## 6. Assessment

1. Measurement: general process of determining the dimensions of an attribute or trait.
Assessment: Processes and procedures for collecting information about human behavior.
Assessment tools: include tests and inventories, rating scales, observation, interview data and other techniques.
Appraisal: appraisal implies going beyond measurement to making judgments about human attributes and behaviors and is used interchangeably with evaluation.
Interpretation: making a statement about the meaning or usefulness of measurement data according to the professional counselor's knowledge and judgment.

## 2. Measures of central tendency

A distribution of scores (measurements on a number of individuals) can be examined using the following measures:
Mean: the arithmetic average symbolized by $X$ or $M$ Median: the middle score in a distribution of scores Mode: the most frequent score in a distribution of scores All three of these fall in the same place (are identical) when the distribution of scores is symmetrical, i.e., normally distributed (not skewed).
(NOTE: You do not need to know any formulas for the CPCE. Furthermore, you do not need to make any calculations and are not allowed to use a calculator)

## 3. Skew

This is the degree to which a distribution of scores is not normally distributed.

## Positive skew



## Negative skew



The relationships between mean, median and mode are indicated above for skewed distributions. The mode is the top of the curve (most frequent score) and median is the middle score. The mean is pulled in the direction of the extreme scores represented by the tail of a skewed distribution.

## 4. Measures of variability

Range: this is the highest score minus the lowest score. Some researchers talk of inclusive range which is the high score minus the low score and adding one (1).

For example: Ten individuals' ages are: 24, 26, 26, 27, 28, 29, 31, 32, 39, 47
$47-24=23+1=24$ This range is inclusive; everyone is included.

- Standard deviation: this value describes the variability within a distribution of scores. We use the symbol SD to signify the standard deviation of a sample. When we talk about the population's variability, we use the symbol a (sigma). Standard deviation is essentially the mean of all the deviations from the mean. It is an excellent measure of the dispersion of scores.
- Variance: this is simply the square of the standard deviation, i.e., SD2. The variance does not describe the dispersion of scores as well as the standard deviation. However, we will see it again in the next section when we talk about analysis of variance.

5. Normal curve (bell-shaped curve): The normal curve essentially distributes the scores (individuals) into six equal parts-three above the mean and three below the mean. Counselors should be familiar with the distribution of scores within the normal curve: $34 \%$ and $34 \%=68 \%$ and comprises one standard deviation and $13.5 \%$ and $13.5 \%=95 \%$ and comprises two standard deviations and $2 \%$ and $2 \%=99 \%$ and comprises three standard deviations.

## 6. Percentile and stanine

Percentile is a value below which a specified percentage of cases fall.
For example: $75 \%$. This score is higher than $74 \%$ of the scores; $25 \%$ of the score are higher than this score.

Stanine, from standard nine, converts a distribution of scores into nine parts (1 to 9) with five in the middle and a standard deviation of about 2.

## 7. Standardized scores

A standardized score scale is like a 'common language' that we can use to compare several different test scores for the sam

For example: A person has a raw score of 60 on a vocabulary test and a raw score of 45 on an arithmetic test. Which performance is better? We cannot tell. Direct comparisons are not possible. After we convert both scores to a standardized scale, we can realistically compare them.

Standardized scores occur by converting raw score distributions. These derived scores provide for constant normative or relative meaning allowing for comparisons between individuals. Specifically, standardized scores express the person's distance from the mean in terms of the standard deviation of that standard score distribution. Standardized scores are continuous and have equality of units. The two most commonly used standardized scores are:
z-score: The mean is 0 ; the standard deviation is 1.0 - (See normal curve figure.) The range for the standard deviation is -3.0 to 3.0.


| C | 200 | 300 | 400 | 500 | 600 | 700 | 800 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| E |  |  |  |  |  |  |  |
| B |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| s |  |  |  |  |  |  |  |
| c |  |  |  |  |  |  |  |
| r |  |  |  |  |  |  |  |
| e |  |  |  |  |  |  |  |

Deviation IQ $<S D=15$ )

| 55 | 70 | 85 | 100 | 115 | 130 | 145 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |

Stanine
1
4\%

Percentile | | | | | | | | | | | | | $\begin{aligned} & \text { | } \\ & 99\end{aligned}$

Figure Relationships among different types of test scores in a normal distribution

The $\mathbf{z}$ in $\mathbf{z}$-score should remind you of zero which is the mean of this distribution.

