Intelligence Testing: Wechsler and Stanford-Binet Scales

Wechsler Adult Intelligence Scale (WAIS)
- Tests made up of items that are appropriate for a wide range of ages
- Subject receives a certain # of points for each item answered correctly.
- 1939 – Wechsler-Bellevue Intelligence Scale – designed as adult-oriented point scale of intelligence
- 1955 – Wechsler Adult Intelligence Scale (WAIS) – most widely used individual test of adult intelligence

Characteristics of the WAIS-III
- 14 separate subtests
- 11 subtests are used to compute a full-scale IQ score, a Verbal IQ score, and a Performance IQ score
- Also yields more specific indices (new to WAIS-III):
  - Verbal Comprehension
  - Perceptual Organization
  - Working Memory
  - Perceptual Speed
Structure of the Scale

**WAIS–III Levels of Performance**

<table>
<thead>
<tr>
<th>FSIQ</th>
<th>VIQ</th>
<th>PIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary, Similarities, Information, Comprehension</td>
<td>Digit Span, Arithmetic, Letter–Number Sequencing</td>
<td>Block Design, Matrix Reasoning, Picture Completion, Picture Arrangement</td>
</tr>
</tbody>
</table>

**WAIS–III Subtests for IQ Scores**

<table>
<thead>
<tr>
<th>Verbal</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary, Similarities, Arithmetic, Digit Span, Information, Comprehension</td>
<td>Picture Completion, Digit Symbol—Coding, Block Design, Matrix Reasoning, Picture Arrangement</td>
</tr>
</tbody>
</table>

**WAIS–III Subtests for Index Scores**

<table>
<thead>
<tr>
<th>Verbal Comprehension</th>
<th>Perceptual Organization</th>
<th>Working Memory</th>
<th>Processing Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary, Similarities, Information</td>
<td>Picture Completion, Block Design, Matrix Reasoning</td>
<td>Arithmetic, Digit Span, Letter–Number Sequencing</td>
<td>Digit Symbol—Coding, Symbol Search</td>
</tr>
</tbody>
</table>

...
**Vocabulary**
- A list of up to 33 words is presented, both orally and in writing (range from concrete to abstract).
- The subject must explain or define the meaning of the word.
  - Factor analysis: loads on Verbal Comprehension
  - Measures: language development and word knowledge
  - Also affected by: cultural opportunities, foreign language background, interests, outside reading, school learning

**Similarities**
- Subject responds to open-ended questions by telling how two things are alike.
- Corresponds best to Spearman’s idea that ability to see relations is critical indicator of $g$.
  - Example: How are shoes and socks alike?
  - Factor analysis: loads on Verbal Comprehension
  - Measures: logical abstractive (categorical) thinking
  - Also affected by: flexibility, interests, negativism “they’re not alike”, overly concrete thinking
  - Of verbal tests, least affected by specific learning and experience

**Arithmetic**
- Set of problems that can be solved without pencil & paper; presented orally
- Thought of as a reasoning test, not test of quantitative performance test
  - Example: How many inches are in 4 feet?
  - Factor Analysis: Loads on Working Memory
  - Measures: computational skill, quantitative reasoning, mathematical achievement
  - Affected by: Attention span, anxiety, concentration, distractibility, learning disabilities, school learning and working under time pressure
**Digit Span**
- Examiner reads aloud list of 2 to 9 digits that subject must repeat.
- Subjects must repeat list in order and reverse order.
  - Factor analysis: loads on working memory
  - Measures: immediate rote recall, memory span, reversibility
  - Also affected by: attention span, anxiety, distractibility, flexibility, learning disabilities, negativism
  - Weakest correlation to $g, r=.57$

**Information**
- Includes items covering knowledge of general information that an adult in our culture could reasonable be expected to acquire (not supposed to be specialized or academic)
  - Example: “Who was Abraham Lincoln?”
  - Factor analysis: loads on Verbal Comprehension
  - Measures range of factual knowledge
  - Also affected by: alertness to environment, cultural opportunities, interests, outside reading, school learning

**Comprehension**
- Open-ended questions asking subject to explain why certain things should be done or by describing what he/she would do in certain situations
  - Example: Why are votes often taken via secret ballot?
  - Factor analysis: loads on Verbal Comprehension
  - Measures: demonstration of practical information, use of past experience, generalization, knowledge of conventional standards of behavior, social maturity, judgment
  - Also affected by: cultural opportunities, moral development, negativism, overtly concrete thinking
**Letter-Number Sequencing**

- New to WAIS-III
- Examiner presents series of letters and numbers in a mixed up order and subject is to repeat list saying numbers first in order, then letters in order
  - Example: X-4-B-2-H-1-C → 1-2-4-B-C-H-X
  - Factor Analysis: loads on Working Memory
  - Measures: facility with over-learned sequences, working memory
  - Also affected by: attention span, anxiety, concentration, distractibility, flexibility, illiteracy/dyslexia, ADHD, negativism, persistence

