

HINT

Within-group differences are typically larger than between-group differences.

highest scoring boy and the lowest scoring boy than between the average boy and the average girl. Furthermore, knowing that boys generally outperform girls on this test tells us nothing about the performance of any particular girl compared with the performance of any particular boy. Therefore, we need to be careful about how we use information about differences between groups. Essentially, we should not use it. We should ignore it and evaluate each person, regardless of group membership, as an individual.

A CAUTIONARY NOTE

It is often said that we live in a testing society. We like to be able to measure things and assign them a number. Therefore, keeping in mind the limitations and extraordinary labeling power of these instruments is particularly important. As we have discussed, the definition of intelligence (and many other concepts) remains hotly debated and many factors affect people's performances on tests. Thus, we need to take care not to ascribe too great a meaning to a test score. Many schools that used to measure all their students' IQs periodically have abandoned that practice. Schools that used to base admission to programs for exceptional children solely on these tests now frequently gather information in other ways as well. When IQ tests are given, the results remain confidential so as not to create expectations about how people *ought* to perform (see the information on self-fulfilling prophecy in Chapter 14). While well-designed tests can be extremely useful, we must recognize their limitations.

Practice Questions

Directions: Each of the questions or incomplete statements below is followed by five suggested answers or completions. Select the one that is best in each case.

1. Paul takes a test in the army to see if he would make a good pilot. Such a test is
 - (A) a standardized test.
 - (B) an aptitude test.
 - (C) an intelligence test.
 - (D) an achievement test.
 - (E) a biased test.

2. If a test is reliable, it means that
 - (A) it is given in the same way every time.
 - (B) it tests what it is supposed to test.
 - (C) it is a fair assessment.
 - (D) it yields consistent results.
 - (E) it is also valid.

3. The standardization sample is
- (A) the group of people who take the test.
 - (B) a random sample of the test takers used to evaluate the performance of others.
 - (C) the people used to represent the population for whom the test was intended.
 - (D) all the people who might ever take the test.
 - (E) the top 15 percent of scores on the test.
4. Which of the following is not one of Howard Gardner's multiple intelligences?
- (A) practical
 - (B) musical
 - (C) interpersonal
 - (D) spatial
 - (E) linguistic
5. Mrs. Cho is careful to make sure that she fairly represents the whole year's work on the final exam for her American literature class. If Mrs. Cho achieves this goal, her test will have
- (A) test-retest reliability.
 - (B) construct validity.
 - (C) content validity.
 - (D) split-half reliability.
 - (E) criterion validity.
6. Astor scores at the 84th percentile on the WISC. Which number most closely expresses his IQ?
- (A) 85
 - (B) 110
 - (C) 115
 - (D) 120
 - (E) 130
7. Spearman argued that intelligence could be boiled down to one ability known as
- (A) *s*.
 - (B) *i*.
 - (C) *g*.
 - (D) *a*.
 - (E) *x*.

8. Which of the following would provide the strongest evidence for the idea that intelligence is highly heritable?
- (A) The IQ scores of parents are positively correlated with the scores of their children.
 - (B) Monozygotic twins separated at birth have extremely similar IQ scores.
 - (C) Dizygotic twins score more similarly on IQ tests than do other siblings.
 - (D) Adopted children's IQ scores are positively correlated with their adopted parents' scores.
 - (E) Different ethnic groups have different average IQ scores.
9. All of the following people are known for their theories of what intelligence is except for
- (A) Thurstone.
 - (B) Gardner.
 - (C) Sternberg.
 - (D) Flynn.
 - (E) Guilford.
10. Which statement is true of power tests?
- (A) They are administered in a short amount of time.
 - (B) They are an example of an individual test.
 - (C) They are a pure measure of achievement.
 - (D) They consist of items of varying difficulty levels.
 - (E) They yield IQ scores.
11. People with high EQs would be likely to
- (A) pursue high-paying occupations.
 - (B) complete college.
 - (C) find jobs well suited to their individual strengths.
 - (D) be creative problem solvers.
 - (E) have a lot of close friends.
12. Although her score on the personality test indicated that Mary was devoid of social grace, painfully shy, and frightened of other people, she is extremely popular and outgoing. This personality test lacks
- (A) reliability.
 - (B) standardization.
 - (C) consistency.
 - (D) validity.
 - (E) practical worth.

13. Santos is 8 years old and, according to the Stanford-Binet, he has a mental age of 10. What is his IQ?
- (A) 80
 - (B) 100
 - (C) 120
 - (D) 125
 - (E) 150
14. The Flynn effect is the finding that
- (A) intelligence seems to increase with every generation.
 - (B) television has decreased intellectual performance.
 - (C) linguistic skills decline with age.
 - (D) within-group differences are larger than between-group differences.
 - (E) the more times people take a test, the better they tend to score.
15. Desmond believes that nature is far more important in shaping personality than nurture. Desmond probably believes in the strong influence of
- (A) environment.
 - (B) learning.
 - (C) reinforcement.
 - (D) genetics.
 - (E) culture.

