Interrater Agreement Using SPSS Dropdown Menus

In this document I explain how to use SPSS to obtain iInterrater agreement indices using SPSS dropdown menus. For instructions on using SPSS syntax, see the document "Obtaining Interrater Agreement Using SPSS Syntax."

The data for these examples are taken from Table 9.1 in the book (p. 211) and are in the SPSS dataset "**rater data.sav**"

For interrater agreement, SPSS will calculate Cohen's kappa and will produce tables such as that shown in Table 9.2 in the book. SPSS versions 26 and higher can also compute Fleiss's (1971) multirater kappa statistic.

I discuss two ways to obtain values of nominal agreement using SPSS.

The first way is to obtain a crosstabulation table for the ratings from the two raters of interest. Then simply count the numbers on the diagonal of the table and divide by the total to obtain the nominal agreement index.

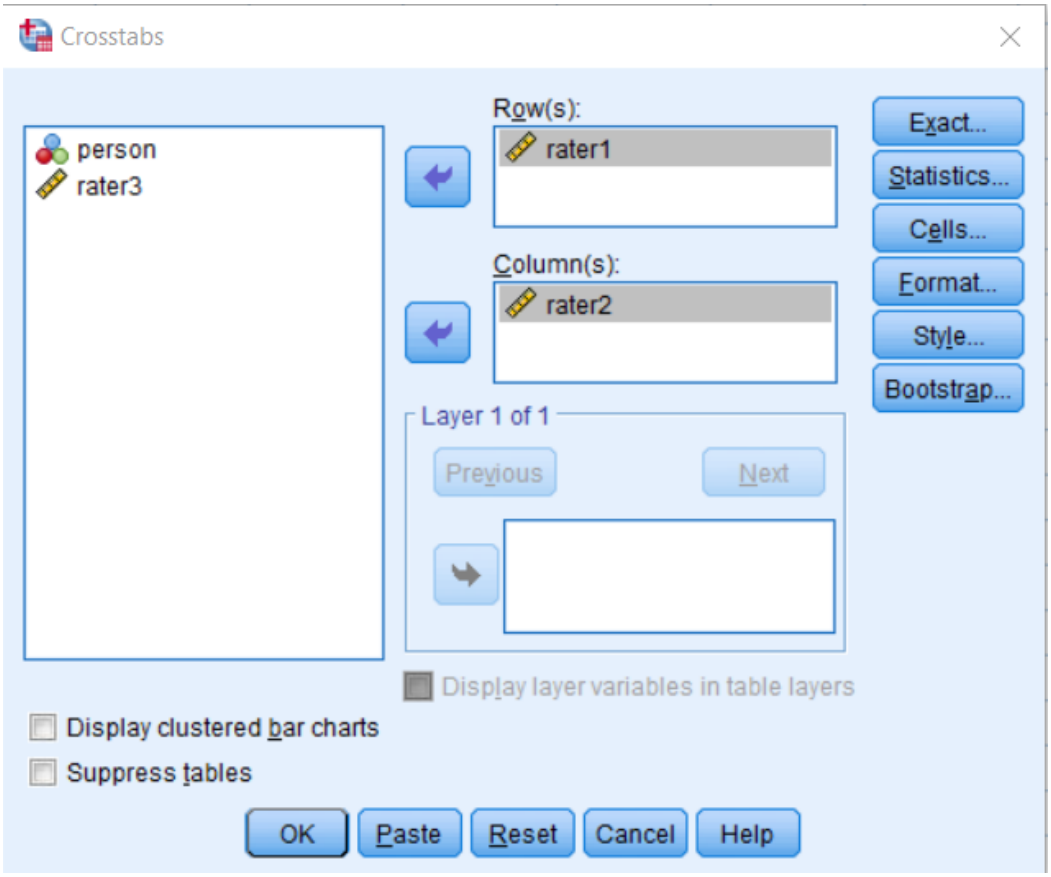
The second way is to use the **compute** menu to calculate agreement between pairs of raters and then use the **descriptives** command to obtain the mean of these agreements. This mean is the nominal agreement.

Nominal Agreement

Nominal agreement using crosstabs

The first way of obtaining values of nominal agreement is to use the **crosstabs** procedure to obtain the necessary table.

From the **analyze** menu, choose **descriptive statistics**, then **crosstabs**. You will get the screen below:



When using **crosstabs** to obtain tables for nominal agreement, only two variables at a time can be analyzed. Here, I have chosen to assess the agreement between raters 1 and 2.

