TESTING AND INDIVIDUAL DIFFERENCES

AP Psychology
CHAPTER LEARNING TARGETS:

- AP students in psychology should be able to do the following:
  - Define intelligence and list characteristics of how psychologists measure intelligence:
    - abstract versus verbal measures;
    - speed of processing.
  - Discuss how culture influences the definition of intelligence.
  - Compare and contrast historic and contemporary theories of intelligence (e.g., Charles Spearman, Howard Gardner, Robert Sternberg).
  - Explain how psychologists design tests, including standardization strategies and other techniques to establish reliability and validity.
  - Interpret the meaning of scores in terms of the normal curve.
  - Describe relevant labels related to intelligence testing (e.g., gifted, cognitively disabled).
  - Debate the appropriate testing practices, particularly in relation to culture-fair test uses.
  - Identify key contributors in intelligence research and testing (e.g., Alfred Binet, Francis Galton, Howard Gardner, Charles Spearman, Robert Sternberg, Louis Terman, David Wechsler).
TODAY’S AGENDA:

1. Problem solving
2. Intelligence test examples
3. Discussion
4. Child Prodigy Video
INTELLIGENCE DISCUSSION:

1. What did you find easy or difficult about the various tests?
2. Do you think these are a good indicator of intelligence? Why or why not?
3. Do you think this test would be appropriate for children as well as adults? Why or why not?
4. If you had to make up an intelligence tests what would it look like?
5. Why didn’t you see any real math or science on these tests?
- Child Prodigies – part 1
- Child Prodigies – part 2
- Any reaction to these video examples?
Lesson One: Objectives

By the end of this lesson, I will be able to:

1. Explain how psychologists design tests, including standardization strategies and other techniques to establish reliability and validity.
DISCUSSION:

- Tests are a part of your everyday life. Some of them even help decide your future choices (ACT)

Questions:
- What makes a good test?
- How do we make sure the test is reliable?
- How do we make sure the test is valid?
- Why is this important?
STANDARDIZATION AND NORMS:

- **Psychometrics** – the measurement of mental traits, abilities, and processes.
- A **psychometrician** would develop tests that would measure some construct or behavior that distinguished people.
- You can’t measure honesty or happiness in feet or meters, so we need other methods of measurement.
PSYCHOLOGICAL TESTS:

- Test abilities, interests, creativity, personality, and intelligence.
- Questions that everyone answer correctly or incorrectly are thrown out – not specific enough
- So, what does a good test look like?
GOOD TESTS:

- Standardized, reliable, and valid
- **Standardization** – established test norms from a large representative sample and they are scored the same way
- Standardized tests are written, edited, pretested, and re-written.
HOW DO WE ESTABLISH NORMS?

- Established from the test results of representative samples
- They are then used for subsequent test takers
- Then you can compare the scores of all test takers
WHAT MAKES A GOOD TEST - RELIABILITY

- If a test is reliable, we should obtain the same score no matter where, when, or how many times we take it

Types of Reliability:

1. **test-retest** – take the test 2X (familiarity)
2. **Split half** - compare evens and odds
3. **Alternative form** – different versions are compared

Note – if the same people took both exams and got the same score = high interrater reliability
WHAT MAKES A GOOD TEST: VALIDITY

- **Validity** – is the test measuring or predicting what it’s supposed to?
- **Example**: Class test that doesn’t test what you’ve talked about in class

**Five types of validity:**
- 1. Face
- 2. Content
- 3. Criterion
- 4. Predictive
- 5. Construct
Face Validity:

- **Face validity** – a measure of the extent to which the content of the test measures all of the knowledge or skills that are supposed to be included on the test. According to the test takers.

- **Example**: You expect that 8-10% of the Cognition material should be on the AP Test.
Content Validity:

- **Content Validity** – same as face validity except “according to the experts.
- **Example**: Tests are completed by experts and sample populations to make sure that they are measuring what they are supposed to
**Criterion Related Validity:**

- **Criterion Related Validity** – How does the test correlate with other tests being given?
- **Example:** Do Mr. B’s tests compare to the AP Psych test?
Predictive Validity:

- **Predictive Validity** – Does the test predict future success?
- **Example**: High ACT = success in college
CONSTRUCT VALIDITY:

- **Construct Validity** – Does the test measure the specific construct / behavior it is supposed to?
- This is arguably the most important type of validity
Testing in Schools:

- Our educational system is based around testing.
- Do the tests really show student learning?
- Video Example:
- Video Example #2:
Lesson Two: Types of Tests

By the end of this lesson, I will be able to:

1. Explain how psychologists design tests, including standardization strategies and other techniques to establish reliability and validity.

2. Debate the appropriate testing practices, particularly in relation to culture-fair test uses.
I feel that standardized tests are the most reliable and valid form of testing.