ANSWERS TO PRACTICE QUESTIONS

1. **(B)** Aptitude tests aim to measure someone's ability or potential. In this case, the test is supposed to show whether Paul has the ability to be a pilot. The test may or may not be standardized or biased. The test is not attempting to measure Paul's intelligence. Since Paul has not yet been trained as a pilot, the test is not an achievement test.
2. **(D)** If a test is reliable, it yields consistent results. Standardized tests are generally given in the same way every time. A test is valid if it measures what it is supposed to measure. While valid tests are reliable, reliable tests are not necessarily valid. Whether a test is fair or biased can be evaluated in several ways as explained in the chapter. However, the fairness of a test is not synonymous with its consistency.
3. **(C)** The standardization sample represents the population for whom the test was intended and is used to construct the test. None of the other choices are referred to by any specific terminology.
4. **(A)** Practical intelligence is part of Sternberg's triarchic theory of intelligence. Gardner's multiple intelligences include linguistic, spatial, logical-mathematical, musical, bodily-kinesthetic, interpersonal, intrapersonal, and naturalist.

5. (C) Mrs. Cho is concerned about the content validity of her test. A test that fairly represents all the material taught in her class has content validity. Validity, in general, measures how well a test measures what it is supposed to measure. In order for Mrs. Cho's test to have construct validity, we would need to know that the test was successful in differentiating between varying levels of achievement in Mrs. Cho's class. If the test has criterion validity, we would have to know that the test successfully identified either those students who had excelled in their study of American literature (concurrent validity) or those students who would excel in the future (predictive validity). Reliability is a measure of how consistent the scores are on a test. Test-retest reliability involves giving the same test to the same population on at least two different occasions and measuring the correlation between the sets of scores. Split-half reliability is when one test is divided into two parts and the correlation between people's scores on the two halves is measured.
6. (C) The WISC (Wechsler intelligence scale for children) yields a deviation IQ score. The mean on the WISC is set at 100. Therefore, someone who scores 100 has scored at the 50th percentile on the test. The standard deviation on the WISC is set at 15. Since approximately 34 percent of the scores in a normal distribution fall between the mean and one standard deviation above the mean, Astor's score at the 84th percentile indicates that he scored almost exactly one standard deviation above the mean. Therefore, to compute Astor's score, we simply have to add the mean (100) to one standard deviation (15).
7. (C) Spearman argued that intelligence could be boiled down to one ability known as *g*. The *g* stands for general intelligence. Spearman also discussed *s*, which stands for specific intelligences. The other letters are all simply distractors.
8. (B) The strongest evidence presented for intelligence to be highly heritable is that monozygotic twins separated at birth have extremely similar IQ scores. Monozygotic twins share 100 percent of their genetic material. If they are separated at birth and therefore raised in different environments, similarity in their IQ scores argues for the influence of nature or heritability. Parents' IQ scores do tend to correlate positively with those of their children, but this similarity could be explained by either genetic or environmental factors. Dizygotic twins and other siblings share the same amount of genetic material on average (50 percent). Therefore, if the former score more similarly on IQ tests, an environmental influence is suggested. For instance, dizygotic twins may be treated more similarly than other siblings and grow up during the same time period. Since adopted children do not share any genetic material with the parents who adopted them, similarities must be due to environmental factors. Differences in average IQ scores between ethnic groups could be explained by either genetic or environmental factors.
9. (D) While Thurstone, Gardner, Sternberg, and Guilford all tried to define intelligence, Flynn is known for his observation that intelligence, as measured by IQ tests, is increasing.

10. (D) Power tests consist of items of varying levels of difficulty because their purpose is to identify the upper limit of a person's ability. Speed tests are given in a small amount of time since they seek to test how quickly someone can solve problems. Power tests could be given individually or in a group. Having a pure measure of achievement is impossible. IQ tests yield IQ scores.
11. (C) People with high EQs would be likely to find jobs well suited to their individual strengths. Emotional intelligence is thought to help people achieve what they want to achieve. Someone who has a high EQ will not necessarily want a high-paying job, go to college, be a creative problem solver, or have many close friends.
12. (D) Since the personality test does not seem to have resulted in an accurate depiction of Mary's personality, the test lacks validity. If repeated administrations of the test yielded similar results, the test could still be reliable. The test may or may not have been standardized. Consistency is generally equated with reliability. If the test lacks validity, it will not have practical worth, but the latter is not a psychological term.
13. (D) Scores on the Stanford-Binet IQ test are computed by dividing mental age by chronological age and multiplying by 100. Since 10 divided by 8 equals 1.25, Santos has an IQ of 125.
14. (A) The Flynn effect is the finding that intelligence seems to be increasing with every generation. Although television is often cast as a great social evil that rots the minds of our nation's youth, one hypothesized contribution to the Flynn effect is the exposure to the complex and rapid visual stimuli that appear on television. Linguistic skills do not decline with age. The statements in choices D and E are true but are not known as the Flynn effect.
15. (D) Adherents to a nature perspective often emphasize the effect of genetic makeup in shaping personality. All the other factors (environment, learning, reinforcement, and culture) are associated with a nurture perspective.