
PSY1102

Introduction to Applied Psychology

Class 8

Intelligence (continued)

Victor Emerson
vemerson@uottawa.ca

Agenda for today

1. Assessing intelligence
 - a. The origins of intelligence testing
 - b. Modern tests of mental abilities
 - c. Principles of test construction
2. The dynamics of intelligence
 - a. Stability or change?
 - b. Extremes of intelligence

1. Assessing intelligence

- a. The origins of intelligence testing
- b. Modern tests of mental abilities
- c. Principles of test construction

1. Assessing intelligence

How should we measure intelligence?

What method(s) do you suggest?

1a. The origins of intelligence testing

- The first recorded mention of interest in differential intelligence is from Plato's *The Republic*, written more than 2000 years ago.
- The first efforts to exploit these differences came in the 1800s.
 - While Charles Darwin was doing his research, his cousin Francis Galton was concerned with measuring human traits.
 - Galton tried to extend Darwin's work on natural selection, and to apply it to human intelligence.
 - In his view, if we could measure human intelligence we could get “intelligent” humans to breed with each other and so “improve” the human race.

1a. The origins of intelligence testing: Galton

- The first mass effort at measurement occurred at the London Exposition of 1884, where more than 10,000 people were tested for:
 - Reaction time
 - Sensory acuity
 - Muscular power
 - Body proportions
- These measures failed to differentiate between “brighter” and “less bright” people. Moreover, the measures were not correlated with each other.
- However, this attempt introduced statistical analysis to intelligence testing.

Are these the variables
you mentioned a few
minutes ago?

1a. The origins of intelligence testing: Binet

- Early in the 20th century, France required that all children attend school.
- The question arose – how can we tell which students have special needs, and may require special classes?
- Two choices existed:
 - Rely on the subjective evaluation by teachers, or
 - Develop an objective test.
- Alfred Binet, who addressed this problem, assumed that all children followed the same developmental course, but at different rates. In his view,
 - A “bright” child performed at the level of an older child, and
 - A “dull” child performed at the level of a younger child.

1a. The origins of intelligence testing: Binet (cont'd.)

- This approach has a couple of important implications:
 - It implies that one can define what is normal at each chronological age, and
 - It implies that we can measure the mental age of a child and compare it to this norm.
- Binet and his colleague Simon developed a test designed to predict which students would do well at schoolwork and which would do more poorly.
- His intention was not to measure “native” intelligence, but to predict performance in school.
- Indeed, Binet believed that the environment shaped intelligence, and that mental exercises would improve performance.

1a. The origins of intelligence testing: Terman

- At Stanford University, Lewis Terman tried to use Binet's tests as a measure of “innate” intelligence.
 - He modified the test and named it the Stanford-Binet test.
- In Germany, Stern developed the intelligence quotient (IQ), which divides the person's “mental age” (as measured by the test) by the person's chronological age, multiplying the result by 100:

$$IQ = \frac{\text{mental age}}{\text{chronological age}} \times 100$$

- Thus, the average IQ was defined as $IQ=100$.
 - A 10-year-old child with a mental age of 8 had an $IQ=80$.
 - A 10-year-old child with a mental age of 12 had an $IQ=120$.

1a. Origins of intelligence testing: Terman (cont'd.)

- Although IQ seemed to work acceptably for children, it did not work as well for adults.
- Accordingly, over time the calculation was changed to represent the person's score compared to the average performance of others of the same age.
- For the distribution of IQ scores, the mean is defined as 100 and the standard deviation is 15.
 - By this definition, 68.26% of all people fall between 85 and 115.

1a. Origins of intelligence testing: Eugenics

- Intelligence testing has a “dark side”.
- Instead of using IQ scores to predict academic or work performance – where the predictive power was useful – Terman aligned with the eugenics movement started in Galton’s time.
- Eugenics proposed that, as with selective breeding to create and maintain different breeds of dogs, it should be possible to breed out “feeble-mindedness”, ridding society of much of crime, poverty, and inefficiency.
 - In the 20th century, the Nazis favoured an accelerated eugenics programme to rid the world of Jews, Catholics, Roma, and other “undesirables” who did not meet the criteria of the Aryan ideal.