1. Yes
2. No
3. Maybe
4. Other
There are three types of tests we will be looking at in this chapter:

1. Performance
2. Observational
3. Self-report
A STANDARDIZATION SAMPLE FOR DEVELOPING A TEST:

1. Should be representative of all the types of people for whom the test is designed
2. Is an early version of the test to determine questions that differentiate individuals
3. Is a set of norms that will determine what score should be considered passing
4. Should include people from all different age groups, ethnic groups, and genders
5. Must include a standard set of directions for administering the test that all students will receive
**PERFORMANCE TESTS:**

- **Performance tests** – the test taker knows what they should do in response to questions or tasks on the test
- Examples: ACT, SAT, Midterm, Finals
- Types:
  1. **Speed** – how many can you answer correctly in a short period of time
  2. **Power** – how many can you answer (varying difficulty) – no time/large limit
**Observational Tests:**

- **Observational tests** – person being tested does not have a single, well defined task to perform
- Instead, they’re assessed on typical behavior or performance in a specific context
- **Examples:** employment interviews, on the job observations or evaluations
EXAMPLE OF RHS TEACHER OBSERVATIONAL TEST:

- 1. Demonstrates knowledge of content
- 2. Creates a climate that promotes fairness
- 3. Makes learning goals and instructional procedures clear to students
- 4. Uses instructional time effectively (time on task)
- 5. Incorporates and implements technology usage in the classroom
IF MRS. DELVECCHIO COMPARED THE SCORES OF STUDENTS ON THE ODD-NUMBERED QUESTIONS ON THE TEST WITH THEIR SCORES FOR THE EVEN-NUMBERED QUESTIONS, SHE WOULD BE ATTEMPTING TO DETERMINE IF THE TEST HAD:

1. Content validity
2. **Split-half reliability**
3. Predictive validity
4. Test-retest reliability
5. Concurrent validity
SELF REPORT TESTS:

- Self-report test – requires the test taker to describe his/her feelings, attitudes, beliefs, mental state, etc. on a survey or poll
ABILITY, INTEREST, AND PERSONALITY TESTS:

- **Aptitude tests** – designed to predict a person’s future performance or capacity to learn (SAT)
- **Achievement tests** – designed to assess what a person has already learned (AP Psychology Test)
GROUP VS. INDIVIDUAL TEST:

- **Group tests** – Mass testing (low cost)
- **Individual tests** – individual tests with an examiner

NOTE: Psychologist must be careful when administering these tests (trust is an issue)
APTITUDE TESTS ARE DESIGNED TO MEASURE:

1. Previously learned facts
2. Future performance
3. Previously learned skills
4. Current competence
5. Your IQ score
Lesson Three: Ethics and Standards in Testing

By the end of this lesson, I will be able to:

1. Debate the appropriate testing practices, particularly in relation to culturally fair test uses.
DISCUSSION STARTER: TURN AND TALK

1. What does “Ethics in testing” mean to you?
2. What kinds of considerations must you make sure to follow when testing?
3. Do you think that tests can be biased towards specific groups of people?
The APA and Testing:

- The APA has specific guidelines detailing appropriate technical and professional standards for:
  - 1. Construction of tests
  - 2. Evaluation of tests
  - 3. Interpretation of tests
  - 4. Application of tests
WHY DO THEY TAKE THESE MEASURES?

- 1. Promote the welfare and best interest of the client
- 2. Guard against the misuse of assessment results
- 3. Respect the client’s right to know the results
- 4. Safeguard the dignity of test takers
Some groups, over time, have scored lower on intelligence tests and other standardized tests (ACT’s).

Critics have argued that these tests are culturally biased.

**Culture Relevant Tests** – questions are based on cultural experiences of the specific test takers.

Is this necessary?

Click the Pic!!!
Test Examples and Discussion Questions:

- How would you feel if this test was used as a standardized intelligence test?
- Is this intelligence tests fair? Why or why not?
- What have you learned from taking this test in terms of how non-middle-class individuals might feel about typical standardized tests?
- How do you measure someone who scores at an IQ of only 90 on a standardized intelligence test but very well on one of the other intelligence tests?
Lesson Four: Intelligence Tests

By the end of this lesson, I will be able to:

1. Define intelligence and list characteristics of how psychologists measure intelligence:
   A. abstract versus verbal measures;
   B. speed of processing.

2. Discuss how culture influences the definition of intelligence.

3. Compare and contrast historic and contemporary theories of intelligence (e.g., Charles Spearman, Howard Gardner, Robert Sternberg).
HOW DO YOU DEFINE INTELLIGENCE?

- **Intelligence** – the capacity of an individual to act purposefully, think rationally, and deal effectively with their environment.