Click on **OK** and you will get the output below:

rater1 * rater2 Crosstabulation

Count

		rater2					
		1	2	3	4	5	Total
rater1	1	1	0	0	0	0	1
	2	1	1	0	0	0	2
	3	0	1	2	0	0	3
	4	0	0	1	1	0	2
	5	0	0	0	0	2	2
Total		2	2	3	1	2	10

This table is that same at Table 9.2 in the book. I have highlighted the values that indicate exact agreement; these appear on the diagonal. Nominal agreement is simply the sum of these agreements divided by the total number of people being rated (10).

This is equal to:

$$P_O = \frac{1}{N}\sum_{i}^{C} n_{ii} = \frac{1}{10}(1 + 1 + 2 + 1 + 2) = \frac{1}{10}(7) = .7$$

P_0 is the nominal agreement, *N* is the number of people being rated, *C* is the number of rating categories, and n_{ii} represents the number on the diagonal of the matrix.

The second way to obtain values of nominal agreement is to use the **compute** procedure to calculate the agreement between the two raters of interest and then calculate the mean of these agreements using the **descriptives** procedure.

Nominal agreement using compute

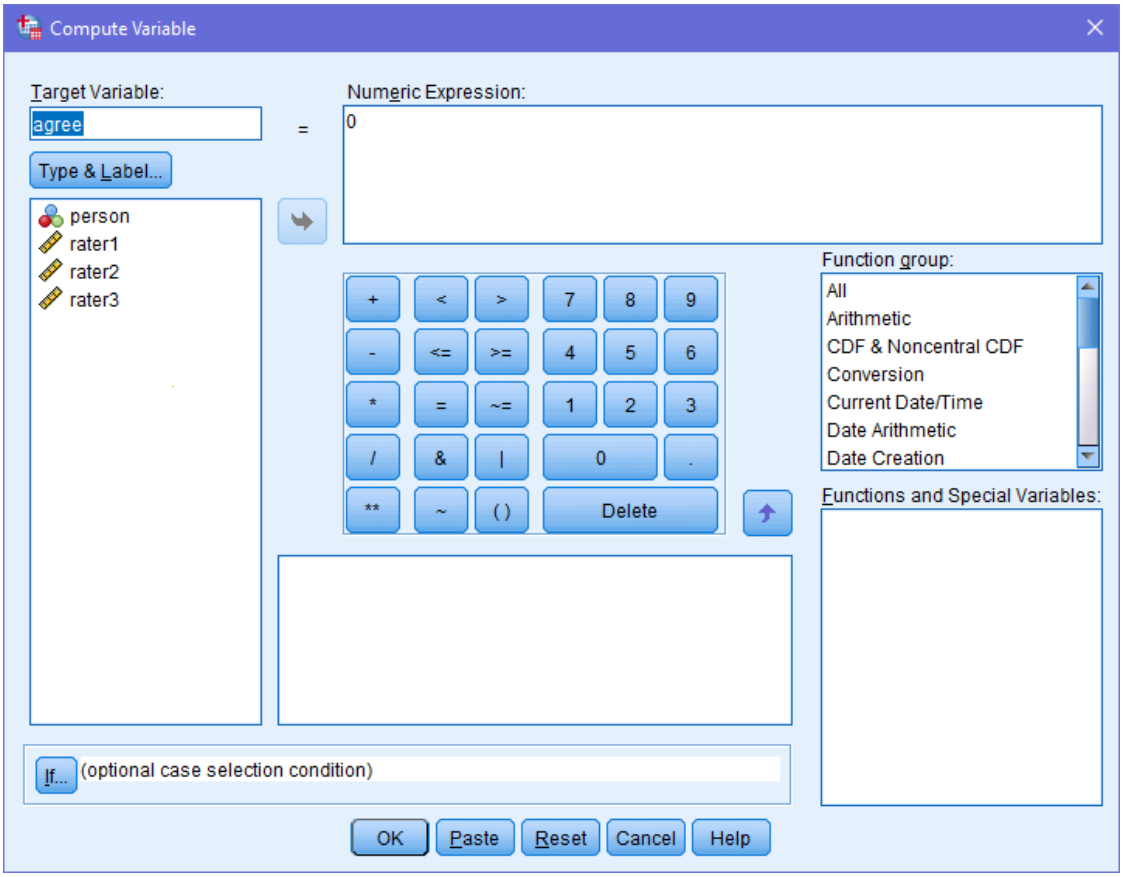
To calculate agreement between two raters, use the **compute** function. Below I explain how to do this using dropdown menus. The procedure is as follows: First, assign everyone the value of 0 on the agreement variable. Then, change that value to 1 for those with the same rating from the two raters. This will result in a variable with values of 0 for those with different ratings and values of 1 for those with the same rating.

