# Web Page for Chapter 17

### MULTIPLE CHOICE QUESTIONS - SET A

- 1 The degree to which a test measures what it purports to measure reflects its:
  - (a) reliability
  - (b) validity
  - (c) objectivity
  - (d) stability
- 2 If a sample of test items adequately represent the subject matter of the given apprenticeship course, the test is said to have:
  - (a) construct validity
  - (b) predictive validity
  - (c) content validity
  - (d) concurrent validity
- 3 A test which yields score that are found to be highly correlated with later performance has:
  - (a) construct validity
  - (b) predictive validity
  - (c) content validity
  - (d) concurrent validity
- 4 A test which yields scores that are highly correlated with performance now has:
  - (a) construct validity
  - (b) predictive validity
  - (c) content validity
  - (d) concurrent validity
- 5 The degree to which a test measures a given hypothetical construct is:
  - (a) structural validity
  - (b) predictive validity
  - (c) content validity
  - (d) none of these
- 6 Cronbach's alpha measures:
  - (a) stability
  - (b) split half reliability
  - (c) reliability of parallel items
  - (d) internal consistency
- 7 A measurement instrument can be valid but not reliable:
  - (a) always false
  - (b) always true
  - (c) sometimes true
  - (d) depends on the type of validity
- 8 The standard error of the measure provides an indication of:
  - (a) the accuracy of the correlation
  - (b) the range of error round an individual value
  - (c) the standard deviation of sampling error
  - (d) none of these

- 9 If the SD = 6 and r = .64, what is the 95% range of estimate round a score of 100?
  - (a) 103.6 96.64
  - (b) 107 93
  - (c) 101.96 98.04
  - (d) none of the above
- 10 If a test of 20 items has a reliability of .80, what is its reliability if it were three times as long?
  - (a) 2.4.
  - (b) 1.00
  - (c) .80
  - (d) .92
- 11 When computing the reliability from a narrow range of performance, would the reliability be higher, the same, or lower than that obtained if the performance range were larger.
  - (a) higher
  - (b) lower
  - (c) the same
  - (d) depends on N
- 12 Which two of the following would be taken as evidence of validity and not of reliability?
  - (a) assessments obtained on two successive evaluations correlated +.95
  - (b) scores on a selection test predicted the final training assessment grades
  - (c) scores on a test correlated highly with scores from another test measuring the same quality
  - (d) even numbered questions were correlated with odd numbered questions
- 13 External validity is concerned with:
  - (a) error from subjective assessments
  - (b) the effect of statistical regression
  - (c) selection bias
  - (d) generalizing the results of a piece of research to the population or other settings
- 14 Which of the following is a threat to internal validity:
  - (a) Hawthorne effect
  - (b) non-random sampling
  - (c) drop out
  - (d) failure to operationalize variables adequately

## SPSS ACTIVITY – CRONBACH ALPHA ANALYSIS

Access SPSS Chapter17 Data File B. It consists of 10 Likert 5 point scale items concerning attitudes towards recent changes in Industrial Relations laws responded to by 98 owners of small businesses. We wish to refine the scale to ensure it has high internal reliability. Carry out a Cronbach alpha analysis, interpret the results and then if necessary conduct a second analysis using a smaller set of items.

## MULTIPLE CHOICE QUESTIONS - SET B

- 1 The consistency of a measure over time is:
  - (a) temporal validity
  - (b) temporal instability
  - (c) validity
  - (d) reliability
- 2 The degree to which a measure actually measures what we think it measures is:
  - (a) generalization
  - (b) validity
  - (c) predictiveness

- (d) reliability
- 3 Face validity is:
  - (a) the degree to which two measures of the same variable produce the same results
  - (b) the degree to which respondents believe the measure is relevant and important
  - (c) the degree to which the measure generalizes
  - (d) the degree to which the items reflect the universe of items
- 4 To ensure a high degree of content validity
  - (a) include a wide range of items
  - (b) write the items clearly and precisely in non-jargon words
  - (c) have the content judged as relevant for the named assessment by external experts
  - (d) correlate the measure with another similar measure
- 5 Cronbach alpha measures:
  - (a) internal reliability
  - (b) external validity
  - (c) internal validity
  - (d) split half reliability
- 6 With an SD of 5 and reliability of .75, what is the approximate 95% range of error round a score of 50?
  - (a) 52.5 47.5
  - (b) 54.9 45.1
  - (c) 50.75 49.25
  - (d) 51.5 48.5

### ADDITIONAL QUESTIONS/PROBLEMS

- 1 If the correlation between the scores from two occasions of testing was +1 what does this imply for reliability and error?
- 2 If test-retest reliability is less than +1, what does this imply?
- 3 Given a test instrument that has a reliability of +0.84, and SD = 10, what is the 95% confidence range around a score of 100?
- 4 For each of the following indicate whether they would contribute to true score or error score variance when a test-retest method is employed:
  - (a) noise from the next room on the first occasion
  - (b) reading ability
  - (c) headache on both occasions
  - (d) broken pencil on second occasion
  - (e) knowledge of specific fact prior to first occasion
  - (f) knowledge of a specific fact prior to second occasion only
  - (g) a football match on the pitch outside the exam room during the retest
  - (h) equipment breakdown during an apprentices' first test
  - (i) a sneezing attack during the re-test
  - (j) practice on similar items between test and retest
  - (k) application of non-standardized instructions on every occasion
  - (l) different time allowed for first and second test

### CLASS DISCUSSION AND ACTIVITIES IN GROUPS OR PAIRS

1 Explain the meaning of Xobs = Xtrue ± Xerror. How does this expression help us to understand the concept of reliability?

- 2 Try to explain to a classmate why it is possible for an assessment to be reliable but not valid but if it not reliable it cannot be valid.
- 3 Explain the usefulness of reporting the standard error of the measurement as well as the reliability of the test.
- 4 Explain why maturation, and statistical regression can affect internal validity.
- 5 Explain why splitting a test onto a first half and a second half is not a good way of splitting a test for a split half reliability assessment.
- 6 Explain what problems would you see as problems stemming from (a) too short a time lapse, and (b) too long a time lapse with human subjects in a test-retest situation.

#### ANSWERS TO MULTIPLE CHOICE QUESTIONS – SET A

1 (b), 2 (c), 3 (b), 4 (d), 5 (d), 6 (d), 7 (a), 8 (b), 9 (b), 10 (d), 11 (b), 12 (b and c), 13 (d), 14 (c).

### ANSWERS TO QUESTIONS IN TEXT

Qu 17.6

Type of Reliability Coefficient				
Source of	Test-retest	Test-retest	Parallel form	Split half
error variance		immediate	delayed	
	37	37	37	N
The procedure/method	Yes	Yes	Yes	No
The person/event/equipment	No	Yes	No	No
The conditions	No	Yes	No	No

Qu 17.9 (b)

*Qu 17.11 (a)* 

### ANSWERS TO MULTIPLE CHOICE QUESTIONS – SET B

1 (d), 2 (b), 3 (b), 4 (c), 5 (a), 6 (b).

#### ANSWERS TO ADDITIONAL QUESTIONS

- 1 r = +1.0 would imply perfect replicability with no measurement error evident.
- 2 As the test-retest correlation declines from unity, we are measuring the effects of increasing random error.
- 3 approximately 92 108.
- 4 b, c and e contribute to true score variance while the rest contribute to error score variance.