- **Behaviors that indicate intelligence:**
  1. ability to learn from experience
  2. ability to solve problems
  3. ability to use information to adapt to the environment
  4. ability to benefit from training
**History of Intelligence Tests:**

- **Francis Galton** – Measurement of Psychosocial Performance (Darwin’s cousin)
- He believed that people with excellent physical abilities are better adapted for survival → highly intelligent
FRANCIS GALTON: cont.

- How did he measure this?
  - 1. Strength
  - 2. Reaction time
  - 3. Sensitivity to pain
  - 4. Weight discrimination

Problem: How does this correlate with reasoning ability?
ALFRED BINET’S MEASUREMENT OF JUDGMENT:

- Binet believed that we answer questions differently depending on our age.
- He assigned children with a “mental age” – this reflects the age at which children typically give similar responses.
- Problem: a 6 year old and an 8 year old each 2 years behind would be proportionally different (6 year old would be further behind).
MENTAL AGE AND INTELLIGENCE QUOTIENT:

- Lewis Terman – Stanford-Binet intelligence scale
- The result of this test = your IQ
- MA / CA X 100
- You take a test – Mental age
- How old are you? – Chronological age
- Multiply X 100 = your IQ
- Present tests are standardized with age groups up to 90 years old
Newest Version of Stanford-Binet:

- Assess five ability areas:
  1. Knowledge
  2. Fluid reasoning
  3. Quantitative reasoning
  4. Verbal Reasoning
  5. Non verbal reasoning
Wechsler Intelligence Scale:

- David Wechsler has developed three age-based intelligence tests:
  - 1. **WPPSI** – Wechsler Preschool and Primary Scale of Intelligence
  - 2. **WISC** – Wechsler Intelligence Scale for Children
  - 3. **WAIS** – Wechsler Adult Intelligence Scale
MORE ABOUT WECHSLER:

- **WAIS III** (adults and adolescents) – latest edition that tests:
  - 1. Vocabulary
  - 2. Similarities
  - 3. Object assembly
  - 4. Block design
  - 5. Picture arrangement and completion

Results are based on deviations (IQ score)

68% of the population will have an IQ score between 85 and 115
DNA AND IQ:

- Video Example:
LESSON FIVE: MENTAL RETARDATION

By the end of this lesson, I will be able to:

1. Describe relevant labels related to intelligence testing (e.g., gifted, cognitively disabled).
**History of Mental Retardation:**

- The Romans used people with MR as forms of amusement.
- This practice started to fade with the introduction of Christianity throughout Europe.
- In 1876, the American Association on Mental Retardation was formed – studied cases of MR.
- Most cases of MR result from chromosomal abnormalities (Down Syndrome).
- The other cases usually involved head trauma.
MENTAL RETARDATION:

- MR or Cognitively Delayed ranges from mild → profound
- Each level of MR requires different levels of assistance
- The vast majority of MR cases are mild (85%)
Mild Retardation: (85%)

- Between 50-70 IQ
- Can usually care for themselves and their homes
- Can achieve a 6th grade education
- Can become an adequate parent (I am Sam)
- Are often mainstreamed into regular education classroom (integration)
- How do you feel about mainstreaming?
MODERATELY RETARDED: (10%)

- Between 35-49 IQ
- Can achieve a 2\textsuperscript{nd} grade education
- Need to be given training for personal skills (hygiene, dressing, etc.)
- Need training to be able to do basic tasks or jobs
SEVERELY RETARDED: (3-4%)

- Between 20-34 IQ
- Limited vocabulary
- Limited self-care skills
- Need assistance with most things
- Very basic education
PROFOUNDLY RETARDED: (1-2%)  

- IQ below 20  
- Require full time care  
- Many live in group homes  
- There has been a push for deinstitutionalization – out of group homes and into regular homes with their families  
- Problem – Some families are unable to take care of them (as they get older)
Savants: 

- **Savants** – individuals that are otherwise considered mentally retarded, have a specific, exceptional skill usually in art, music, or calculation (about half are autistic)

- **Example:** Raymond Babbitt (Rainman)

- Knows 7600 books by heart

- Knows all US area codes, Zip codes, and TV stations
DOWN SYNDROME: A CLOSER LOOK

- Video Example:
LESSON SIX: KINDS OF INTELLIGENCE

By the end of this lesson, I will be able to:

1. Compare and contrast historic and contemporary theories of intelligence (e.g., Charles Spearman, Howard Gardner, Robert Sternberg).

2. Identify key contributors in intelligence research and testing (e.g., Alfred Binet, Francis Galton, Howard Gardner, Charles Spearman, Robert Sternberg, Louis Terman, David Wechsler).
Kinds of Intelligence:

- Is there one underlying capacity for intelligence or are there different ways to be intelligent?

- **Charles Spearman** – he believed (using factor analysis) that one important factor (g) – general factor - underlies all intelligence.