To create the agreement variable, click on **transform**, then **compute**.

In the box labeled "**Target Variable**" type the name for the agreement variable (here, I use "agree").

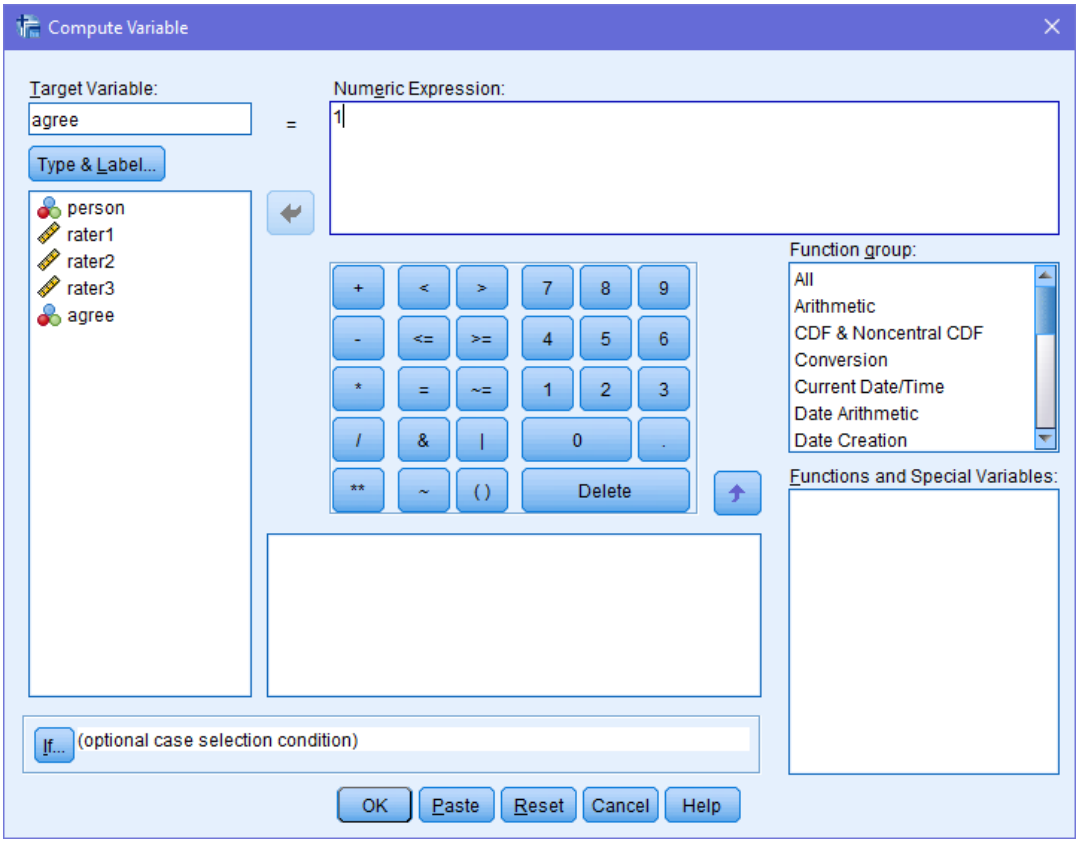
In the box labeled "**Numeric Expression**" type 0. Then click on ***OK*.**

This will create a new variable in named "agree" in your SPSS dataset whose value will equal 0 for everyone.



