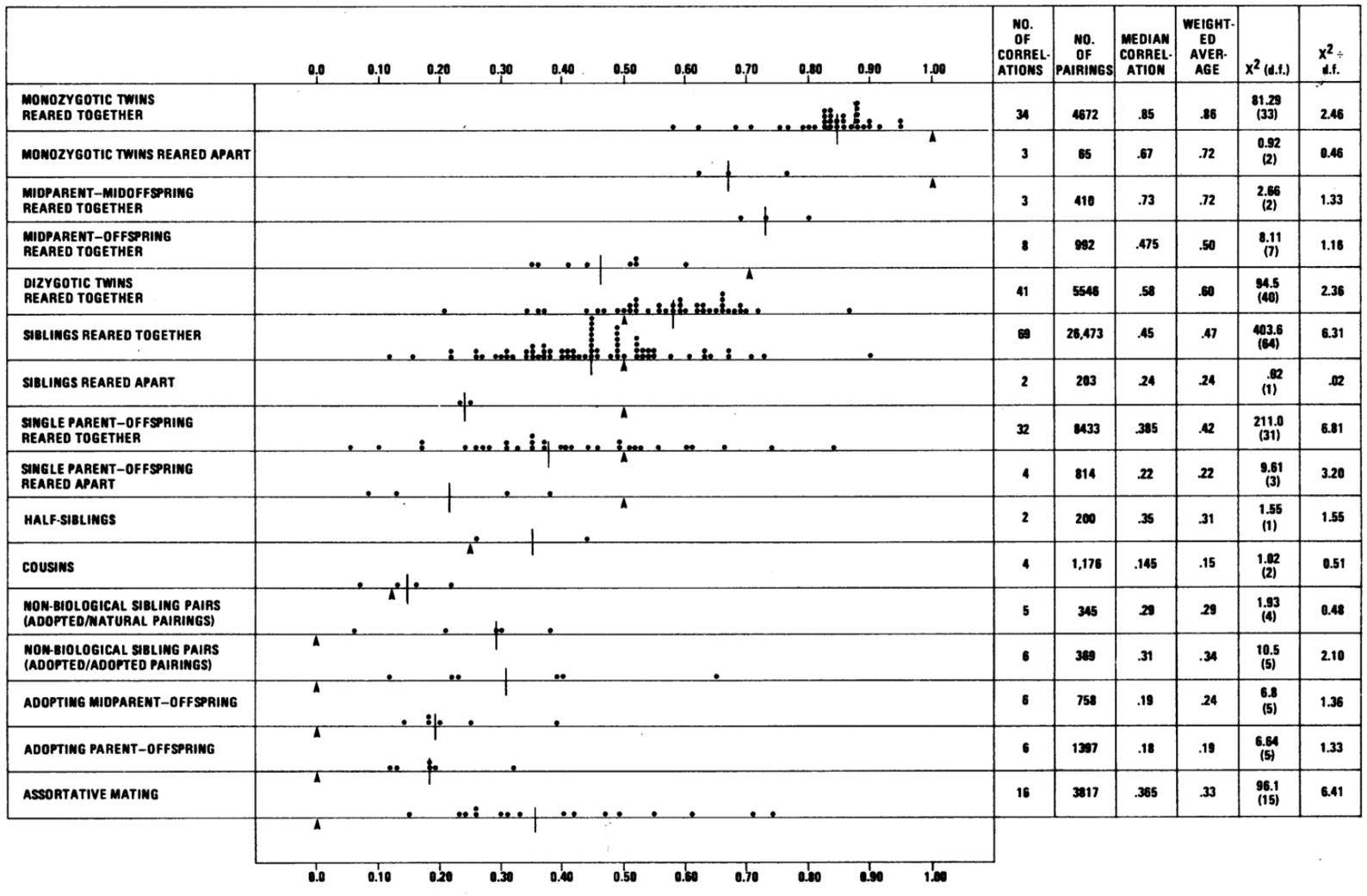
Methodological and empirical advances, behavioral genetics, and consensus

- **1950:** APA publishes *Ethical standards for the distribution of psychological tests and diagnostic aids*
- **1951:** Lee Cronbach introduces the alpha coefficient to measure consistency
- **1954:** *Technical recommendations for psychological tests and diagnostic aids* by APA (chaired by L. Cronbach)
- **1955:** Wechsler Adult Intelligence Scale (WAIS)
- **1963:** Theory of fluid and crystallized intelligence by R.B. Cattell
- **1980:** *Applications of Item Response Theory to Practical Testing Problems* by Frederic Lord
- **1981:** *Familial studies of intelligence* by Bouchard & McGue in *Science*

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

Behavioral genetics

"A summary of 111 studies identified in a survey of the world literature on familial resemblances in measured intelligence reveals a profile of average correlations consistent with a polygenic mode of inheritance. There is, however, a marked degree of heterogeneity of the correlations within familial groupings, which is not moderated by sex of familial pairing or by type of intelligence test used."



