

Psych./Neuro. 201
SPSS Instructions for Research Project #2 (Q-naire Construction)

Part I: Assessing the Internal Reliability of Your Questionnaire

You will be using SPSS to calculate a measure of internal reliability called “Cronbach’s alpha (α).” Cronbach’s α is a reliability coefficient based on the average covariance among items in a scale. We assume that items on a scale are positively correlated with each other because they’re all tapping into the same construct; that is, they’re all measuring a common entity (some items may need to be reverse scored so that they are all positively correlated). The average correlation of an item with ALL OTHER ITEMS IN THE SCALE tells us about the extent to which the items are all measuring the common entity. Because α can be interpreted as a correlation coefficient, it ranges in value from 0 to 1. (Negative α values can occur when items aren’t positively correlated among themselves and the reliability model is violated.)

Before you can analyze your scale’s reliability, however, you must do the following:

1. If you entered your data by hand you need to check for errors! One quick way to check for major mistakes is to run a **frequency** analysis on each variable. (See other SPSS instruction handout for how to do this.) A frequency analysis will tell you, for each item, how many people responded “1,” how many people responded “2,” and so on. If you find that one person responded “9,” you’ll know you’ve made an error if 9 wasn’t one of the choices. By examining your output, you can also detect if any of your items had **ceiling** or **floor effects**.
2. **REMEMBER TO RECODE THE QUESTIONS THAT ARE NEGATIVELY KEYED!** To do so, follow these instructions for using the “recode” command:

As we discussed in class, questionnaires are often constructed such that approximately half of the questions are framed positively and the other half are framed negatively. For example, in Rosenberg’s Self-Esteem scale (one of the most widely used measures of self-esteem), items 1, 2, 4, 6, and 7 are written so that high numbers correspond to high self-esteem (e.g., Item 7: “On the whole, I am satisfied with myself”) whereas items 3, 5, 8, 9, and 10 are written so that high numbers correspond to low self-esteem (e.g., Item 3: “All in all, I am inclined to feel that I am a failure”). If we wanted to compute an average self-esteem score for each participant based on his/her responses to all 10 items on the scale, we would need to reverse-score half the items so that they were all coded in the same direction (i.e., high numbers correspond to high self-esteem). To do this, we would go up to the “Transform” menu and scroll down to “Recode”. You get two options: “Into same variables” and “Into different variables.” Choose “Into different variables.” You should get the window you see below:

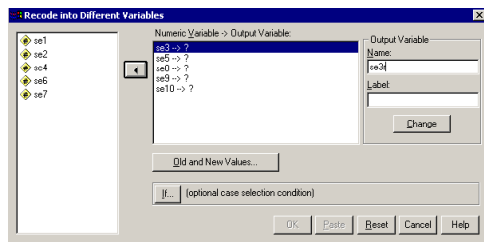


Figure 1

The names of all your variables will appear in the list on the left part of the window. Scroll down until you find the variables you’re looking for. In this case, we want to recode se3, se5, se8, se9, and se10. Simply click on the variable you want, then click on the arrow (which will be pointing the opposite of the way it’s pointing in the figure). Do this for all the variables you want to recode the same way. Now you need to tell the computer what you want the new (recoded) variables to be called. To make life easier, I usually just add “r” to the original variable name, to remind myself that the variables have been recoded. So, for example, “se3” becomes “se3r.” Type this new name in the “output variable” box on the right of the window and click on the “Change” button. Do this for all the variables you need to recode. Now you’re ready to tell the computer how to recode the variables. Since the Rosenberg Self-Esteem Scale uses a 4-point scale, you need to recode in the following way: 1=4, 2=3, 3=2, and 4=1. That is, in order for a response of “1” on question 3 to indicate high self-esteem, it needs to be recoded as a “4”

(you're just reversing the scale). To do this, first click on the "Old and New Values" box. You'll get the screen you see below:

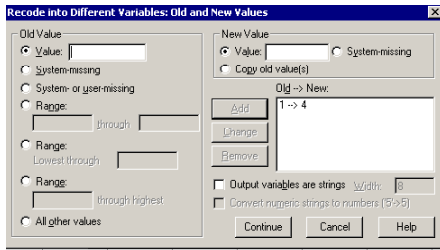


Figure 2

Note how I've already recoded 1 into 4; you can see it on the right side of the window in the "Old→New" box. To change 2 into 3, simply type 2 in the "Old Value" box on the left and type 3 in the "New Value" box on the right. Then click "Add." Keep going until you've recoded all 4 scores, then click "Continue." This will take you back to the window in Figure 1, and you can simply click "OK." If you now scroll to the end of your data file, you'll find five new columns there, one for each of the recoded variables.

