

Principles of Test Construction

Principles of Test Construction

For a psychological test to be acceptable it must fulfill the following three criteria:

1. Standardization
2. Reliability
3. Validity

Standardization

Standardization:

Standardizing a test involves administering the test to a representative sample of future test takers in order to establish a basis for meaningful comparison.

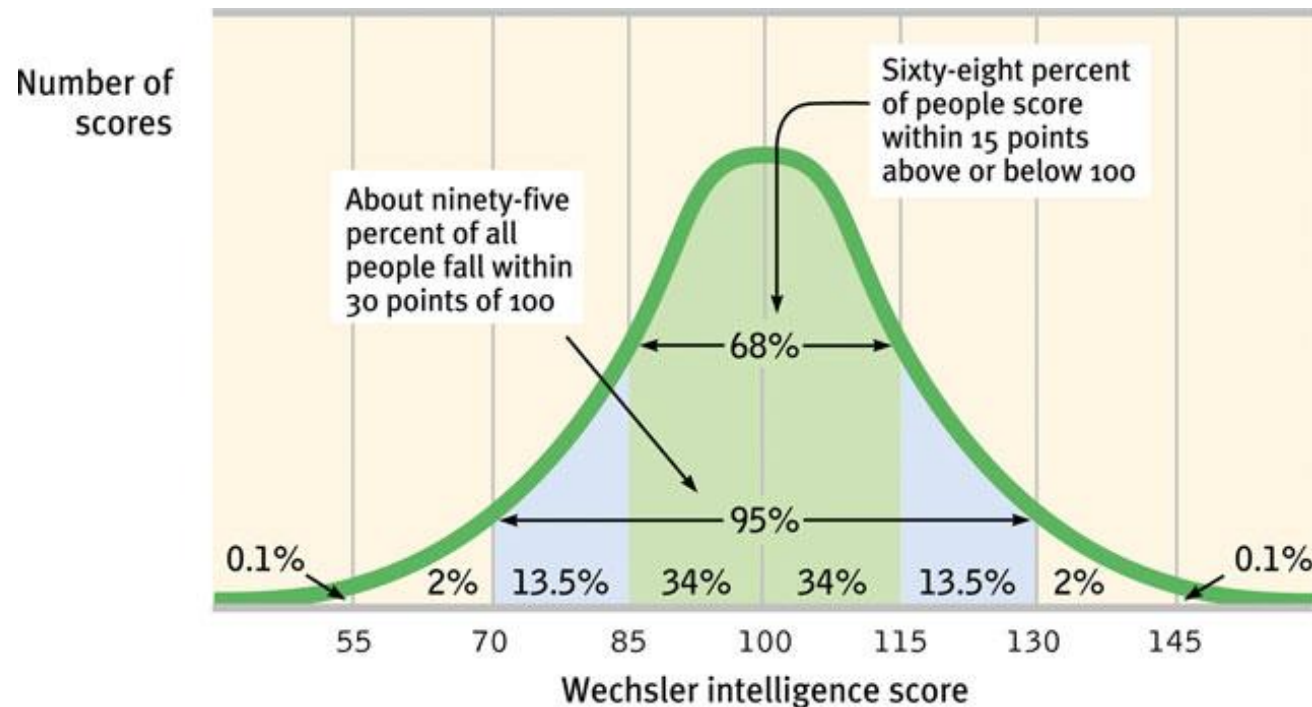
- AP Exams, Intelligence Tests, and SAT (your scores are ranked in average of many test takers)

Norms:

Average score for a designated group of people

Normal Curve

Standardized tests establish a normal distribution of scores on a tested population in a bell-shaped pattern called the **normal curve**.