	NO. OF CORREL-ATIONS	NO. OF PAIRINGS	MEDIAN CORREL-ATION	WEIGHT-ED AVER-AGE	x^2 (d.f.)	$x^2 \div$ d.f.
MONOZYGOTIC TWINS REARED TOGETHER	34	4672	.85	.86	81.29 (33)	2.46
MONOZYGOTIC TWINS REARED APART	3	65	.67	.72	0.92 (2)	0.46
MIDPARENT–MIDOFFSPRING REARED TOGETHER	3	410	.73	.72	2.66 (2)	1.33
MIDPARENT–OFFSPRING REARED TOGETHER	8	992	.475	.50	8.11 (7)	1.16
DIZYGOTIC TWINS REARED TOGETHER	41	5546	.58	.60	94.5 (40)	2.36
SIBLINGS REARED TOGETHER	69	26,473	.45	.47	403.6 (64)	6.31
SIBLINGS REARED APART	2	203	.24	.24	.92 (1)	.02
SINGLE PARENT–OFFSPRING REARED TOGETHER	32	8433	.385	.42	211.0 (31)	6.81
SINGLE PARENT–OFFSPRING REARED APART	4	814	.22	.22	9.61 (3)	3.20
HALF-SIBLINGS	2	200	.35	.31	1.55 (1)	1.55
COUSINS	4	1,176	.145	.15	1.02 (2)	0.51
NON-BIOLOGICAL SIBLING PAIRS (ADOPTED/NATURAL PAIRINGS)	5	345	.29	.29	1.93 (4)	0.48
NON-BIOLOGICAL SIBLING PAIRS (ADOPTED/ADOPTED PAIRINGS)	6	369	.31	.34	10.5 (5)	2.10
ADOPTING MIDPARENT–OFFSPRING	6	758	.19	.24	6.8 (5)	1.36
ADOPTING PARENT–OFFSPRING	6	1397	.18	.19	6.64 (5)	1.33
ASSORTATIVE MATING	16	3817	.365	.33	96.1 (15)	6.41

0.0 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00

Fig. 1. Familial correlations for IQ. The vertical bar in each distribution indicates the median correlation; the arrow, the correlation predicted by a simple polygenic model.

Behavioral Genetics

==> the field of study that examines the role of genetic and environmental influences on animal (including human) behaviour. Behavioural geneticists study the inheritance of behavioural traits. In humans, this information is often gathered through the use of the *twin study or adoption study*.

Heritability

==> a statistic used in breeding and genetics works that estimates how much *variation in a phenotypic trait in a population is due to genetic variation* among individuals in that population.

Bouchard Jr, T. J., & McGue, M. (1981). Familial studies of intelligence: A review. *Science*, *212*(4498), 1055-1059. Link to paper

Psychological testing and psychometrics: Four phases and major events

PHASE I (1880s - 1890s)

The Origins:
Standardisation and objectivity in psychological testing

- **1883:** Francis Galton publishes *Inquiries into Human Faculty and Its Development*
- **1984:** Francis Galton administers a test battery to thousands of volunteers at the International Health Exhibition, London
- **1890:** James McKeen Cattell publishes *Mental Tests and Measurements*
- **1896:** Karl Pearson develops the product-moment correlation

PHASE II (1900s - 1910s)

The Birth of IQ:
Norms and predictive validity

- **1901:** Applications of correlation by Clark Wissler (Cattell's student) suggest that "brass measures" have little predictive validity (e.g., for academic performance)
- **1904:** Charles Spearman publishes *The proof and measurement of the association between two things (rank order!)*
- **1904:** *General Intelligence* by Charles Spearman, introduces *Correlational psychology* and *g* factor (psychometric *g*)
- **1905:** Binet-Simon scale, 30 ordered items
- **1914:** William Stern publishes *The psychological methods of testing intelligence*
- **1914/1916:** Stanford-Binet Scale through translation of Binet Scale by Terman, introduction of IQ.
- **1917:** Woodworth Personal Data Sheet (WPDS)
- **1917:** Army Alpha & Beta, group tests for military recruits, Committee led by Yerkes (advisors, Terman, Cattell, Goddard)

PHASE III (1920s - 1940s)