**Picture Completion**

- Shown pictures in which critical detail is missing and subject must tell what is missing in the picture.
  - Factor Analysis: loads on Processing Speed
  - Measures: flexibility of closure, visual alertness, visual recognition and identification (long-term visual memory)
  - Also affected by: Ability to respond when uncertain, alertness to environment, concentration, negativism, working under time pressure
Digit Symbol Coding

- Code substitution test: one of oldest, best established psychological tests
- A subject is given a code sheet that pairs nine symbols (e.g., #) with digits. The subject is given a list of 133 digits and, using the code sheet, is required to substitute the appropriate symbol for each digit. Strictly timed test (2 minutes)
  - Factor Analysis: loads on Processing Speed
  - Measures: ability to follow directions, clerical speed and accuracy, psychomotor speed, visual short-term memory
  - Also affected by: anxiety, compulsive concern for accuracy and detail, distractibility, persistence, working under time pressure

Block Design

- A set of 9 pictures with geometric designs in red and white corresponds to set of blocks. Subject must arrange 4 – 9 blocks to duplicate pictures.
- Measure of non-verbal intelligence
  - Factor Analysis: Perceptual Organization
  - Measures: analysis of whole into component parts, nonverbal concept formation, spatial visualization
  - Also affected by: visual-perceptual problems, working under time pressure

Example Item–Block Design
Matrix Reasoning
- Another new subtest, in which subjects identify recurring patterns or relationships between figural stimuli drawn on each of 26 cards
  - Factor Analysis: loads on Perceptual Organization
  - Measures: analogic reasoning, induction, nonverbal problem solving without time limit
  - Also affected by: ability to respond when uncertain, color blindness, flexibility, motivation level, negativism, overly concrete thinking, persistence, visual-perceptual problems

Sample Items—Matrix Reasoning

Picture Arrangement
- Given set of 3 to 6 cards, which when arranged in proper order tell a story. Cards presented out of sequence and subject is to arrange in proper order.
  - Factor Analysis: loads on Perceptual Organization
  - Measures: anticipation of consequences, planning ability, temporal sequencing and time concepts
  - Also affected by: creativity, cultural opportunities, exposure to comic strips, working under time pressure
Symbol Search

- New subtest in which subjects scan a series of paired groups of symbols, with each pair consisting of a target group and a search group. The subject's task is to indicate, as quickly as possible, whether the target symbol also appears in the search group.
  - Factor Analysis: loads on Processing Speed
  - Measures: speed of visual search, perceptual speed
  - Also affected by: anxiety, distractibility, ADHD, motivation level, obsessive concern with accuracy and detail, persistence, visual-perceptual problems, working under time pressure

Object Assembly

- Flat cardboard representations of common objects cut up to make puzzle – assembled by subject
- Optional subtest
  - Factor Analysis: loads on perceptual organization
  - Measures: closure speed, ability to benefit from sensory-motor feedback, anticipation of relationships among parts
  - Also affected by: ability to respond when uncertain, experience with puzzles, flexibility, persistence
### Administration
- Given individually by a trained examiner
- Subtests given one at a time, alternating between verbal and performance tests
- Items arranged from easy to difficult
- Don’t give all items – have different starting points for different ages
- Reversal rules – if don’t pass first item, administer preceding items in reverse order until 3 in a row are passed

### Scales and Norms for the WAIS
- Determine raw score for each subtest.
- Convert raw scores to standard scores, called scaled scores (M=10, SD=3)
- There are conversions for 13 age groups. This method of conversion obscures any differences in performance by age.
- Subtest scaled scores are added, then converted to WAIS-III composite scores.
- Three composite scores: verbal, performance, full scale, each with M=100, SD=15
- Four index scores: verbal comprehension, perceptual organization, working memory, processing speed

### Converting Raw Scores to Scaled Scores

<table>
<thead>
<tr>
<th>SUBTESTS</th>
<th>Raw Score</th>
<th>VERBAL</th>
<th>PERF.</th>
<th>VIQ</th>
<th>PI Q</th>
<th>WI Q</th>
<th>PS</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picture Completion</td>
<td>22</td>
<td></td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>56</td>
<td>7</td>
<td>12</td>
<td>17</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Digit Symbol–Coding</td>
<td>54</td>
<td>5</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Analogies</td>
<td>29</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Digit Symbol–Coding</td>
<td>54</td>
<td>5</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Similarities</td>
<td>29</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Digit Symbol–Coding</td>
<td>54</td>
<td>5</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Arithmetic</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Matrix Reasoning</td>
<td>20</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
Converting Sums of Scaled Scores to IQ and Index Scores

<table>
<thead>
<tr>
<th>IQ/INDEX SCORES</th>
<th>VIQ</th>
<th>PIQ</th>
<th>FSIQ</th>
<th>VCI</th>
<th>PGI</th>
<th>WMI</th>
<th>PSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of Scaled Scores</td>
<td>101</td>
<td>106</td>
<td>103</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentiles</td>
<td>93</td>
<td>98</td>
<td>94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percentiles: 52, 66, 58

Confidence Intervals: 96%, 99%, 99%

Plotting Subtest Scores (IQ)

Plotting Subtest Scores (Index)
Evaluation of the WAIS-III

- Are the norms adequate?
  - Norms determine how well a test-taker did in comparison to others.
- Is there evidence to support reliability?
- Is there evidence to support validity?

