

1. What makes a study an experiment? Why do researchers go to the effort of designing a study to be an experiment? Using variables from the questionnaire that you completed during the first lab, what might be an interesting experiment that you could run?
2. Eight applicants for a firefighter's job are asked to take a test twice. Their set of scores at time 1 was as follows: 27, 44, 82, 70, 43, 62, 41 and 89. Their set of scores at time 2 were as follows: 74, 71, 90, 11, 40, 3, 50, 29. What is the statistical reliability of this test? Show your analysis with SPSS output. Assuming that a reliability coefficient of at least 0.70 is required for the test to be useable for selecting applicants for this job, would you consider this test useable? What is the technical term for how useable the test is for a particular purpose?
3. A US citizenship test consists of the following 10 multiple-choice items:  
Item 1: What is the longest river in the United States east of the Mississippi?  
Item 2: Who was the 12<sup>th</sup> president of the United States?  
Item 3: What were some of the grievances that US colonists held against Britain in the mid-1700s?  
Item 4: What is the capital of Virginia?  
Item 5: What is the tallest mountain in the United States?  
Item 6: Montpelier is the capital of what state?  
Item 7: Can the US Supreme Court give advice to the president?  
Item 8: Which branch of Congress approves members of the president's Cabinet?  
Item 9: What were the major causes of the Civil War?  
Item 10: Who was the vice-president for Martin Van Buren?

10 people took this test. Each test-taker received 1 point for each item correctly answered and 0 points for each item incorrectly answered. Each person scored as follows on the 10 items (in item order):

Person 1: 1, 0, 0, 1, 1, 1, 0, 0, 1, 0	Person 6: 0, 1, 1, 1, 0, 1, 0, 0, 1, 1
Person 2: 0, 0, 0, 0, 0, 0, 0, 0, 0, 0	Person 7: 0, 0, 0, 1, 1, 1, 0, 1, 0, 0
Person 3: 1, 1, 1, 1, 1, 1, 1, 1, 1, 1	Person 8: 1, 1, 1, 1, 1, 1, 1, 1, 0, 0
Person 4: 1, 1, 0, 0, 1, 0, 0, 1, 0, 1	Person 9: 1, 1, 1, 0, 1, 0, 0, 0, 1, 1
Person 5: 0, 0, 1, 1, 1, 1, 1, 0, 1, 0	Person 10: 1, 0, 0, 1, 1, 0, 1, 1, 0, 0

- a. What is the most rational basis for splitting the test in half to calculate a split-half reliability? Why? Using this basis for splitting the test, calculate a split-half reliability for this test using SPSS, and show your SPSS output.
- b. By hand, calculate a split-half reliability for this test by splitting the test into odd-numbered items versus even-numbered items. Show all of your work. Do you obtain the same result for this split-half reliability as you did in part a, above?
- c. Let's assume that the rational basis for determining how to split the test in half that you provided for part a did not exist. Based on your findings for parts a and b, do you see any problem with the split-half method of calculating reliability?
- d. Next, adjust your first calculation of split-half reliability in part a to the original test length of 10 items. What is the numerical value of this adjusted reliability? Show your work. Next, adjust your calculation of split-half reliability to a test length of 20 items. What is the numerical value of this adjusted reliability? Show your work. Do you notice a pattern between the adjusted reliability values and the number of items on the test? (Please, do tell!)