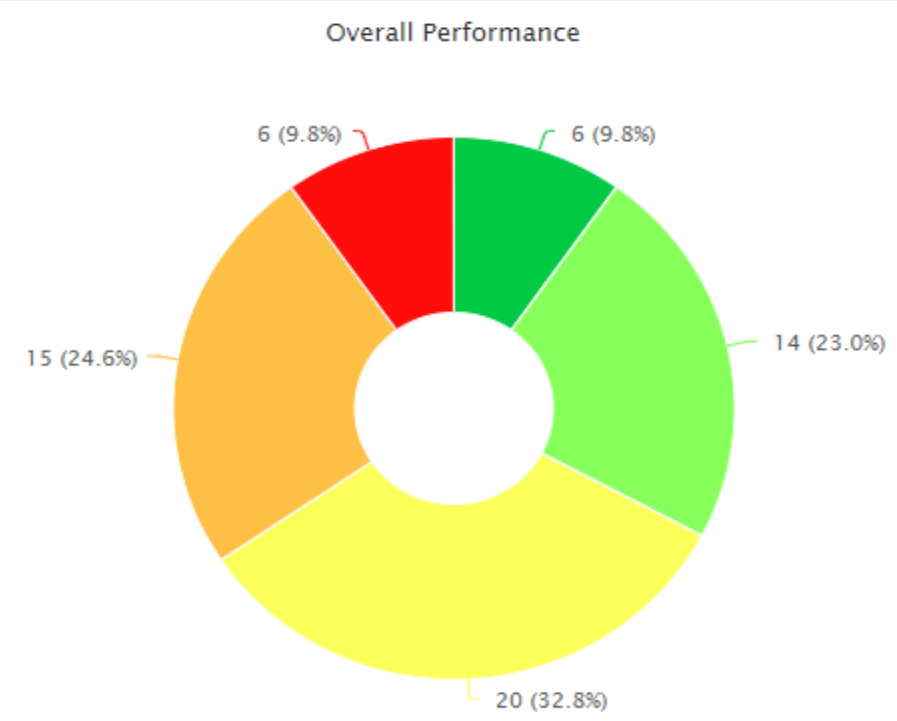


Abnormal Psychology Exam Version II

[\(Add description\)](#)

Created by (1515) Lee, Isaac
Scope: [\(Choose a Scope\)](#)
Subject Area: [\(Choose a Subject Area\)](#)
Questions: 49
Standards: [\(Find and Link Standards\)](#)
Question Groups: [Disorders, Treatments, Depression & Bipolar, Anxiety & ...](#)

[Portal](#) [Data Director](#) [Custom Reports](#) [Curriculum Associate](#) [Edusoft](#) [ALS](#)
[Fluence](#) [Shared](#) [Online Testing](#) [Itembank](#) [Versions](#)



Summary

Avg. % Correct	Students	% Not Mastered	% Mastered
74.5%	61	67.2%	32.8%

Reliability

A test is *reliable* when it yields consistent results. To establish reliability researchers establish different procedures:

- Test scores should be same or similar each time the same person takes it.
 - 1st attempt SAT – 1800 , then 2nd attempt 1000 is this reliable?
1. **Test-Retest Reliability:** Using the same test on two occasions to measure consistency.
 2. **Split-half Reliability:** Dividing the test into two equal halves and assessing how consistent the scores are.
 3. **Reliability using different tests:** Using different forms of the test to measure consistency between them.

Example of Test Re-test Method

- when determining the effect of coloring on stress, a researcher may choose to evaluate the subjects' blood pressure before, during, and after coloring a set number of pages. This gives researchers a base line from which to draw later conclusions.

Validity

Reliability of a test does not ensure validity.

Validity of a test refers to the degree in which the test is supposed to measure or predict.

1. **Content Validity:** Refers to the extent which a test measures your definition of the construct or behavior of interest?
 - Does a physical test measure your knowledge of psychology?
 - Does the AP Psychology exam measure your knowledge of psychology?
 - Does a physical test measure how athletic you are?

Validity

2. **Criterion-related validity**: Relationship between scores on a test and actual performance

- Predictive Validity: Refers to the function of a test in predicting a particular behavior or trait.
 - SAT Scores to College Grade
- Convergent Validity: The degree to which scores on a test correlate with (or are related to) scores on other tests that are designed to assess the same construct.
 - Is your SAT scores similar to ACT scores? Is there a relationship (correlation)?

Validity

3. **Construct Validity** refers to the ability of a measurement tool (e.g., a survey, test, etc) to actually measure the psychological concept being studied.

For example, if we want to know our height we would use a tape measure and not a bathroom scale because all height measurements are expressed in inches and not in pounds.

Intelligence Test

Assessing Intelligence

Psychologists define intelligence testing as a method for assessing an individual's mental aptitudes and comparing them with others using numerical scores.

Alfred Binet

Alfred Binet and his colleague Théodore Simon practiced a more modern form of intelligence testing by developing questions that would predict children's future progress in the Paris school system.



Alfred Binet

Alfred Binet's goal became measuring children's mental age, the level of performance typically associated with a certain age.