Conducting a Reliability Analysis (Cronbach's α)

Now you're ready to check the reliability of your scale. Go up to the "Analyze" menu and choose "Scale," then "Reliability Analysis..." You should see a window like the one in Figure 3. Place the variables you want into the box labeled "Items" by clicking on each one in the list on the left and then clicking on the arrow. Make sure to include the recoded items instead of the originals, where appropriate. When all variables have been entered, click the box marked "Statistics." You'll get a window like the one in Figure 4.

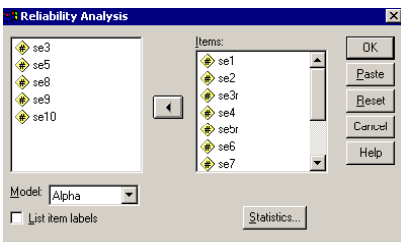


Figure 3

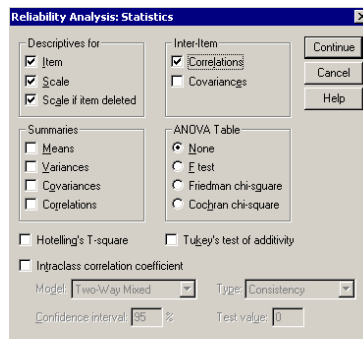


Figure 4

In the upper right-hand corner, check the box labeled Inter-Item Correlations. This will give you a correlation matrix for all items in your scale, allowing you to look at how highly each item correlates with each other item. In the upper left-hand corner, click in all 3 boxes under "Descriptives for" (i.e., scale, item, and scale if item deleted). This will give you the mean and standard deviation for each item, for the scale as a whole, and for the whole scale if each item were to be deleted. In addition, you will get Cronbach's alpha, which I described above. A high alpha (.70 and higher) would be consistent with your hypothesis that all of your scale items are measuring the same construct.

If you have a low alpha (below .70), check the column of your output labeled "Cronbach's alpha if item deleted." Would your alpha go up if you got rid of any items? Which item(s)? You should examine the column labeled "Corrected item-total correlations." Note that items with low item-total correlations are the ones that would allow you to increase the scale's alpha by deleting them. Go to your questionnaire to take a look at the items with low item-total correlations. Can you see what might be wrong with these items or why they might not be measuring what all the other items are measuring? Use your judgment in determining which items to delete. If deleting an item increases the alpha by only a miniscule amount, don't delete it.

Using the Compute Command to Create a Mean Score (Composite Variable)

Once you have determined which items should remain in your final scale, you will need to compute what is known as a "composite variable" for the scale. That is, you will compute a variable that represents each participant's mean of all the scale items. Let's return to our example of the Rosenberg Self-Esteem Scale. The scale has 10 items, so if I wanted to

compute each participant’s mean self-esteem score based on those 10 items, I could use a compute statement. Don’t forget that we need to use the recoded items instead of the originals for questions 3, 5, 8, 9, and 10. To create our compute statement, we would go up to the Transform menu again, and scroll down to “Compute.” You should then see a window that looks like the one below:

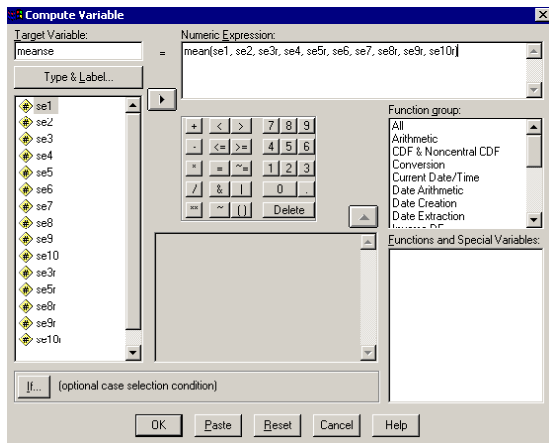


Figure 5

Note how I’ve typed “meanse” in the “Target Variable” box. That means I want SPSS to create a new column labeled “meanse” that will represent each participant’s mean on the self-esteem scale. In the “Numeric Expression” box, I’ve indicated that I want this new variable to equal the mean of the items on the self-esteem scale, substituting the recoded items for the originals where appropriate. You can use the compute command to perform any of a number of functions. I’ve chosen to calculate a mean. [Note: If you wanted to save this formula to use again later, you could click “Paste,” and the formula would get pasted into the Syntax window, where you could run the command but also save it.] When you’re all set, click “OK.” If we scrolled to the end of the data file containing the SE scores, we’d see a new column labeled “meanse.” This variable would represent each participant’s average SE score based on the 10-item scale.

NOTE: In class, each group will calculate Cronbach’s α for its own scale, and will decide which scale items to retain in the final scale. Because all groups will need to have composite variables for each of the groups’ scales, we will ask each group to write on the board which items were included in the final scale. That way, each group can then create a composite variable for each of the scales using the Compute command.