1a. Origins of intelligence testing: Eugenics (cont'd.)

- When IQ tests were administered to immigrants to the US in the early 20th century, it was discovered that immigrants from southern and eastern Europe had a high proportion of “unsuitable” people compared to immigrants from northern and western Europe.
 - In other words, the people who scored best on the test were those immigrants from the same geographic area that produced the people who developed the test.
 - By contrast, immigrants from southern and eastern Europe did not perform as well, and so were seen to be “unsuitable”.
- As a result, immigration quotas were imposed to limit immigration from southern and eastern Europe.

1b. Modern tests of mental abilities

- We've all experienced tests of our capabilities:
 - Achievement tests are designed to measure what we have learned – driver's tests, final exams, etc.;
 - Aptitude tests are designed to predict our ability to learn a new skill – will you be a good mechanic? a good teacher? Are you suited for university? etc.
- There is a high correlation ($r=0.82$) between scores on intelligence tests and scores on aptitude tests.
- However, the relation between aptitude and achievement tests is not as clear.
 - For example, your aptitude test score is influenced by your vocabulary, which you have learned.
 - Also, your aptitude for academic work and test-taking influences your score on achievement tests.

1b. Modern tests of mental abilities: WAIS

- The most widely used intelligence test for adults is the WAIS (Wechsler Adult Intelligence Test); for children, it is the WISC (Wechsler Intelligence Scale for Children), and there is a third version for pre-schoolers.
- Page 419 of the textbook shows examples of test items.
- The WAIS has 11 subtests, including both performance and verbal tests.
 - As a result, the WAIS yields an overall intelligence score as well as scores for the subtests (verbal comprehension, processing speed, working memory, perceptual organisation).
 - Thus, differential scores on subtests should provide guidance to teachers about where more work is required.

1c. Principles of test construction

- Of course, test results are only useful if we can trust them.
- Tests are not “just made up” by people writing out a list of questions. Rather, tests must be reliable, valid, and standardised.

1c. Principles of test construction: reliability

- A reliable test is one that produces the same result over repeated testing.
- For example, if I administer an intelligence test to you today and another one in 3 months, the results of these two tests should be very similar, although not necessarily identical.
- Reliability can be verified by:
 - Administering the same test to the same people (test-retest reliability); or by
 - Splitting the test in half (e.g., odd vs. even questions) and comparing performance on the two halves (split-half reliability)
- The Stanford-Binet, the WAIS, and the WISC all have reliabilities around 0.9, which is very high.

1c. Principles of test construction: validity

- Validity means that the test represents reality.
- One can easily have reliability without validity. For example, I could set my bathroom scale so that it reads 10 kg light. The scale is reliable – that is, it gives the same result each time – but it is not valid.
- Tests with content validity relate to the behaviour under study (criterion). A driver's test has content validity.
- Intelligence tests must have predictive validity – that is, they must predict future performance.
 - For aptitude tests, the predictive power is good early in school years, but is weaker in later years.

1c. Principles of test construction: validity (cont'd.)

- It should not be surprising that predictive intelligence tests lose their predictive power as they are applied to people who have gone farther in the educational process.
 - In grade school, there is a wide variety of skills and abilities;
 - In university, people without an aptitude for school have moved into other occupations, and so only the “academics” are being tested.
 - By graduate school, the sample is even more restricted.
 - Restricting the range of variability necessarily reduces the correlation between the test score and prediction of future achievement.
- In a similar way, an aptitude test for hockey skills would have greater predictive power for 6-year-olds than for NHL players.

1c. Principles of test construction: standardisation

- Because of its prevalence in nature, the “bell curve” (more formally known as the normal distribution or the Gaussian distribution) is commonly used to describe the distribution of scores on variables such as:
 - Heights
 - Weights
 - Test scores
- The normal distribution is appropriate for a population. For a sample, which typically has many fewer observations, we expect scores on a standardised test to approximate a normal distribution.