T score: The mean of this standardized score scale is 50 and the standard deviation is 10. By (T)ransforming this standard score, negative scores are eliminated unlike the zscore.(See normal curve figure.) The T should remind you of ten which is the standard deviation unit of this distribution.

## 8. Correlation coefficient

The Pearson Product-Moment Correlation Coefficient ( $r$ ) is frequently used.
A correlation coefficient ranges from -1.00 (a perfect negative correlation) to 1.00 (a perfect positive correlation). This is a statistical index which shows the relationship between two sets of numbers. When a very strong correlation exists, if you know one score of an individual you can predict (to a large degree) the other score of that person. A correlation between two variables is called bivariate; between three or more variables, it is called multivariate. The correlation coefficient tells you nothing about cause and effect, only the degree of relationship.

## 9. Reliability

Reliability is the consistency of a test or measure the degree to which the test can be expected to provide similar results for the same subjects on repeated administrations. Reliability can be viewed as the extent to which a measure is free from error. If the instrument has little error, it is reliable. A correlation coefficient is used to determine reliability. If the reliability coefficient is high, about .70 or higher, test scores have little error and the instrument is said to be reliable.
Reliability is a necessary psychometric property of tests and measures.

## 10. Types of reliability

- Stability: this is test-retest reliability obtained using the same instrument on both occasions -same group tested twice. The results of the two administrations are correlated. The length of time and intervening experiences may influence stability reliability. Two weeks is a good time between test administrations .
- Equivalence: alternate forms of the same test are administered to the same group and the correlation between them is calculated. How comparable the forms of the tests are will influence this reliability. Intervening events and experiences may also influence reliability.
- Internal consistency: In this split-half method, the test is divided into two halves. The correlation between these two halves is calculated. Because you reduce the length of the test (one-half versus one-half) you necessarily reduce its measured reliability. Consequently, you may apply the Spearman-Brown formula (sometimes called 'prophecy' formula) to see how reliable the test would be had you not split it in two.
- Internal consistency may also be determined by measuring inter-item consistency. The more homogeneous the items, the more reliable the test. Kuder-Richardson formulas (there are two) are used if the test contains
dichotomous items (such as true-false, yes-no). If the instrument contains nondichotomous items (such as multiple choice, essay), Cronbach alpha coefficient is applied.


## 11. True and error variance

Tests measure "true" and "error" variance. You want to measure true variance, the actual psychological trait or characteristic that the test is measuring.


For example: Two tests are administered. Each one measures true variance (T1 and T2) and error variance (E1 and E2).
If the correlation between two tests or two forms of the same test is, for example:.90, then the amount of true variance measured in common is the correlation squared ( $.902=81 \%$ ).

Coefficient of Determination is the degree of common variance. It is the index ( $81 \%$ ) that results from squaring the correlation (.90). Coefficient of Nondetermination is the unique variance, not common. For the above example, it would be $19 \%$ and represents the error variance.

## 12. Standard error of measurement

The standard error of measurement (SEM) is another measure of reliability and useful in interpreting the test scores of an individual. The SEM may also be referred to as Confidence Band or Confidence Limits. The standard error of measurement helps determine the range within which an individual's test score probably falls.

