Standardized Tests, Intelligence & IQ, and Standardized Scores



Alphabet Soup!!

ACT's SAT's ITBS
GRE's
WISC-IV WAIS-IV
WRAT
MCAT LSAT

IMA RAT

Uses/Functions of Standardized Tests

- **Selection and Placement**
- **Diagnosis**
- **Evaluation of Progress/Effectiveness**
- **Program Evaluation/(School Improvement)**
- **Accaccountability**

Types of Standardized Tests

- **Aptitude Tests**
  - designed to assess general abilities
  - predict future performance (learning or task).
  - examples: WISC-III, ACT, GRE
- **Achievement Tests**
  - assess what has been learned
  - examples: ITBS, PIAT,

Types of Standardized Tests (cont.)

- **Norm-Referenced Tests**
  - compares individual performance to group norms
  - norms : average or typical group scores obtained from a specific sample in test development
  - answers question: How well did this person do in comparison to other similar persons?

Types of Standardized Tests (cont.)

- **Criterion-Referenced Tests**
  - measures the extent to which a student has mastered a specific set of learning objectives
  - compared to a standard (criterion) of mastery, not a norm group
  - Answers the question: How close did the person come to meeting the standard of mastery
  - Examples: teacher-developed tests, developmental screeners

Intelligence & IQ Tests

Scores are part of basis for decisions:

1. **Disabled? - Entitled to SSI benefits?**
2. **Diagnosis of learning and psych. problems**
3. **Eligibility & placement decisions - special ed.**
4. **Selection of individuals into Army, Navy, etc.**
5. **Job selection and promotion, etc. etc.**

- **still disagreement about what intelligence is.**
- **Disagreement about how to measure it!!**
- **Intelligence is a construct.**
- **NOT directly observable…must be inferred from overt behavior.**

Intelligence - acting or thinking in ways that are goal-directed and adaptive.

Theorists agree that intelligence is:

- Adaptive – used flexibly to respond to various situations and problems
- Is related to learning ability
- Involves use of prior knowledge to analyze and understand new situations effectively
- Involves many different mental processes
- Is culture-specific

- **Psychometric theories -- intelligence made up of mental factors**
- **e.g., Verbal factor**
- **Statistical tests - factor analyses**
- **Or, e.g., a Performance Factor**

Spearman (1927) - 2 kinds of factors:

- **general factor (g) which influences performance on all intellectual tasks.**
- **specific factors (s) to a certain task.**

Guilford (1967, 1988) - 180 Factors

- **6 Mental Operations**
- **5 Contents**
- **6 Products**

Thurstone (1938) - 7 primary mental abilities

1. **verbal comprehension**
2. **verbal fluency**
3. **number**
4. **spatial visualization**
5. **memory**
6. **reasoning**
7. **perceptual speed**

Catell's (1963, 1971) fluid and crystallized abilities

- **Fluid Intell.**
- **Crystallized**

- **1st IQ test (France) - *Binet* and *Simon* 1905, 1908).**
- **In USA, *Lewis Terman* (1916) - Stanford-Binet.**
- ***David Wechsler* (1940's) - Wechsler scales**
  - **most widely used individually administered IQ tests today:**
- **WISC IV (Wechsler Intelligence Scale for Children - IV)**
- **WAIS IV (Wechsler Adult Intelligence Scale - IV)**
- **WPPSI-III (Wechsler Preschool & Primary Scale of Intelligence - III)**

Old Formula for IQ:

- **IQ = MA/CA X 100 (it's a quotient!!)**
  - **i.e. 10 year old child**
- **What is IQ if:**
  - **MA = 10??**
  - **MA = 7??**
  - **MA = 14???**
- **IQ is now a *deviation IQ* - a standard score.**
- **Mean IQ = 100; standard deviation = 15**

WISC - IV

This is the 4th edition of:

- **The Wechsler Intelligence Scale for Children**
- **WISC – 1949**
- **WISC-R – 1974**
- **WISC-III – 1991**
- **WISC-IV -- 2003**

WISC - IV

- **Individually administered**
- **For Children 6 - 16 years of age**
- **1 hour to 1 1/2 hours (avg.=1 hour, 15 min.)**
- **Standardized on 2,200 children**
- **Stratified Sample:**
  - Age
  - Sex
  - Race
  - Parent education level
  - Geographic region

WISC IV

10 Main Subtests
5 Alternative/Supplemental Subtests

Organized into 5 Composite Scores

Verbal Comprehension Index (3 & 2)
Perceptual Reasoning Index (3 & 1)
Working Memory Index (2 & 1)
Processing Speed Index (2 & 1)
Full Scale IQ (10 main subtests)
(Mean/Avg. of each = 100; s.d. = 15)

WISC - IV

Verbal Comprehension
Similarities
Vocabulary
Comprehension
(Information)
(Word Reasoning)

Perceptual Reasoning
Block Design
Picture Concepts
Matrix Reasoning
(Picture Completion)

Working Memory
Digit Span
Letter-Number Sequencing
(Arithmetic)

Processing Speed
Coding
Symbol Search
(Cancellation)

Comments about IQ & IQ Tests

Test Reliability

- 5 to 8 points - No test is perfectly reliable.

Test Validity

- IQ tests are supposed to predict success in school.
- And, IQ tests do this job pretty well

- Environment can produce much variablity in IQ scores
- Intellectual perfomance based on:
  - biological factors
  - general education
  - life experiences
  - motivation
  - personality
- IQ is not like hat size - It's a range of performance (IQ +/- 5-8 pts)

Misconceptions about IQ Tests

1. **IQ tests measure innate intelligence - NOT TRUE**
2. **IQ's are fixed and never change - NOT TRUE**
3. **Intelligence tests are perfectly reliable - NOT TRUE**

"There is a 90% chance that the child's IQ falls between __ and __."