Identify children who will need extra need
(Special Ed.)



Lewis Terman

In the US, Lewis Terman adapted Binet's test for American school children and named the test the **Stanford-Binet Test**. The following is the formula of **Intelligence Quotient (IQ)**, introduced by William Stern:

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



Stanford-Binet scale

- Stanford-Binet scale
 - Terman's adaptation of the Binet-Simon scale
 - introduced the "I.Q." score
 - A score of 100 is considered average

Sample Questions

If a four-year-old girl correctly answered questions on an intelligence exam similar to a five-year-old girl, she would be said to have a mental age of five. In this case her intelligence quotient (IQ) would be:

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

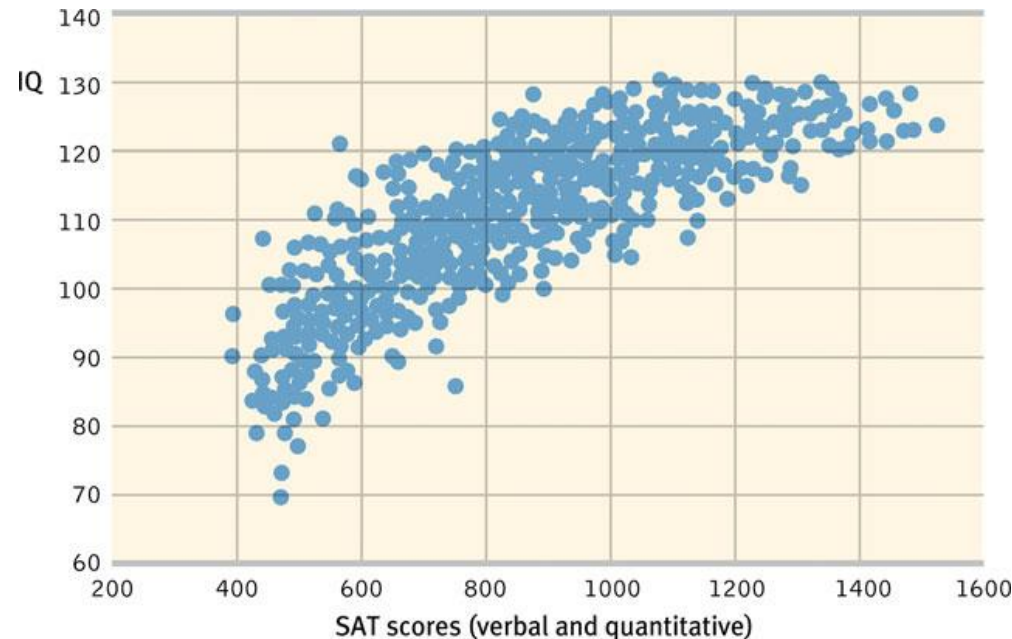
Sample Question

- If a five-year-old girl correctly answered questions on an intelligence exam similar to a 6-year-old girl, she would be said to have a mental age of five. In this case her intelligence quotient (IQ) would be:

Aptitude and Achievement Tests

Aptitude tests are intended to *predict* your ability to learn a new skill

achievement tests are intended to *reflect* what you have already learned.



Examples

Where would these exams fall under?

1.Pop Quizzes

2.Unit Exams

3.SATs (Enter College)

4.ACTs (Enter College)

5.ASVAB (Enter Military)

6.AP Exams

Examples

- Achievement Tests
 - Unit Exams and AP Exams (Assess what you have learned)
- Aptitude Tests
 - SATs and ACTs (predicts how well you will do in your first year in college)
 - ASVAB (What kind of job would be appropriate for you in the military)




David Wechsler

Wechsler developed the Wechsler Adult Intelligence Scale (WAIS) and later the Wechsler Intelligence Scale for Children (WISC), an intelligence test for preschoolers.



WAIS

WAIS measures overall intelligence and 11 other aspects related to intelligence that are designed to assess clinical and educational problems.