For example: A person scores a 92 on a test. The test's SEM $=5.0$. Chances are about 2 in 3 ( $67 \%$ ) that the person's score falls between 87 and 97. (Refer to the normal curve: $34 \%$ and $34 \%$ of the cases fall within one standard deviation, positive and negative, for a total of $68 \%$ )

For the same test with the same SEM of 5.0, you can say that $95 \%$ of the time the person's score would fall within the range of 82 and 102. Every test has its own unique
value of SEM which is calculated in advance and may be reported on the test's score profile.

## 13. Validity

Validity is the degree to which a test measures what it purports to measure for the specific purpose for which it is used. In other words, validity is situation specific depending on the purpose and population. An instrument may be valid for some purposes and not others.

## 14. Types of validity

Face: the instrument looks valid
For example: A math test has math items. This 'validity' could be important from the test-taker's perspective.

Content: the instrument contains items drawn from the domain of items which could be included.

For example: Two professors of Psychology 101, devise a final exam which covers the important content that they both teach.

Predictive: the predictions made by the test are confirmed by later behavior (criterion).
For example: The scores on the Graduate Record Exam predict later grade point average.

Concurrent: the results of the test are compared with other tests' results or behaviors (criteria) at or about the same time.

For example: Scores of an art aptitude test may be compared to grades already assigned to students in an art class.

## Construct:

- A test has construct validity to the extent it measures some hypothetical construct such as anxiety, creativity, etc. Usually several tests or instruments are used to measure different components of the construct or of the hypothesized relationships between that construct and other constructs.
- Convergent validation occurs when there is high correlation between the construct under investigation and others.
- Discriminant validation occurs when there is no significant correlation between the construct under investigation and others.
- The construct validation process is best when multiple traits are being measured using a variety of methods.


## 15. Tests may be reliable but not valid.

Valid tests are reliable unless of course there is a change in the underlying trait or characteristic which might occur through maturation, training or development.

## 16. Tests may be:

- Power based: no time limits or very generous ones (such as the NCE and CPCE).
- Speed based: timed, and the emphasis is placed on speed and accuracy. Examples are measures of intelligence, ability and aptitude.


## 17. Assessment may be:

- Norm referenced: comparing individuals to others who have taken the test before. Norms may be national, state or local. In norm-referenced testing, how you compare with others is more important than what you know.
- Criterion referenced: comparing an individual's performance to some predetermined criterion which has been established as important. The National Counselor Exam's cut-off score is an example. For the CPCE, university programs are allowed to determine the criterion (cut-off score). Criterion referenced is sometimes called domain referenced.
- Ipsatively interpreted: comparing the results on the test within the individual. For example, looking at an individual's highs and lows on an aptitude battery which measures several aptitudes. There is no comparison with others. Another example of ipsative is when an individual's score on a second test is compared to the score on the first test.
- A maximal performance test may generate a person's best performance on an aptitude or achievement test and a typical performance may occur on an interest or personality test.


## 18. Purposes/rationale for using tests

- Help the counselor decide if the client's needs are within the range of his or her services.
- help the client gain self-understanding.
- help the counselor gain a better understanding of the client.
- assist the counselor in determining which counseling methods, approaches or techniques will be suitable.
- assist the counselee to predict future performance in education, training or work.
- help counselees make decisions about their educational or work futures.
- help identify interests not previously known.
- help evaluate the outcomes of counseling.

19. Circumstances under which testing may be useful:
20. placement-in education or work settings
21. admissions-such as undergraduate, graduate or professional schools
22. diagnosis
23. counseling
24. educational planning
25. evaluation
26. licensure and certification
27. self-understanding

## 20. Regression toward the mean

Statistical regression means that if one earns a very low score (at 15\% or lower) or very high score (at $85 \%$ or higher) on a pretest; the individual will probably earn a score closer to the mean on the posttest. This is because of the error occurring due to chance, personal and environmental factors. These factors can reliably be expected to be different on the posttest.