Part II: Assessing the Convergent and Discriminant Validity of Your Questionnaire

Before you begin to assess the convergent and discriminant validity of your measure, you should already have decided, based on your nomological net, which scale(s) should correlate most and least highly with your own. Once you have made these predictions, you are ready to begin examining the correlations among the scales.

Before you conduct a Pearson’s r , it is very important to examine a scatterplot of the two variables to get an idea about whether they are linearly related to one another. In SPSS, go under the “Graphs” menu to “Legacy Dialogs” and choose “Scatter/Dot...” You will see the window in Figure 6. Choose “Simple Scatter” (it should already be selected) and click on “Define.” You will now see a window like the one in Figure 7.

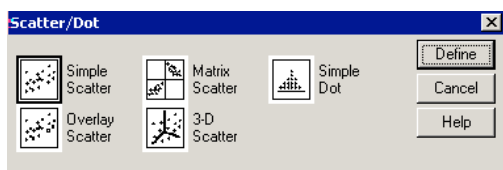


Figure 6

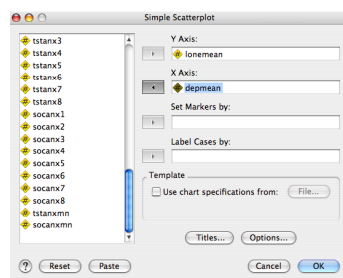


Figure 7

Scroll down the list of variables on the left until you find the variable you want on the Y axis (the criterion variable). Click on this variable, then click the top arrow. The variable name should now appear in the box labeled “Y Axis.” Now find the variable you want on the X axis (the predictor variable). Click on this variable, then click on the arrow next to the “X axis” box. Click OK. SPSS will bump you into the output window, where you’ll see your scatterplot. Take a look at it. Are there any outliers? Does the relationship appear linear? (If it’s curvilinear, a Pearson’s r is inappropriate to use.) Double-click on the graph, and when SPSS bumps you into the chart editor, go under the “Elements” menu and choose “Fit line at total.” [You can close the “Properties” box that pops up.] SPSS will draw in the least-squares line, which will better help you see whether the data are linearly related, and whether the relationship is positive or negative. If the scatterplot looks OK, you are ready to proceed to calculating the correlation.

To do so, go to the “Analyze” menu, scroll down to “Correlate” and choose “Bivariate.” You will see a window like the one below:

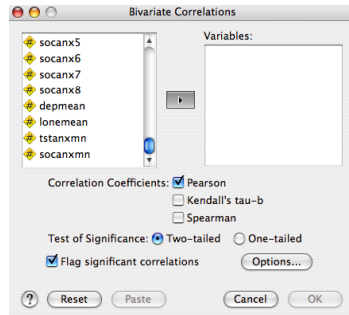


Figure 8

Find the two variables you used above. Select each and place them both in the “Variables” box. Where it says, “Correlation Coefficients,” make sure that “Pearson” is selected (this is the default). Now click OK. SPSS will bump you into the output window and produce a correlation matrix. The top value is Pearson’s r (the correlation). Below that, the “Sig (2-tailed)” value is the p value, or the probability of getting a correlation that big if there really was no relationship between the two variables in the population. In psychology, we use $p < .05$ as our criterion for statistical significance; that is, if $p < .05$, there is only a 5 in 100 (5%) chance that the two variables are not really correlated with each other in the population—an error rate we are willing to live with. Finally, the bottom number is N (the sample size). The proper way to report a correlation is as follows (for example): $r(58) = .43, p = .005$. [Where the number in parentheses is df , which is $N - 2$ for a correlation.]

Make sure you correlate your scale with the two or three others you have chosen for the purposes of determining convergent and discriminant validity. Look at the pattern of r values. Does your scale correlate relatively strongly with the scale you used to determine convergent validity? Does it correlate less strongly with the scale you used to determine discriminant validity?

A note about linear regression (but you do not need this for the Project 2 paper; this is just for your reference): If two variables are correlated, then scores on one variable can be used to predict scores on the other. To conduct a linear regression in SPSS, go under “Analyze,” “Regression,” “Linear.” Put the criterion (y) variable in the box labeled “Dependent” and the predictor (x) variable in the box labeled “Independent(s).” Click OK. In your output, you’ll see that the Multiple R is the same as what you’d get if you conducted a correlation. The R Square (r^2) is the coefficient of determination, or strength of the relationship between x and y . The Std. Error of the Estimate is not quite the same as the one described in chapter 5 of your text; rather, it is an *estimate* of the standard error of estimate in the population, discussed in chapter 14 of your book. Under the part of the printout labeled, “Coefficients,” you’ll see numbers listed under the column “B” for your predictor variable and for (Constant). The number under B next to your predictor variable is the SLOPE of the regression line. The number under B next to (Constant) is the Y INTERCEPT of the regression line.