<p>VERBAL</p> <p>General Information What day of the year is Independence Day?</p> <p>Similarities In what way are <i>wool</i> and <i>cotton</i> alike?</p> <p>Arithmetic Reasoning If eggs cost 60 cents a dozen, what does 1 egg cost?</p> <p>Vocabulary Tell me the meaning of corrupt.</p> <p>Comprehension Why do people buy fire insurance?</p> <p>Digit Span Listen carefully, and when I am through, say the numbers right after me.</p> <p style="text-align: center;">7 3 4 1 8 6</p> <p>Now I am going to say some more numbers, but I want you to say them backward.</p> <p style="text-align: center;">3 8 4 1 6</p> <p>PERFORMANCE</p> <p>Picture Completion I am going to show you a picture with an important part missing. Tell me what is missing.</p> <p style="text-align: center;">'85</p> <table border="1" style="margin-left: auto; margin-right: auto;"><thead><tr><th>SUN</th><th>MON</th><th>TUE</th><th>WED</th><th>THU</th><th>FR</th><th>SAT</th></tr></thead><tbody><tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr><tr><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td></tr><tr><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td></tr><tr><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td></tr><tr><td>29</td><td>30</td><td></td><td></td><td></td><td></td><td></td></tr></tbody></table>	SUN	MON	TUE	WED	THU	FR	SAT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30						<p>Picture Arrangement The pictures below tell a story. Put them in the right order to tell the story.</p>  <p>Block Design Using the four blocks, make one just like this.</p>  <p>Object Assembly If these pieces are put together correctly, they will make something. Go ahead and put them together as quickly as you can.</p>  <p>Digit-Symbol Substitution</p> <p>Code</p> <table border="1" style="margin-left: auto; margin-right: auto;"><tr><td>△</td><td>○</td><td>◇</td><td>×</td><td>◇</td></tr><tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr></table> <p>Test</p> <table border="1" style="margin-left: auto; margin-right: auto;"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>1</td><td>5</td><td>4</td><td>2</td><td>1</td><td>3</td><td>5</td><td>4</td><td>1</td><td>5</td></tr></table>	△	○	◇	×	◇	1	2	3	4	5											1	5	4	2	1	3	5	4	1	5
SUN	MON	TUE	WED	THU	FR	SAT																																																																			
1	2	3	4	5	6	7																																																																			
8	9	10	11	12	13	14																																																																			
15	16	17	18	19	20	21																																																																			
22	23	24	25	26	27	28																																																																			
29	30																																																																								
△	○	◇	×	◇																																																																					
1	2	3	4	5																																																																					
1	5	4	2	1	3	5	4	1	5																																																																

Intelligence Tests

- The Wechsler Intelligence Scales
 - The Wechsler Adult Intelligence Scale – Third Edition is the most commonly used test of intelligence for adults
 - WAIS-III is divided into to parts, one that focuses on verbal abilities and one that focuses on performance skills
 - Also a version for children, Wechsler Intelligence Scale for Children – Third Edition

Intelligence Tests

- Individual Tests
 - Intelligence tests that are taken by one individual
 - IQ Tests and Personality Tests
- Group Tests
 - Intelligence tests that can be given to large groups
 - Advantages
 - Quick scoring
 - No examiner bias
 - Easier to establish norms
 - Disadvantages
 - Less likely to detect someone who is ill or confused
 - Might make people nervous
 - Learning disabled children often perform worse

Intelligence Tests

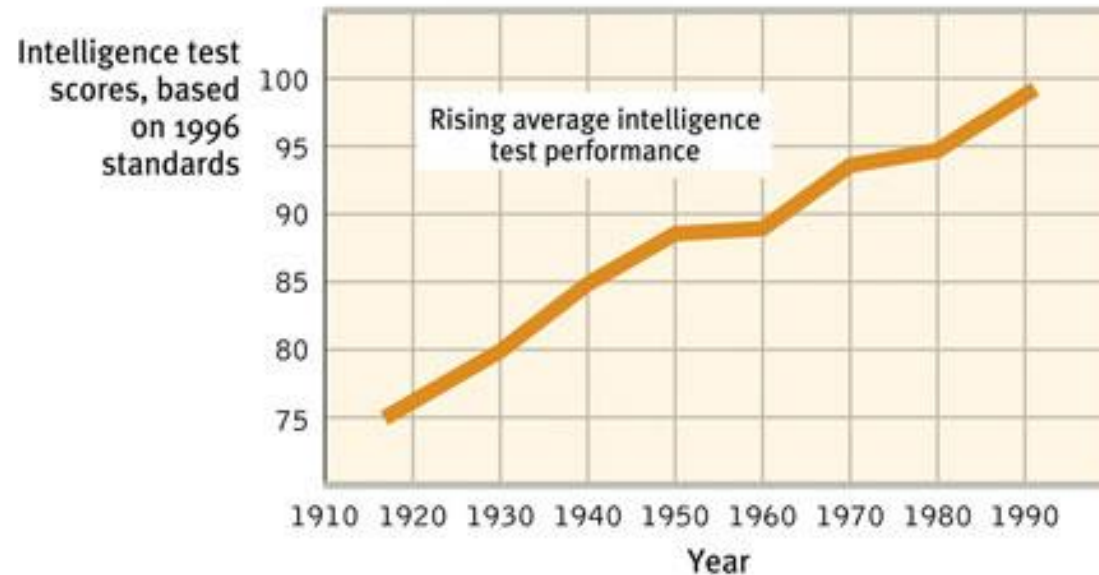
- Speed Tests
 - Large number of questions are given in a short period time
 - How quickly can you solve the problems?
- Power Tests
 - Consists of items that get progressively more difficult
 - To what degree of difficulty can you solve the problems?