## 21 Standardized vs. nonstandardized assessment

- Standardized: the instruments are administered in a formal, structured procedure and the scoring is specified.
- Nonstandardized: there are no formal or routine instructions for administration or for scoring. Some examples may be checklists or rating scales.


## 22. Tests and inventories

Tests and inventories are typically mentioned on the NCE and CPCE in the context of an application-type exam question. Perhaps only two or three of these will be mentioned on the exam. Examples of tests in several areas are listed here and details about any test or inventory will not be the focus on the exam.

1. Intelligence is the ability to think in abstract terms; to learn. Some also believe it is the ability to adapt to the environment and adjust to it. It is also called general ability or cognitive ability.
A. Intelligence Tests

- Stanford-Binet Intelligence Scales
- Wechsler Adult Intelligence Scale (WAIS - IV)
- Wechsler Intelligence Scale for Children (WISC-IV)
- Cognitive Abilities Test
B. Specialized Ability Tests
- Kaufman Assessment Battery for Children - II
- System of Multicultural Pluralistic Assessment (SOMPA). It measures medical, social systems and pluralistic factors.
- ACT (American College Test Program)
- SAT Reasoning Test
- Miller Analogies Test (MAT) Graduate Record Exam (GRE)

2. Achievement: measures the effects of learning or a set of experiences. These tests may be used diagnostically. Many states have their own K-12 achievement tests. A national measure of academic performance is National Assessment of Educational Progress (NAEP). Other tests available include:

- California Achievement Tests
- lowa Tests of Basic Skills
- Stanford Achievement Test


## Specialized Achievement Tests:

- General Education Development (GED)
- College Board 's Advanced Placement Program
- College-Level Examination Program ( CLEP)

3. Aptitude: also called ability tests, these measure the effects of general learning and are used to predict future performance. Each of those listed here measures several abilities or aptitudes.

- Differential Aptitude Tests (DAT)
- O*Net Ability Profiler (formerly, General Aptitude Test Battery, GATB)
- Armed Services Vocational Aptitude Battery (ASVAB)
- Career Ability Placement Survey (CAPS)

4. Personality: is the dynamic product of genetic factors, environmental experiences, and learning to include traits and characteristics.
A. Projectives: (These tests present a relatively unstructured task or stimulus. The person projects thought processes, needs, anxieties, etc.)

- Rorschach
- Thematic Apperception Test (TAT)
- Rotter Incomplete Sentences Blank
- Draw-A-Person Test


## B. Inventories:

- Minnesota Multiphasic Personality Inventory
- California Psychological Inventory (CPI)
- NEO Personality Inventory - Revised
- Beck Depression Inventory
- Myers -Briggs Type Indicator
C. Specialized:
- Tennessee Self Concept Scale
- Bender Visual-Motor Gestalt Test

5. Interests: preferences, likes and dislikes of an individual and more broadly includes values. Interests are often not stable in the teen years.

- Strong Interest Inventory
- Self-Directed Search
- Career Assessment Inventory
- Campbell Interests and Skills Survey
- O*Net Interest Profiler


## 23. Semantic differential

This scale asks respondents to report where they are on a dichotomous range between two affective polar opposites. For example: "Think about the value of this Comps Preparation Tool."
Very Good $\qquad$ Very Bad
Responses can be codified and added to those of others. The adjective pairs selected can usually be classified as having an evaluative, potency, and activity underlying structure thus providing for a second level of analysis.

## 24. Intrusive and unobtrusive measurement

- Intrusive (or reactive) measurement means the participant knows he or she is being watched or questioned and this knowledge may affect his or her performance. Examples are questionnaires, interviews or observation.
- Unobtrusive (or nonreactive) measurement means data is collected without the awareness of the individual, or without changing the natural course of events. Examples are reviewing existing records or unobtrusive observation.


## 25. Observation as appraisal technique

With this technique, you observe samples from a stream of behavior. In observation, you may use schedules, coding systems, and record forms.