Modern Tests:
Point scales, verbal, non-verbal, and 'culture-free' tests

- **1921:** Psychological Corporation founded by J.M. Cattell, R.B. Woodworth, & E.L. Thorndike
- **1926:** Scholastic Aptitude Test (SAT) to predict college success (but only normed later in 1941)
- **1933:** Publication of *Vectors of Mind* by L.L. Thurstone suggestion reformulation of Spearman's approach
- **1935:** Foundation of the Psychometric Society
- **1938:** Ravens' progressive matrices
- **1938:** Wechsler-Bellevue Intelligence Test (published by Psych Corp.)
- **1943:** Minnesota Multiphasic Personality Inventory (MMPI)

PHASE IV (1950s - present)

Methodological and empirical advances, behavioral genetics, and consensus

- **1950:** APA publishes *Ethical standards for the distribution of psychological tests and diagnostic aids*
- **1951:** Lee Cronbach introduces the alpha coefficient to measure consistency
- **1954:** *Technical recommendations for psychological tests and diagnostic aids* by APA (chaired by L. Cronbach)
- **1955:** Wechsler Adult Intelligence Scale (WAIS)
- **1963:** Theory of fluid and crystallized intelligence by R.B. Cattell
- **1980:** *Applications of Item Response Theory to Practical Testing Problems* by Frederic Lord
- **1981:** *Familial studies of intelligence* by Bouchard & McGue in *Science*
- **1997:** 52 intelligence researchers sign editorial on a consensus about intellgence and mental testing

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. Link to ebook
Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. Link to chapter

25-point consensus on intelligence and testing

"1. Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings—"catching on," "making sense" of things, or "figuring out" what to do."

"2. Intelligence, so defined, can be measured, and intelligence tests measure it well. They are among the most accurate (in technical terms, reliable and valid) of all psychological tests and assessments. They do not measure creativity, character, personality, or other important differences among individuals, nor are they intended to."

…

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence, 24*(1), 13-23. Link to paper

More consensus on intelligence and testing

Table 1
Predictive Validity for Overall Job Performance of General Mental Ability (GMA) Scores Combined With a Second Predictor Using (Standardized) Multiple Regression

Personnel measures	Validity (r)	Multiple R	Gain in validity from adding supplement	% increase in validity	Standardized regression weights	
					GMA	Supplement
GMA tests[a]	.51					
Work sample tests[b]	.54	.63	.12	24%	.36	.41
Integrity tests[c]	.41	.65	.14	27%	.51	.41
Conscientiousness tests[d]	.31	.60	.09	18%	.51	.31
Employment interviews (structured)[e]	.51	.63	.12	24%	.39	.39
Employment interviews (unstructured)[f]	.38	.55	.04	8%	.43	.22
Job knowledge tests[g]	.48	.58	.07	14%	.36	.31
Job tryout procedure[h]	.44	.58	.07	14%	.40	.20
Peer ratings[i]	.49	.58	.07	14%	.35	.31
T & E behavioral consistency method[j]	.45	.58	.07	14%	.39	.31
Reference checks[k]	.26	.57	.06	12%	.51	.26
Job experience (years)[l]	.18	.54	.03	6%	.51	.18
Biographical data measures[m]	.35	.52	.01	2%	.45	.13
Assessment centers[n]	.37	.53	.02	4%	.43	.15
T & E point method[o]	.11	.52	.01	2%	.39	.29
Years of education[p]	.10	.52	.01	2%	.51	.10
Interests[q]	.10	.52	.01	2%	.51	.10
Graphology[r]	.02	.51	.00	0%	.51	.02
Age[s]	−.01	.51	.00	0%	.51	−.01

➡ Mental tests (i.e., general mental ability) have predictive validity in the real world, such as in job performance!

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological bulletin*, *124*(2), 262. Link to paper

Summary

- **Psychological testing:** initially fuelled by eugenics and need for objective measurement; dual focus on theory of psychological faculties and applications to real-world problems (selection)

- **Measurement:** researchers first emphasised standardisation of measurement conditions and later standardisation of scoring; introduction of points and homogenisation across batteries (i.e., IQ) facilitated determining (age appropriate) norms, and distinction of components (e.g., verbal vs. non-verbal)

- **Methods:** development of statistical methods to quantify association between variables (e.g., Pearson correlation) and perform dimensionality reduction (e.g., factor analysis/principal component analysis)

- **Conceptual issues:** theoretical consensus by the end of the 20th century, some agreement about biological (behavioural genetics) and environmental bases (still debate about the nature of *g*, neural basis, etc.).

- **Main applications:** personnel selection/job performance, academic achievement

Key reading

Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3-55). Guilford Press. https://johndwasserman.com/index_htm_files/Wasserman%202012%20A%20History%20of%20Intelligence%20Assessment%20The%20Unfinished%20Tapestry.pdf

Additional reading (optional)

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Pearson Education India. **Link to ebook**

Additional materials

See this webpage for a demonstration / graphical interpretation of a similar statistical procedure - principal component analysis: http://setosa.io/ev/principal-component-analysis/