Intelligence Tests

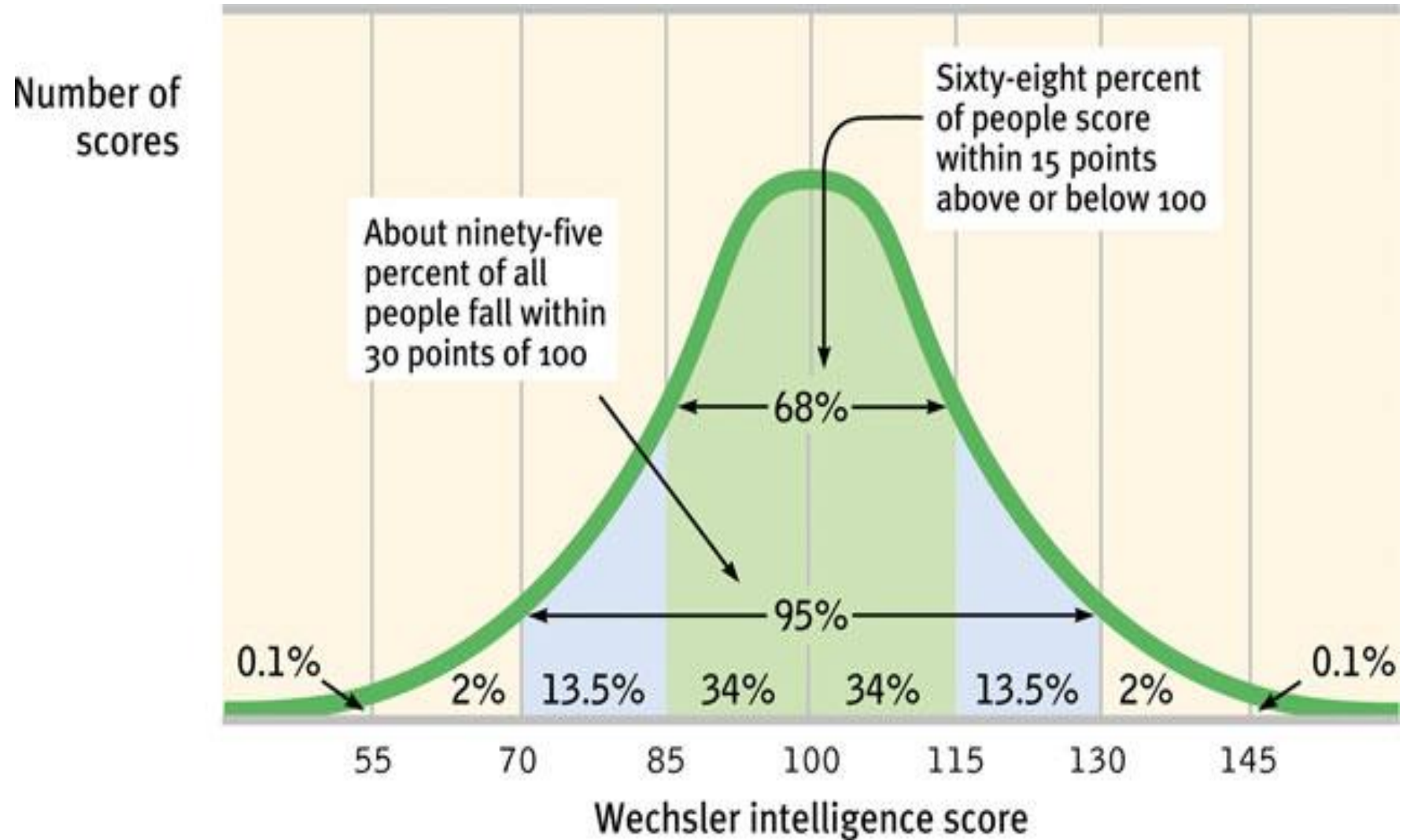
- Performance tests
 - Tests that minimize the use of language
 - Used to test very young children or people with retardation
 - Also can be used to test those unfamiliar with English
- Culture-fair tests
 - Tests designed to reduce cultural bias
 - Minimize skills and values that vary from one culture to another