## 26. Case/historical study; rating scales; interview

- Case or historical study: This may be an analytical and/or diagnostic investigation of a person or group.
- Rating scales: these may be used to report the degree to which an attribute or characteristic is present.


## 27. Sociometry

- Sociometry can be used to identify isolates, rejectees or stars (popular individuals). You can measure the structure and organization of social groups which could be a classroom of fourth graders who have been together for a few months, or a work unit. It requires revealing personal feelings about others.
- Sociogram: a figure or map showing the interrelationships or structure of the group.


## 28. Social desirability

This is the tendency for test takers to respond in ways they perceive to be socially desirable.

## 29. Using and interpreting test scores

- You need training in test theory and background information about the tests you use.
- You must study the test's technical manual.
- Prepare for the test interpretation. Understand the scores, profiles, and implications of the results before you counsel the individual.
- Describe the test to the person in nontechnical terms and explain what was being measured.
- Describe the nature of the scores you are reporting. Explain percentiles, stanines or any other technical terms.
- Organize the data so it makes the most sense to the client. Show profiles, charts and comparative data if appropriate. Consider and explain interrelationships between scores and between tests if more than one was used.
- Provide an interpretation to the client and ask for reactions and feelings. Help the client integrate the test results with existing information.
- Remind the client that test scores are additional data for them to consider and are not infallible. Test data may be useful in decision in making or obtaining some objective.
- Go slowly you may have used similar words in test interpretations hundreds of times. It may be the first time the client is hearing these words.


## 30 Grade and age equivalent scores

In school settings, scores on an achievement test are often reported as grade equivalent scores. If a student correctly completes the number of items on a test that the average sixth grader completes, that student has a grade equivalent score of 6 . Age equivalent scores work in a similar manner. An individual's score is compared to the average score of others at the same age. For example, if a 7.6 year old student earned a score equivalent to 8.0 year old students that would be his or her age equivalent score.

## 31. Percentile ranks

An individual's score can be compared to a group (norm group) already examined. The individual's percentile rank indicates what percentage of individuals in that group has scores above or below this individual. A percentile rank of 35 means that this individual's score is higher than 34 percent of the individuals in the norm group. On the other hand, $65 \%$ of the individuals in that norm group have scores higher than this individual.

## 32. Computer-based assessment

Much academic as well as professional testing is now done on computer. Some of the advantages and disadvantages include the following:

## Advantages:

- standardizes administration and scoring
- feedback and results may be available immediately
- assuming computers are available, costs will be reduced
- profiles of results and reports can be generated


## Disadvantages:

- not all assessments are available on computer
- testing by computer may be a scary proposition for some test takers
- if not available, computer equipment is expensive to purchase
- personal contact with a test administrator or proctor may not be available


## 33. Ethical issues in testing

- Tests may be biased against non-whites, females, and those of other cultures. Many tests were originally developed and normed on white middle-class males.
- Counselors must be trained and competent to select and administer tests and to interpret test results and information.
- Test results should be released only to competent professionals and with the consent of the test taker.
- Tests may be used to label and stereotype; they may invade privacy.
- Confidentiality of test results may be an issue especially with computerization.
- Computerized testing (on-line) may raise issues of validity. Is the test the same on the computer as it is on paper?
- If a test is given, its results should be interpreted. Many say the test results belong to the client.
- Review the measurement and evaluation section of the ethical standard periodically.


## 34. Assessment resources

- The Mental Measurements Yearbook which comes from the Buros Institute, contains critical reviews of tests and it lists published references of tests. The nineteenth edition was published in 2014.
- Tests in Print VIII (2011), has information on approximately 3,000 testing instruments.
- A Counselor's Guide to Career Assessment Instruments (6 ${ }^{\text {th }}$ Edition) was edited by Chris Wood and Danica G. Hays (2013), and is published by the National Career Development Association.


## 35. Association for Assessment and Research

The Association for Assessment and Research is one of the 20 divisions of the American Counseling Association.