Standardization of the WAIS

- Standardized on a stratified sample of 2,450 adults representative of the US population aged 16-89.
- There were 200 cases per age group, except for the smaller numbers in the two oldest groups.
- Still difficult to know the effects of self-selection since participants had to be invited and accept to be included.

Reliability of the WAIS

- Internal consistency and test-retest reliabilities are about .95 or higher for full scale and verbal scores.
- They're about .90 for performance and three other index scores: perceptual organization, working memory, and processing speed.
- Internal consistency reliability for the subtests range from upper .70s to low .90s. Test-retest is about .83.
- Generally, performance reliabilities are lower than verbal reliabilities on the subtests.
Validity of the WAIS

- Great deal of information on criterion-related and construct validity.
- Factors analyses support use of 4 index scores.
- Comparison studies show the pattern of WAIS-III scores for many special groups, e.g., Alzheimer’s Disease, Parkinson’s, learning disabled, brain injury.
- Is the top test used today

Common Characteristics of Individual Intelligence Tests

- individually administered
- administration requires advanced training
- tests cover wide range of age and ability
- examiner must establish rapport
- immediate scoring of items
- usually requires about one hour
- allows opportunity for observation

Two Main Individually Administered Intelligence Tests

Stanford-Binet
- He wanted to create a process for identifying intellectually limited children so they could be removed from the regular classroom and put in special education.

Wechsler scales
- Developed in response to the perceived shortcomings of the Stanford-Binet
Stanford-Binet Fifth Edition

<table>
<thead>
<tr>
<th>Fluid Reasoning</th>
<th>Verbal</th>
<th>Nonverbal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Fluid Reasoning</td>
<td>Early Reasoning (2–7)</td>
<td>Nonverbal Fluid Reasoning</td>
</tr>
<tr>
<td>Verbal Abstractions (4)</td>
<td></td>
<td>Object-Sets/Reasoning</td>
</tr>
<tr>
<td>Verbal Analogies (5–6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>Verbal Knowledge</td>
<td>Nonverbal Knowledge</td>
</tr>
<tr>
<td>Verbal Vocabulary</td>
<td></td>
<td>Procedural Knowledge (2–3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Picture Analogies (4–6)</td>
</tr>
<tr>
<td>Quantitative Reasoning</td>
<td>Verbal Quantitative</td>
<td>Nonverbal Quantitative</td>
</tr>
<tr>
<td>Reasoning (2–6)</td>
<td>Quantitative Reasoning (2–6)</td>
<td></td>
</tr>
<tr>
<td>Visual-Spatial Processing</td>
<td>Visual/Spatial Processing</td>
<td>Visual/Spatial Processing</td>
</tr>
<tr>
<td></td>
<td>Position and Direction (2–6)</td>
<td>Forms Board (1–2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Forms Patterns (3–6)</td>
</tr>
<tr>
<td>Working Memory</td>
<td>Verbal Working Memory</td>
<td>Nonverbal Working Memory</td>
</tr>
<tr>
<td></td>
<td>Memory for Sentences (2–3)</td>
<td>Delayed Response (1)</td>
</tr>
<tr>
<td></td>
<td>Last Word (4–6)</td>
<td>Block Span (2–4)</td>
</tr>
</tbody>
</table>

Stanford-Binet scale:
- Subtests have mean of 10 and SD of 3
- Composite has mean of 100 and DS of 15
- **Adaptive** test – test taker only takes the part that is appropriate for his/her developmental level
- Starting points determined by [routing tests](#)

- Standardization sample has 4,800 subjects ages 2 to 85+.
- Sample stratified based on 2000 census – age, sex, race / ethnicity, geographic region, and socioeconomic level.
- Split-half reliability is .98 for composite
- The five factor index scores were all above 90, and were higher than the subtest scales, which were however comparable to other cognitive tests with ranges from .84 to .89.
- Factor analysis supports the structure of the test.
- Correlations with other IQ tests ranged from .78 to .84
WISC-IV (2003)

- Wechsler Intelligence Scale for Children
- Is the most popular test for assessing intellectual ability of children ages 6 years, 0 months to 16 years, 11 months.
- Similar to structure of the WAIS, with items in order of difficulty (a few subtests are different)
- Yields 4 composite scores and a full-scale IQ
  - Verbal Comprehension
  - Perceptual reasoning
  - Working Memory
  - Processing Speed

Psychometric Properties of the WISC-III

- Standardization program involved 2,200 cases selected to represent the US population of children aged 6-16, based on 2000 census.
- Composite scores generally have internal consistency reliabilities in the mid-.90s and test-retest reliabilities ranged from .79 (Processing Speed) to .89 (Full Scale and Verbal Comprehension).
- Subtest reliabilities are generally in the mid-.80s.