Flynn Effect

In the past 60 years, intelligence scores have risen steadily by an average of 27 points. This phenomenon is known as the Flynn effect.



Review



OA 3

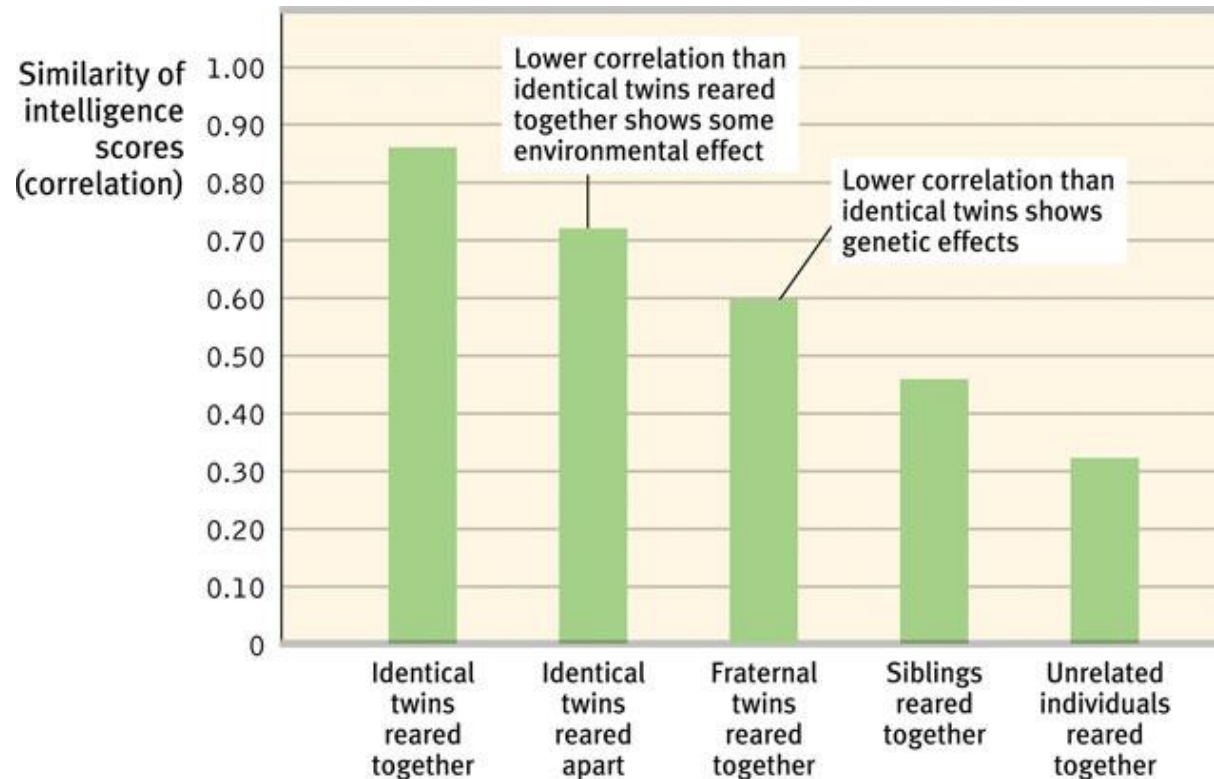
1. What % of people will receive 0-100 on the IQ test?
2. What % of people will receive 85-100 on the IQ test?
3. What % of people will fall under standard 1 deviation On the IQ test?
4. What % of people will receive 70-100 on the IQ test?
5. What % of people will fall under standard 2 deviation on the IQ test?
6. What % of people will receive 130- up on the IQ test?