## Assessment Study Questions

1. An employment counselor uses the General Aptitude Test Battery which measures Nine aptitudes important for different jobs. His clients are typically uneducated. Because the counselor is concerned how his clients' scores stack up against others and how they would perceive their aptitudes, he interprets the scores to each of his clients
A. ipsatively.
B. using national norms.
C. using local norms.
D. in relationship to other tests.
2. Historically, there were many influences on the development of the testing movement. Which one of the following was NOT an early influence?
A. Civil rights movement.
B. Compulsory school attendance.
C. World War I.
D. Advances in statistical methodology.
3. To establish its validity, a new test was administered to the same group which took a well-established test. The correlation coefficient between the two tests was . 70 .
The amount of variance common to both instruments was
A. 61 to 70 percent.
B. 50 to 60 percent.
C. less than 50 percent.
D. more than 70 percent.
4. Shortly after Christmas, a third grade teacher realized that several students in her class had not established any friendships with other children although they did classroom activities together when she put them in groups. The teacher could clearly identify some leaders, some loners and others. She decided to formally examine the social structure in the classroom and chose $\qquad$ as her method for doing this.
A. the interview
B. case study
C. personality testing
D. sociometry
5. Jenna is beginning her junior year in high school and is college bound. Her ability for success in college is excellent but she is still uncertain as to what her college major should be. Given this information about Jenna, the counselor would recommend
A. ability testing
B. aptitude assessment.
C. interest assessment.
D. intelligence testing.
6. A $\qquad$ occurs when all the members of some set of individuals have an equal chance of being selected.
A. population
B. collection
C. sample of convenience
D. random sample
7. A group of individuals completed two assessment devices. The correlation between the two sets of scores was .81. An accurate interpretation of that correlation coefficient is
A. the higher the individuals scored on one test, the lower they scored on the other.
B. the individuals scored high on both tests, or they scored low on both tests.
C. scoring high on one test caused the high score on the other test.
D. an accurate interpretation cannot be made.
8. A personnel director of a medium-sized manufacturing firm was instructed to increase the percentage of hires who could successfully perform the delicate and complex jobs requiring finger dexterity. The personnel director was reluctant to add a test because the hiring procedures were already extensive. In reviewing possible tests to use in employee selection, the personnel director was most concerned about
A. standard error of measurement.
B. stability reliability.
C. face validity.
D. the norm group.
9. David a gifted and talented fourth grader, completed an achievement test. His score on the math portion was two standard deviations above the mean. David's score was higher than $\qquad$ of others taking the same test.
A. 64 percent
B. 97 percent
C. 95 percent
D. 68 percent
10. A counselor educator gave the same final exam to two classes studying Human Growth and Development. In Class A, the final exam scores were normally distributed and in Class B there was a definite negative skew in exam results. The counselor educator would be correct in making the following statement.
A. More students mastered the material in Class A than Class B.
B. Over $50 \%$ of students in Class A scored higher than the mean.
C. The mean, median, and mode are the same for Class A students
D. A different final exam should be used in each class.
11. Given: Test Z Mean $=100$ SD $=12$ SE Meas. $=4$