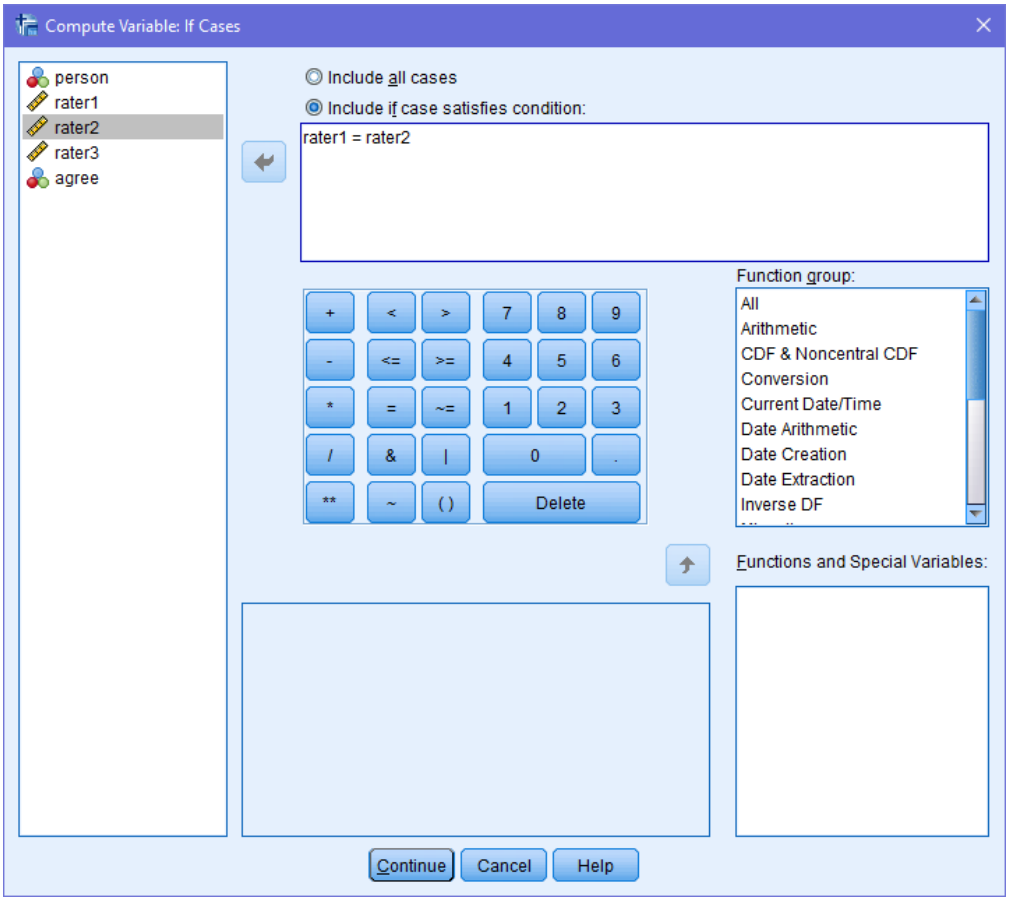
In the next step we will change the value of "agree" to equal 1 if a person obtains the same rating from each of the two raters.

Repeat the steps above, but type "1" in the box for **numeric expression** instead of 0. Keep the name of the variable the same (agree).



Then click on the **If** button at the bottom of the page. You will get the screen below.

Click on the radio button for **include if case satisfies condition** and type in “rater1 = rater2.”

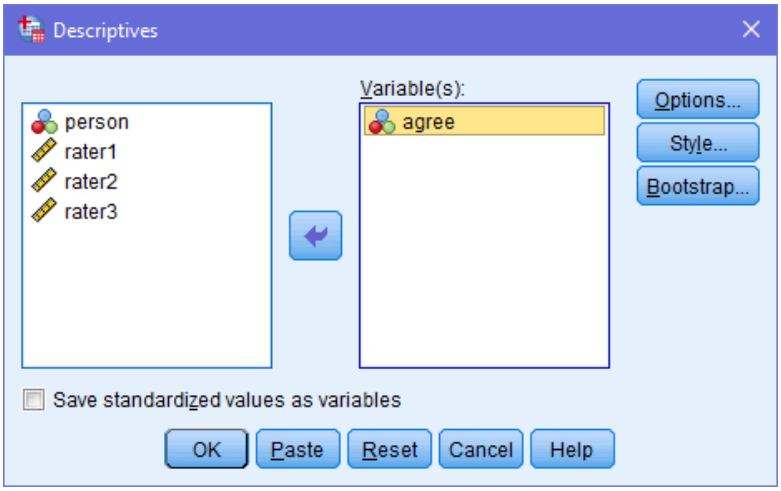


Click on **continue** and then **OK**. When you are asked "**change existing variable?**" click **yes**.

The variable "agree" at the end of your SPSS dataset should now have values of 0 for those with different values for rater1 and rater2, and values of 1 for those with the same values from the two raters.

The final step is to use the **descriptives** command to calculate the mean of the variable "agree." This mean will be the value of nominal agreement.

Choose **analyze, descriptive statistics**, then **descriptives**. Then choose the variable "agree."