Nature AND Nurture

Both Genetics & Environment play a
role in our intelligence

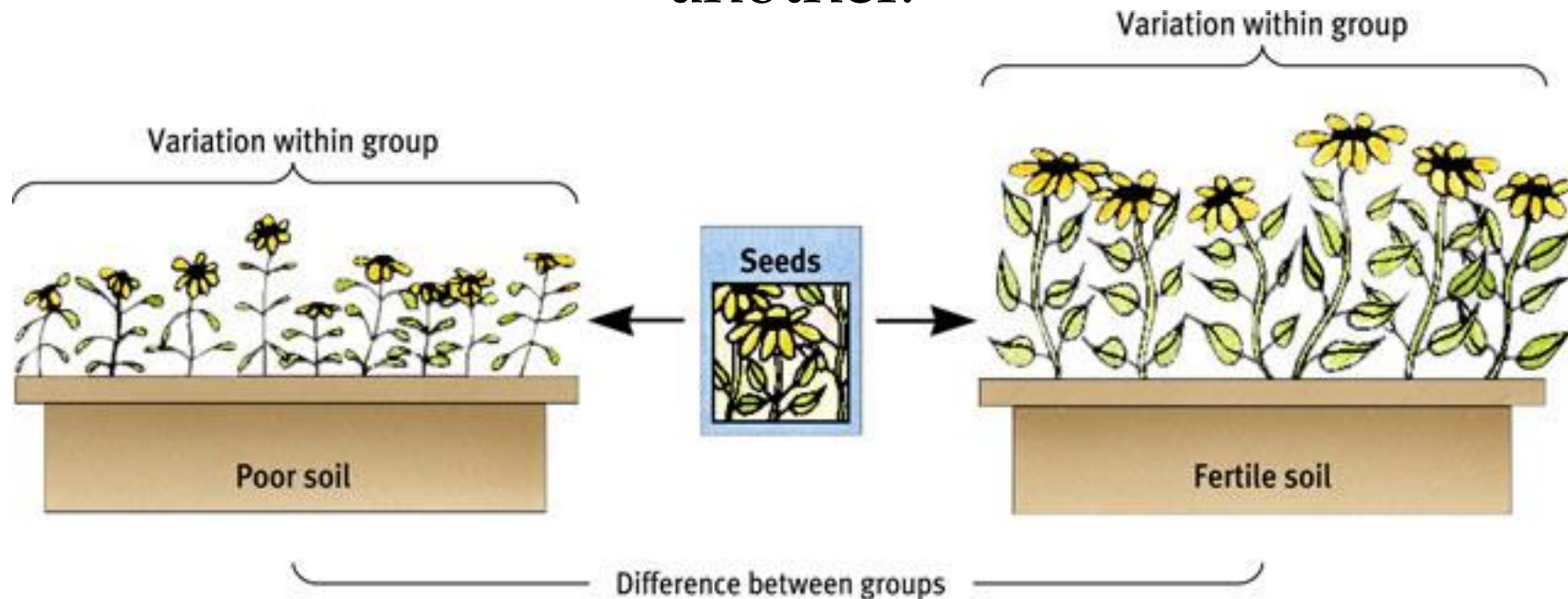
Genetic Influences

Studies of twins, family members, and adopted children together support the idea that there is a significant genetic contribution to intelligence.



Environmental Effects

Differences in intelligence among these groups are largely environmental, as if one environment is more fertile in developing these abilities than another.



Reaction Range

- Intelligence may increase or decrease as a result of environment
- Extent to which traits, abilities, or IQ scores may increase or decrease as a result of environmental factors
 - Have you received tutoring as a child?
 - Have your attended Preschool?
 - Do you think these factors would have had an impact on your abilities or IQ scores?

Early Intervention Effects

Early neglect from caregivers leads children to develop a lack of personal control over the environment, and it **impoverishes their intelligence**.



Romanian orphans with minimal human interaction are delayed in their development.

Schooling Effects

Schooling is an experience that pays dividends, which is reflected in intelligence scores. **Increased schooling correlates with higher intelligence scores.**



To increase readiness for schoolwork, projects like Head Start facilitate learning.

Stereotype Threat

- Situational predicament in which people are or feel themselves to be at risk of confirming negative stereotypes about their social group.
- Black students who was proctored by a white teacher scored poorly than black students who was proctored by a black teacher.
 - Real or imagined sense of discrimination