1c. Test construction: standardisation (continued)

- For intelligence tests, the mean is set to 100, and the standard deviation (a measure of the spread of scores) is set to 15.
- Because of the nature of the normal distribution, we know that 68.26% of the population will fall within one standard deviation of the mean. (See figure on p. 420.)
- Over the decades, the mean score on the WAIS has been improving; this is called the Flynn effect.
- Knowing this, it has been possible to restandardise the WAIS to compensate.
- The causes of the Flynn effect are unknown, but could be attributable to improved diet or education, for example – or even “hybrid vigor”, the mixing of different peoples.

2. The dynamics of intelligence

- a. Stability or change?
- b. Extremes of intelligence

2a. Stability or change?

“How stable are intelligence scores over the life span?” (p. 422)

- Why is this issue important? Imagine what your reaction would be if it was the case that your score on an aptitude test you took today had no predictive value for a year from now, or a month from now, or next week. What would be the point of the test?
- At this point, we have no good way of predicting later performance based on a study of infant characteristics.
- Scores of tests administered around age 4 begin to have some predictive value, and scores begin to stabilise around age 7.

2a. Stability or change? (continued)

- On tests administered at separate times, verbal scores correlate well over the years ($r=0.86$).
- Math scores also correlate well over the years ($r=0.86$).
- However, verbal and math scores do not correlate well with each other, suggesting that these are indeed different capabilities.
- Finally, the ultimate in longitudinal studies arose partly by accident in Scotland, where follow-up testing was done nearly 70 years later on people originally tested at age 11 in 1932.
 - People who scored highly at 11 were more likely to live independently and be free of Alzheimer's disease at 80.
 - Also, more high-scorers were alive at age 80 than low-scorers.

2b. Extremes of intelligence

- Recall that intelligence scores are distributed in a bell curve. This basis enables us to provide a quantitative definition of both extremes:
 - The low extreme, and
 - The high extreme
- We'll consider these separately.

2b. Extremes of intelligence: the low extreme

- By definition, the low extreme of the distribution of intelligence scores is 70 or below (2 standard deviations below the mean of 100).
- Intellectual disability (formerly called mental retardation) requires a low test score and difficulty with the normal demands of everyday life.
 - About 1% of the population meets both criteria;
 - Males outnumber females by 50%.
- Down syndrome (also called trisomy 21) is caused by having a third copy of the 21st chromosome. Disability related to Down syndrome can vary from mild to severe.

2b. The low extreme of intelligence (cont'd.)

- Historically, persons with an intellectual disability were cared for at home, where they might work in an agrarian society.
- Later, residential schools were established for slow learners. Unfortunately, these developed into places where the residents were dropped off and forgotten.
- More recently, the pendulum has swing back to normalisation, with integration into the community and schools.
- After graduation, many move into group homes.

2b. Extremes of intelligence: the low extreme

Level	Approximate intelligence scores	Adaptation to demands of life
Mild	50-70	May learn academic skills to grade 6 level. Adults may, with assistance, achieve self-supporting social and vocational skills.
Moderate	35-50	May progress to grade 2 level academically. Adults may contribute to their own support by labouring in sheltered workshops.
Severe	20-35	May learn to talk and to perform simple work tasks under close supervision but are generally unable to profit from vocational training.
Profound	Below 20	Require constant aid and supervision.

2b. Extremes of intelligence: the high extreme

- Are people with high intelligence test scores weird? The evidence says “no”.
 - Terman studied more than 1500 school kids with scores of at least 135 and found them to be healthy, well adjusted, and successful at academics.
- More recent work shows that high-scoring people were more likely to have patents and Ph.D.s than lower-scoring people.
- Moreover, unlike 13-year-olds scoring well on the math test, there was a tendency for those scoring well on the verbal test to do well in the humanities later in life.
- A predictor of good performance: having parents who read to children, or making such opportunities available in school.

Summary: Class 8

1. Assessing intelligence
 - a. The origins of intelligence testing
 - b. Modern tests of mental abilities
 - c. Principles of test construction
2. The dynamics of intelligence
 - a. Stability or change?
 - b. Extremes of intelligence