- (s) – less important specialized abilities
Another Opinion: Louis Thurstone

- *Thurstone* – Disagreed with Spearman and came up with seven distinct factors he called – **Primary Mental Abilities**.

- Inductive reasoning, word fluency, perceptual speed, verbal comprehension, spatial visualization, numerical ability, and associate memory.
1. There are many factors that determine intelligence, but genetics in the most important one

2. The internal validity of an intelligence test in \( g \)

3. A general intelligence that underlies success on a wide variety of tasks is \( g \)

4. Giftedness is determined by both innate ability to perform and experiences one has in life

5. The \( g \) is measured by the speed with which one can process information
Believed that Spearman’s (g) should be divided into two factors of intelligence

- **Fluid intelligence** – cognitive abilities requiring speed or rapid learning (diminish with age)
- **Crystallized intelligence** – learned knowledge and skills that tend to increase with age. (vocab)
MULTIPLE INTELLIGENCES: HOWARD GARDNER

- **Howard Gardner** – critic of (g) approach
- Believed that there are 8 approaches to intelligence
- 3 are measured on traditional intelligence tests
- 5 are not usually measured on any standardized test
- He believed that these abilities represent ways that people process information differently in the world.
Multiple Intelligences

- **Verbal/Linguistic**
  - The capacity to use language to express what's on your mind and to understand other people.

- **Logical/Mathematical**
  - The ability to understand the underlying principles of some kind of causal system.

- **Visual/Spatial**
  - The ability to present the spatial world internally in your mind.

- **Bodily/Kinesthetic**
  - The capacity to use your whole or parts of your body, to solve a problem, make something, or put on a production.

- **Naturalist**
  - The ability to discriminate among living things as well as sensitivity to other features of the natural world.

- **Interpersonal**
  - The ability to understand other people.

- **Existential**
  - To exhibit the proclivity to pose and ponder questions about life, death and ultimate realities.

- **Musical/Rhythmic**
  - The capacity to think in music, to be able to hear patterns, recognize them, and perhaps manipulate them.

*Multiple Intelligences* 1999

M.I. Smart
Robert Sternberg: Triarchic Theory of Intelligence

1. **Analytic** – What is tested by traditional intelligence tests
2. **Creative** – How we adapt to tough situations (threshold theory – we need a certain amount of intelligence to be creative but it is not the only thing we need)
3. **Practical** – “streetsmarts” – ability to read and perceive people, figure out directions, etc.
EMOTIONAL INTELLIGENCE:

- **Peter Salovey and John Mayer** – combines Gardner’s inter and intrapersonal approaches
- Uses **MEIS** – Multifactor Emotional Intelligence Scale
- Tests the person’s ability to perceive, understand, and regulate emotions
According to Sternberg, which of the following types of intelligence in his triarchic theory are measured by standard IQ tests?

1. Analytic
2. Practical
3. Predictive
4. Creative
5. Concurrent
6. 1, 2, and 5
7. 1, 2, and 4
FOOD FOR THOUGHT:

- Is it better to have a high IQ or high EQ?
- Does it depend on what you will be doing with your life?
CONCLUDING THOUGHT:

- Psychologists, educators, and Psychometricians agree that intelligence tests measure the ability to take tests well.
- They do not agree that all intelligence tests actually measure intelligence.
- Do you agree with them?
By the end of this lesson, I will be able to:

1. Discuss how environment and culture influences the definition of intelligence
To what extent is intelligence innate / learned?

Can intelligence be improved with practice or education?

The Head Start Program – has been shown to decrease children’s placement in special education classes.

Critics argue that these gains are short lived and that the student’s real deficiencies appear over time.
STUDIES OF CHILDREN AND TWINS: THE RESULTS

- Identical twins have much similar scores on IQ tests (even if raised apart)
- Intelligence scores of adopted children are more like those of their biological parents
- So, genetics seem to play a large role in IQ
ENVIRONMENTAL INFLUENCES:

- IQ scores of children that move from deprived environments to middle/upper class environments improved.
- School attendance seems to result in increased IQ scores.
- Intelligence tests have gotten more difficult and people are scoring the same (James Flynn).
- Could be attributed to: better nutrition, health care, advances in technology, better parenting, etc.
Racial differences in IQ scores show that African Americans, Native Americans, and Hispanics all score 10-15 pts. lower than white children.

Reasons: socio-economic status, possible test bias

Stereotype threat – anxiety that influences members of a group concerned about their performance on a test will confirm a negative stereotype.
When Studying Groups:

- **Within-group differences** – range of scores for variables being measured for a group of individuals (Hispanics)
- **Between-group differences** – usually the difference between means of two groups of individuals for a common variable. (Hispanics and African Americans)
WE’RE DONE!!

- Take Home Test for this Chapter due Tuesday
- Next Unit of Study → Abnormal Psychology (Chapter 14)
- Take Home Test for Ch 14 Due Friday the 13th!