A student obtained a score of 112. An equivalent form of Test $Z$ was administered to him. We could expect his score on the second test to fall into which range of scores about two-thirds of the time?
a. 96-104.
B. 108-116.
C. 104-120.
D. 88-112.
12. Assume you have a normal distribution of raw scores. Which of the following standard scores is farthest from the mean?
A. T-score $=35$.
B. Stanine $=6$.
C. $z=-2.0$.
D. Deviation $I Q=120$.
13. A family counselor decides to visit the family home and observe the behavior of one of the children who appears seriously disturbed. In setting up the visitation schedule, the counselor is concerned about all of the following EXCEPT
A. time sampling.
B. observing in approximately uniform periods each day.
C. making sure that the family members who are at home do not change on every
visit.
D. making sure all family members are present every time.
14. A high school principal asks the lead teacher in the English Department to design an exam to assess English usage and writing skills to be administered to tenth graders. After the teacher designs the exam, she gives copies to all the other English teachers and asks them to review it and judge its appropriateness, clarity, accuracy, and fairness. By going through this process, she is determining $\qquad$ validity.
A. predictive
B. concurrent
C. discriminant
D. content
15. A counselor at a mental health agency noticed that the clients who complained of anxiety also tended to introversion. As part of the agency's intake process, test scores were collected from all clients on these two, as well as other, variables. The counselor decided to examine the relationship between these two variables and her statistic of choice was a
A. $t$-test.
B. canonical correlation.
C. multivariate correlation coefficient.
D. bivariate correlation
16. Each year the Director of Counseling at a small liberal arts college in a Midwestern state collects data on the Alcohol and Drug Awareness Scale (ADAS) from the freshmen class. This year the director found that the students' results on the scale reached the $70^{\text {th }}$ percentile of state norms. They were at the 40 lh percentile of national norms. He can accurately conclude
A. his students score higher than national students on the ADAS.
B. national students score lower on the ADAS than state students.
C. his students score higher than state students on the ADAS.
D. not enough data is available to make valid comparisons.
17. A counselor working in a juvenile group home designed an inventory of social skills which each resident takes. The purpose of this inventory is to substitute for other ways of estimating the social skills level of the residents. Since the purpose of this inventory is to substitute for other methods of measurement, the most appropriate type of validity we are concerned with here is
A. construct.
B. concurrent.
C. content.
D. face.
18. Li and May took the same test of academic ability. The mean of the test is 100 and the test's standard deviation is 10 . Li's score is equivalent to a $z$ score of -1.0 and May's score is equivalent to a T score of 45 . On the test
A. Li's score was 90 and May's was 95 .
B. Li's score was 99 and May's was 90 .
C. Li's score was 95 and May's was 140.
D. There is not enough data to make the calculation.
19. A standard technique often used by family counselors who are interested in determining intergenerational influences is
A. sculpting.
B. structural map.
C. determining coalitions.
D. genogram.
20. A group of 100 individuals completed two tests. Their scores on both tests are plotted on a chart. If the slope of the regression line (line of best fit) angles from top left to bottom right, the counselor reviewing this data could conclude there is a correlation between the two variables measured.
A. zero
B. positive
C. negative
D. parametric
21. Laura was applying to a doctoral program in counselor education. Along with her application, she submitted the following three test scores: GRE Verbal: 450; a tolerance for ambiguity test score at T-score 45; and a z-score of -.5 on a test of trait anxiety. The professor reviewing her file characterized all her scores at the following stanine
A. 3.2.
B. 4.0 .
C. 5.2.
D. 6.0
22. Convergent and discriminant are concepts associated with which of the following validities?
A. Predictive.
B. Content.
C. Concurrent.
D. Construct.
23. Which of the following names are NOT associated with the reliability of tests?
A. Kaufman.
B. Spearman-Brown.
C. Cronbach.
D. Kuder-Richardson.
24. A school counselor is assisting an eighth grader in determining which specialized high school she should attend next fall. The counselor might be LEAST concerned with
A. maximal performance tests.
B. sociometric assessments
C. grade equivalent scores
D. age equivalent scores
25. A newspaper opinion writer hopes to sway readers of her article about the high cost of homes in a community. Although she has several measures of central tendency of house values, she chooses to use the $\qquad$ of house values.
A. mode
B. median
C. mean
D. range

## Assessment Question Answers

1. A
2. $A$
3. $C$
4. D
5. B
6. D
7. B
8. B
9. $B$
10. C
11.B
11. C
13.D
14.D
12. D
13. C
17.B
14. A
19.D
15. C
21.B
16. D
17. A
24.B
18. C
