

Chapter 6 Selected Homework Answers

1. Observed score = true score + error score.

Observed score variance = true score variance + error score variance.

Reliability = true score variance/observed score variance.

For this test: observed score variance = $20^2 = 400$, and error score variance = $10^2 = 100$.

Thus, true score variance = $400 - 100 = 300$.

Test reliability = $300/400 = .75$.

3. Note that the entire test length is now 120 items, so $k = 120/30 = 4$. Use the Spearman-Brown prophecy formula on page 87 or table 6-4 on page 89 to obtain .80.

5. Note that reliability is .91, which is equal to true score variance/observed score variance. Observed score variance is $20^2 = 400$. Since 91% (.91) of the observed score variance is true score variance, true score variance = $400 \times .91 = 364$. Error score variance = observed score variance – true score variance, so error score variance = $400 - 364 = 36$.

7. Recall that observed score variance is simply the standard deviation squared: $10^2 = 100$.

9. Note that reliability is .91, which is equal to true score variance/observed score variance. Observed score variance is $10^2 = 100$. Since 91% (.91) of the observed score variance is true score variance, true score variance = $100 \times .91 = 91$. Error score variance = observed score variance – true score variance, so error score variance = $100 - 91 = 9$.

11. Use table 6-4 and work “backward.” You have increased the size of the test by a multiple of 5 (i.e., $k = 5$). Upon increasing the size by five, you obtained an estimated reliability of .84. Look under $k = 5$ and find .84. Read to the left and see that the original reliability (r_{11}) had to have been .52.

Use the alpha coefficient on page 90. $k = 5$ trials.

13. Calculate the standard error of measurement (SEM) from formula on page 93. The SEM is 30. You can be 95% confident that Jamie’s true score is his observed score ± 2 SEM.

$500 \pm 2 \times 30 = 500 \pm 60$; so Jamie's true score should fall between 440 and 560.

15. Objectivity is a special case of reliability. Objectivity is *interrater* reliability.

17. Reliability and relevance.

19. You could continually use the Spearman-Brown prophecy formula on page 87 or table 6-4 on page 89. Using table 6-4, note that increasing from 10 items to 50 ($k = 50/10 = 5$) and starting with $r_{11} = .28$ only raises the estimated reliability to .66. Since table 6-4

does not have $r_{11} = .66$, you can use the actual formula on page 87 or estimate the next step from table 6-4. Using the estimation method with table 6-4 suggests that the 50 items must be increased to 100 items (i.e., $k = 2$) in order to obtain a reliability of approximately .80.

21. Note that reliability is true score variance/observed score variance, so reliability = $34/50 = .68$.

23. The index of reliability (page 92) is simply the square root of the reliability. Thus, the index of reliability is .90. This means that the theoretical correlation between observed scores and true scores on the test is .90 (quite good).

25. Use equation page 84 to confirm the first two answers. Lowest value for reliability is 0 (zero). Highest value for reliability is 1.0 The minimum desired reliability for a written test is .80.

27. Use the Spearman-Brown prophecy formula: $k = 150/50 = 3$. The predicted reliability is .84.

29. Use the Spearman-Brown prophecy formula. $k =$ number of items you *want* reliability for divided by number of items you *have* reliability for: $k = 60/20 = 3$. Again, use the Spearman-Brown prophecy table or formula. Answer = .84.

31. Consistency.

33. Bias.

35. Consider the definitional formula of reliability (see page 84). Observed score variance = 100; true score variance = 84; error score variance = 16 (this is 4^2 [i.e., the SEM squared is the variance error]). Note that reliability is $84/100$. Total (observed) variance = true score variance + error score variance (i.e., $100 = 84 + 16$).

37. Use the z score formula. Note that the mean is Mary's score and the standard deviation is the SEM. z score = $(42 - 46)/4 = -4/4 = -1.0$. The answer is $50.00 - 34.13 = 15.87\%$.

39. Consider the definitional formula of reliability (see page 84). Observed score variance = 900 (i.e., 30^2); true score variance = 792; error score variance = 108 (this is 10.39^2 [i.e., the SEM squared is the variance error]). Note that reliability is $88/100$.

41. A little over 2. Using the Spearman-Brown prophecy formula with $k = 2$ results in .94, and with $k = 3$ results in .96.

43. Use the Spearman-Brown prophecy formula "backward." She needs 60 items. She has 20. So she must write 40 items.

45. Your best guess of Susan's *true* score is 940 because she has only taken the test once. That one score is all the information you have.

47. Consider the definitional formula of reliability on page 84. Reliability is .84 (84% of observed or total variance); observed (total) variance = 400 (i.e., 20^2); true score variance is 336 (i.e., 84% of 400); error score variance is 64 (i.e., the SEM squared or 8^2).

49. Use the z score formula with John's score as the mean and the SEM as the s . Then use the z table. z score = $(128 - 120)/8 = 8/8 = 1.0$. The answer is $50.00 - 34.13 = 15.87\%$.

51. Concurrent or predictive validity, because each of these has a criterion.

53. Use the equation on page 90 to find the alpha coefficient. The answer is .84. To estimate the reliability use the Spearman-Brown prophecy formula with $k = 15/5 = 3$. The answer is .94. You would expect this because longer tests are more reliable than shorter tests.

55. Use the Spearman-Brown prophecy formula with $k = 1/3 = .333$. The answer is $.198/.604 = .328$.

57. Using the reliability function of SPSS, you will see that these 5 trials have an alpha coefficient of .842.

59. The only difference between concurrent and predictive validity is the *time* at which the criterion is measured.

With concurrent validity, the criterion is measured at the same (or nearly) time.

With predictive validity, the criterion is measured at some time in the future.

An example is physical activity. Physical activity could be correlated with $\dot{V}O_2\max$ measured at the same time; this is concurrent validity.

One's present level of physical activity could be used to predict the development of heart disease in later life; this is predictive validity.