Click **OK** and you will get the output below:

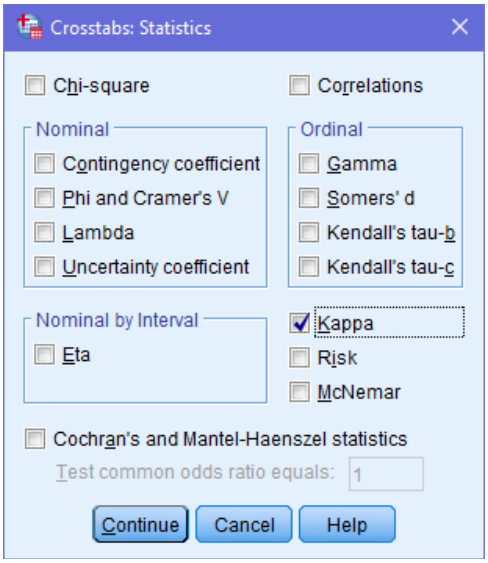
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
agree	10	.00	1.00	.7000	.48305
Valid N (listwise)	10				

The value of .7000 under "mean" is the value for nominal agreement. Note that this is the same value we obtained by counting the entries on the diagonal and dividing by the total number of ratings.

Cohen's Kappa

To obtain values of Cohen's kappa, use the **crosstabs** procedure as before, but click on the **statistics** button and choose **kappa**.



Click on **continue** and then **OK**. You will get the same crosstabulation table as before, plus the output below.

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.620	.182	3.922	.000
N of Valid Cases		10			

a. Not assuming the null hypothesis.

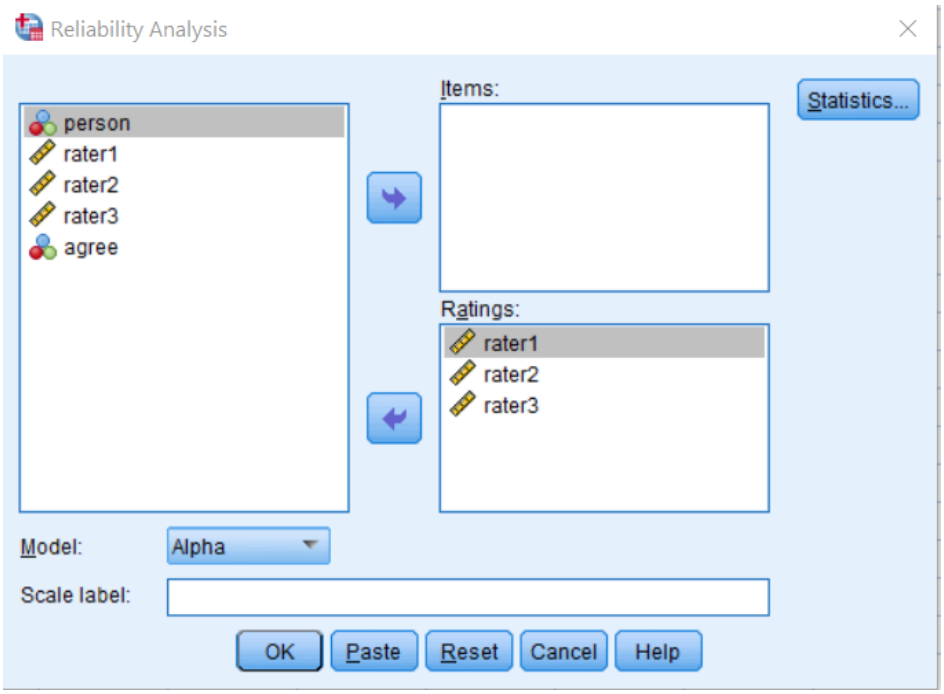
b. Using the asymptotic standard error assuming the null hypothesis.

The value of kappa is .620, as calculated in the book on page 214.

Fleiss's Multirater Kappa

To obtain Fleiss's multirater kappa, choose **scale** from the **analyze** menu. The choose **reliability analysis.**

You will get the screen below. Choose the three rater variables and click on the arrow to put them in the box labeled **ratings**. The click on **OK**.



You will get the table below.

Overall Agreement[a]

	Kappa	Asymptotic Standard Error	Asymptotic z	Asymptotic Sig.	Asymptotic 95% Confidence Interval Lower Bound	Asymptotic 95% Confidence Interval Upper Bound
Overall Agreement	.363	.094	3.850	.000	.357	.369

a. Sample data contains 10 effective subjects and 3 raters.

The value of Fleiss's multirater kappa is .363 for the three raters.