4. **IQ tests measure all we need to know about a person's intelligence - NOT TRUE**
5. **IQ's obtained from a variety of tests are interchangeable - NOT TRUE**

Derived or Transformed Scores

- **<u>Percentile Ranks</u> - percentage of students in norm group that scored lower than a particular score (<u>not</u> percent correct)**
- **<u>Grade Equivalents</u> - relate students' raw score to average scores obtained by norm group at different grade levels *(Beware!!)***
- ***GE = 6.3 ⟶ 3rd month of 6th grade***
- **2nd grader obtains GE = 5.8????**

Standard Scores

- **Definition** - derived scores that are based on their position on the normal curve.
- **Normal Curve/Distribution** - symmetrical distribution of scores with the majority falling near the **mean** (average) and progressively fewer away from the mean.
- We know and/or can predict things about normally distributed scores

- **Mean = avg.** (sum and divide by # of scores)

  Example: (4, 90, 5, 1)

  Mean (Avg.) = 100/4 = 25

- influenced by extreme scores

- **Median**

(Odd #) Middle score when in ranked order

(Even #) Mid-point between 2 middle scores in ranked order

Example: (4, 2, 7, 1, 22)

(1, 2, 4, 7, 22 → median = 4)

Example: (4, 90, 5, 1)

(1, 4, 5, 90 → median = 4.5)

Example: (8, 2, 12, 6)

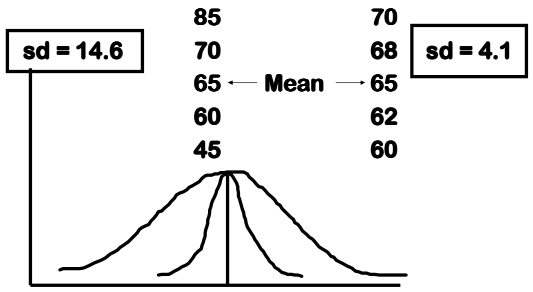
(2, 6, 8, 12 → median = 7)

- Median ***not*** as influence by extreme scores
- **Standard Deviation** - measure of how the scores "spread out" around the mean
- IQ's have mean = 100; sd = 15

1. Median???
8, 2, 6, 1, 3

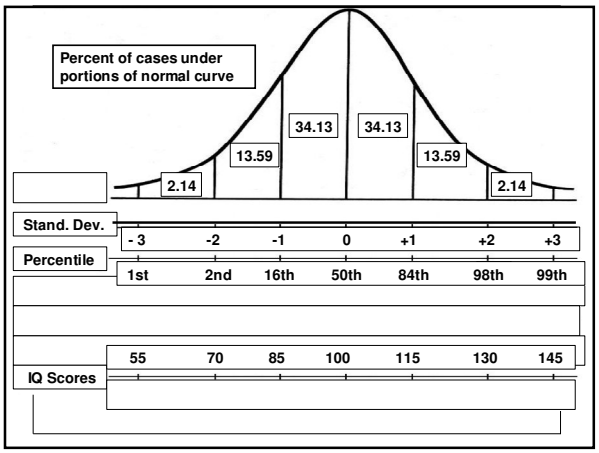
2. An 8-year-old boy obtains a mental age of 8 years on an IQ test. What is his IQ according to the "old formula?"

Distributions with different sd's



Important Standard Scores

- Mean IQ = 100 (sd = 15)
- 68% of IQ's are between +1 and -1 sd
- (68% of IQ's are between 85 - 115)
- 84th %tile = 115 (1 sd > mean)
- 50th %tile = 100 (Mean/Avg)
- 16th %tile = 85 (1 sd < mean)
- 2nd %tile = 70 (2 sd's < mean)
- 98 th %tile = 130 (2 sd's > mean)



- 4,1,3,9,2
- What is Median???

- 3,5,1,2,9,10
- What is Median???

- 50% of IQ scores are below what score
- A student scores 1 standard deviation below the mean on an achievement test – what is his/her percentile score??

- 1000 people take an IQ test (mean = 100 & s.d. = 15)….approx. how many would we expect to score between 85 and 115??
  - A student scores at the ___ percentile on I Q test…..what is his score??
  - 16th percentile = ? IQ
  - 2nd percentile = ? IQ
  - 98th percentile = ? IQ
  - 84th percentile = ? IQ

- Mike takes an achievement test. The mean of scores is 300, and the standard deviation is 25. If Mike does better than 84% of the people his age who take the test, his standard score is….

Test Reliability

- **Definition - consistency or accuracy of test scores by same person at different times, with different sets of equivalent items, or across test items.**
  - Reliability increases with more items
  - Usually expressed as a reliability coefficient (correlation!!) on scale from .00 to 1.00;
  - r=.90 (High)/r=.60 (Moderate)/r=.30 (Low)
  - r=.80 and above acceptable

- The reliability coefficient expresses the degree to which there is consistency in measurement of the test scores.
- Reliability is also related to:

Error of Measurement (s.e.m.)

3 Major Types of Reliability

1. Test-Retest Reliability
2. Alternate Form Reliability
3. Internal Consistency Reliability

Test Validity

Definition - degree to which a test measures or accomplishes what it was supposed to measure or accomplish.

- No test is simply valid or not - rather it is said to be valid or not for a specific purpose.

Test Validity

3 Major Types of Validity

1. **Content Validity**
2. **Construct Validity**
3. **Criterion-Related Validity**
   a. Predictive
   b. Concurrent

- Validity implies reliability, but not vice versa

End of
Standardized Tests
Ch 14