

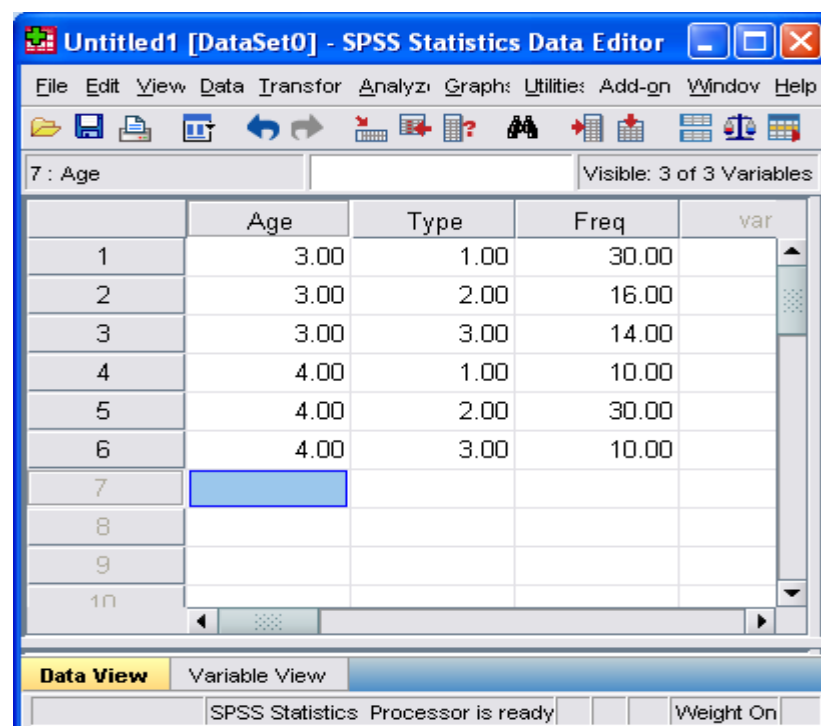
Research Methods

Two variable Chi-squared test on SPSS

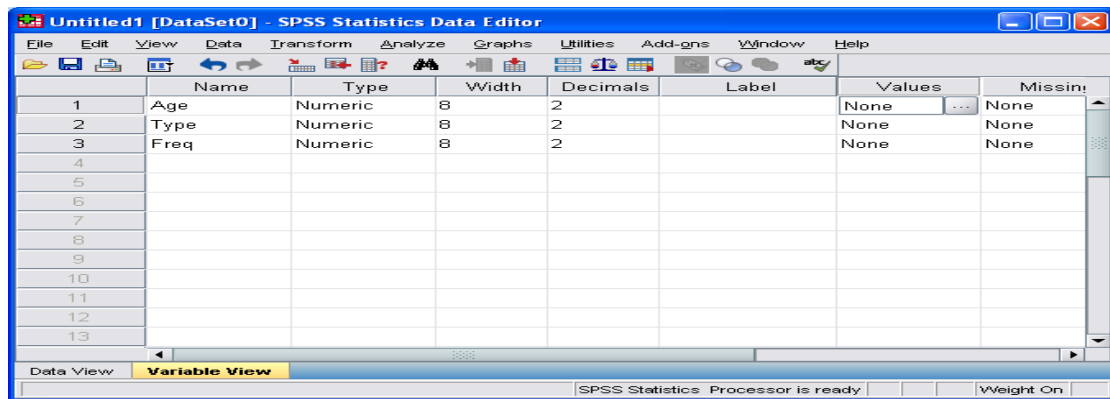
We will use SPSS to do a chi-squared test on the two way table used in the lecture (the *age x type of play* cross-tabulation). The table is reproduced below:

	Associative	Cooperative	Parallel
3 year-olds	30	16	14
4 year-olds	10	30	10

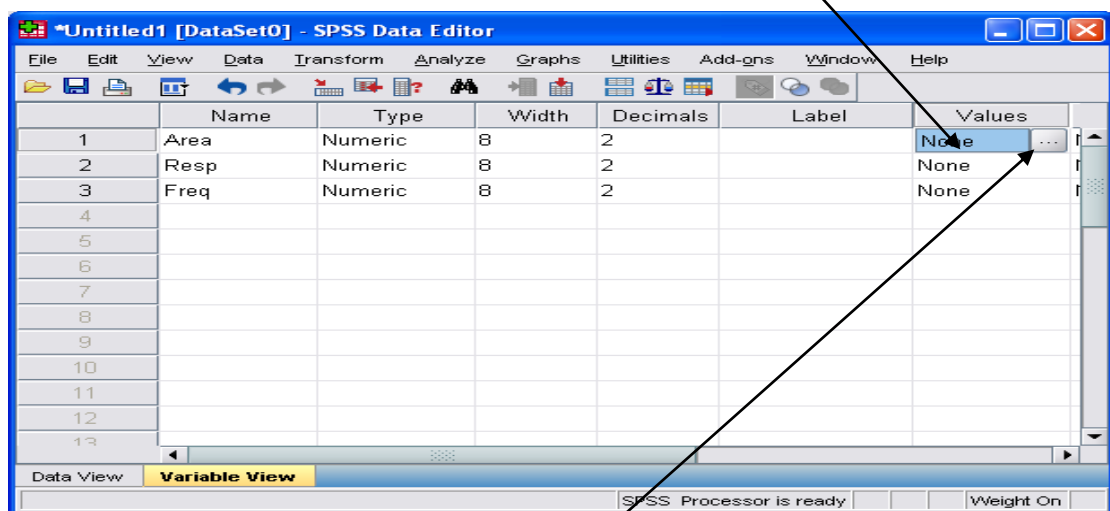
Firstly, this table has to be represented in the SPSS data editor. The manner in which you will do this is similar to the way the data file was structured for the unrelated t-test, except that here, there are two between subjects variables (*age 3/4* and *associative/cooperative/parallel play*). Therefore, you will need **two coding variables**, and a variable to represent the **frequencies**. So in the data editor, define three variables: *age* (with *three year olds* represented by a '3', *4 year olds* by a '4'), *type* (*associative* = '1', *cooperative* = '2', *parallel* = 3), '*freq*' (the individual frequencies). The final data file will look as follows:



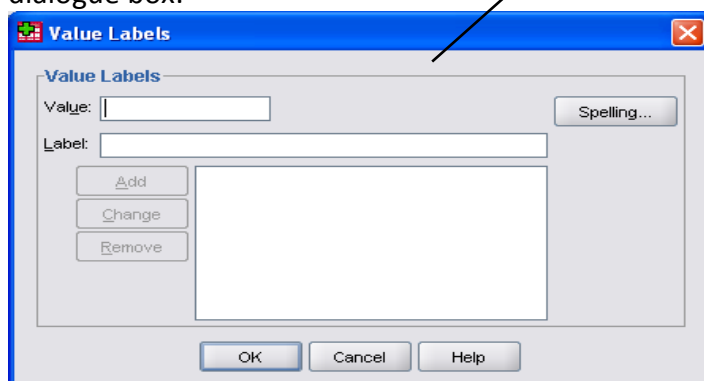
So the first row represents the 30 children in aged 3 who played in an associative manner and so on. To make the output more readable, you can also as SPSS to replace the numerical codes in the variables **Age** and **Type** with proper labels. To do this, go back to the **Variable View** tab:



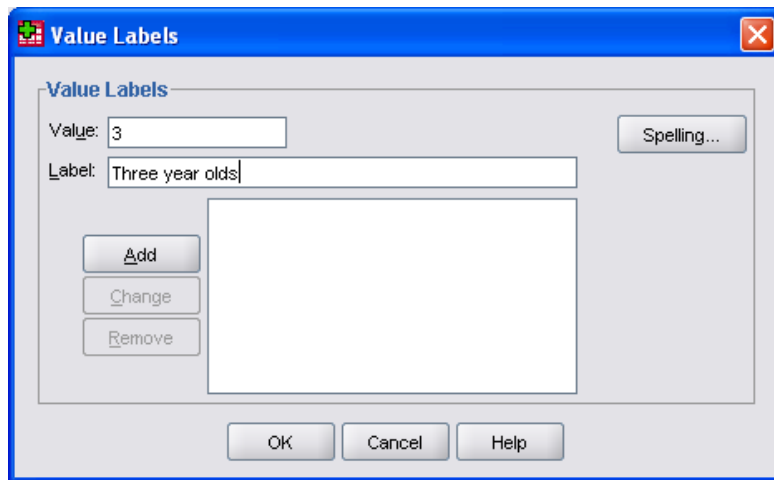
In the first row (describing the variable Area, click on the **Values** cell for this row:



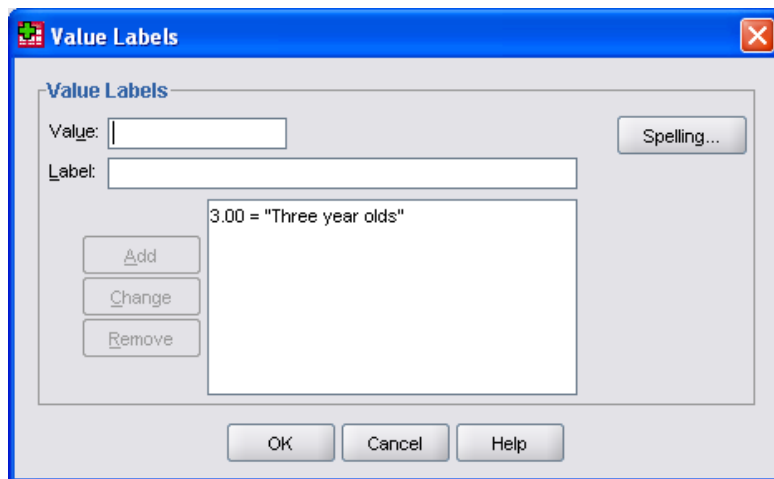
Note that a little grey box appears. Click on this box to produce the following dialogue box:



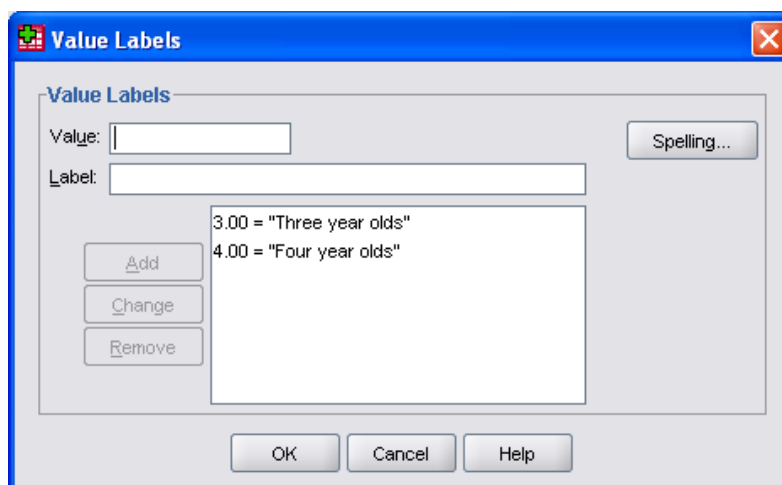
Instead of referring to the two age groups as '3' and '4' in the output, you would like SPSS to use the labels 'Three year olds/Four year olds'. So firstly, ask SPSS to replace the value 3 in this variable with 'Three year olds'. In the **Value** box, type 1 and in the **Label** box type Area A:



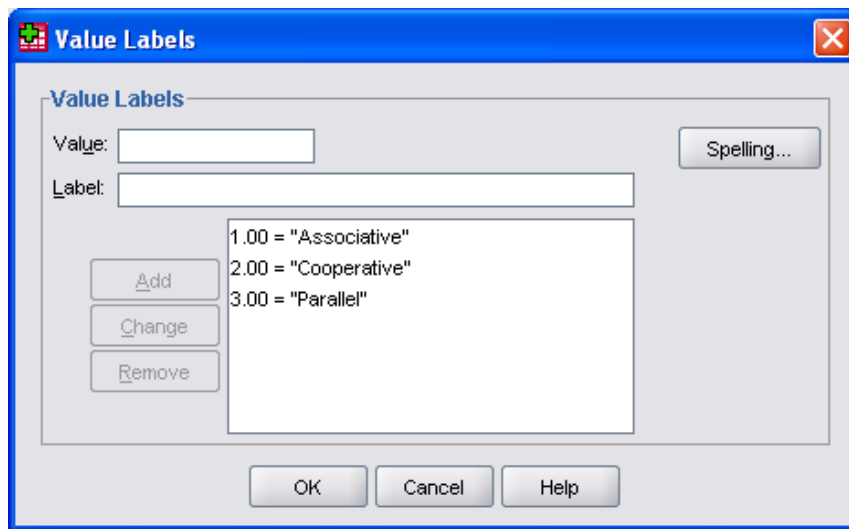
When you have done this click **Add**. The dialogue box changes as follows:



Now repeat the process to replace '4' by *Four year olds*. When you have done this click **Add** and the two new labels are displayed in the dialogue box:



Now repeat this process for the variable Type, replacing '1' with 'associative', '2' with 'cooperative' and '3' with 'parallel'. The final dialogue box should look as follows:



Now return to the data file, which should now look as follows:

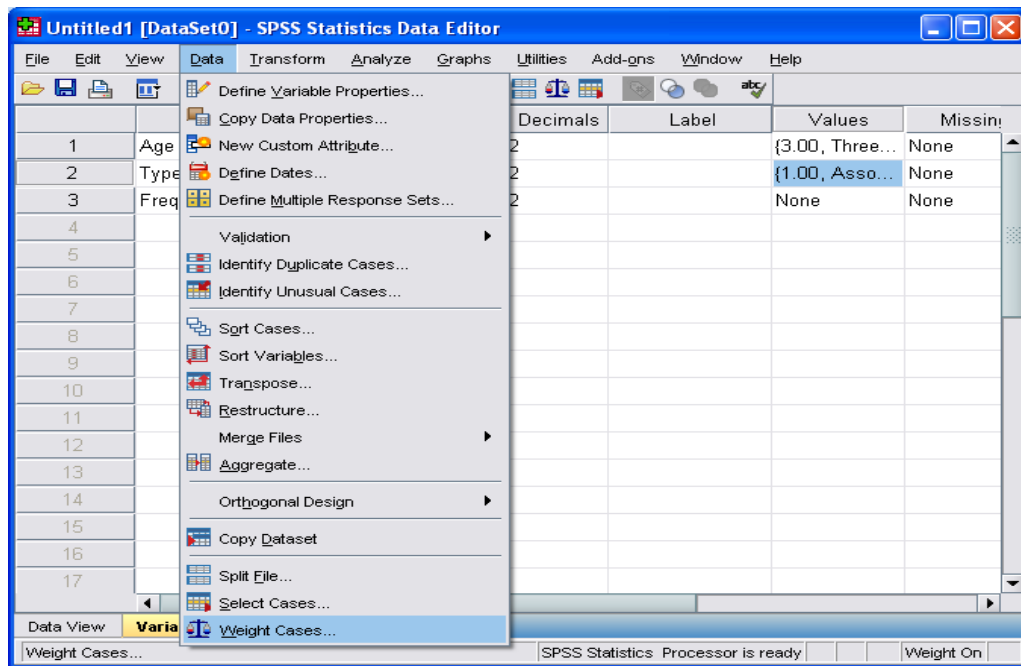
The SPSS Data Editor window shows the data file with the following data:

	Age	Type	Freq	var	var	var	var	var
1	Three year ...	Associative	30.00					
2	Three year ...	Cooperative	16.00					
3	Three year ...	Parallel	14.00					
4	Four year ...	Associative	10.00					
5	Four year ...	Cooperative	30.00					
6	Four year ...	Parallel	10.00					
7	.	.	.					
8	.	.	.					
9	.	.	.					
10	.	.	.					
11	.	.	.					
12	.	.	.					
13	.	.	.					

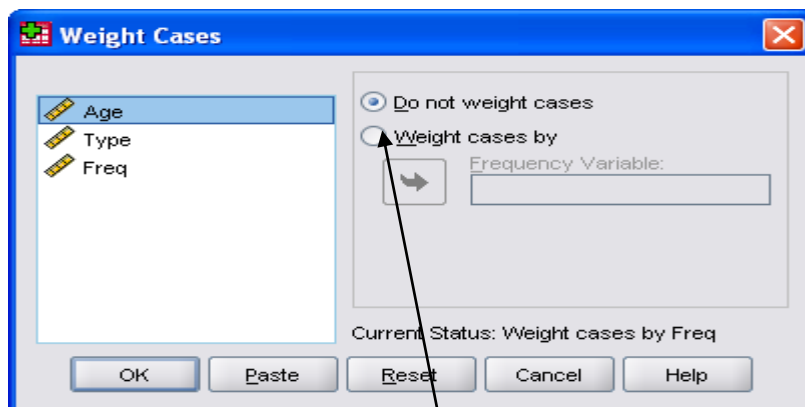
If the Age and Type columns are still displaying numerical data, click on the **Value Labels** icon, and this will change the columns to display the category names.

Do the chi-square analysis

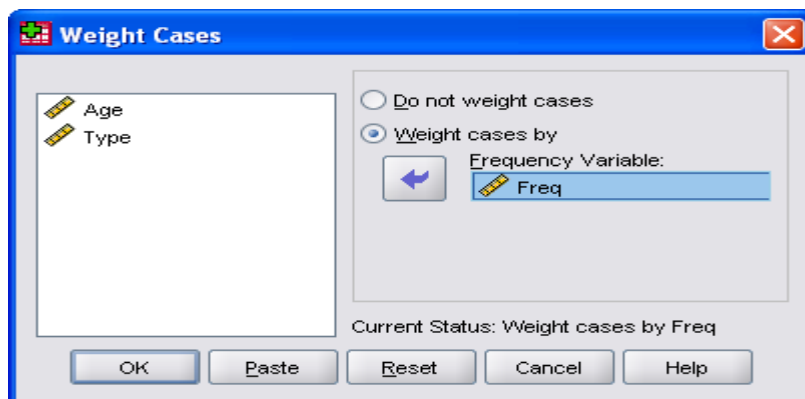
Firstly you have to tell SPSS that the variable 'freq' represents the frequencies of occurrence in the above table. To do this go to the '**Data**' menu at the top of the window, and from this select '**Weight cases**':



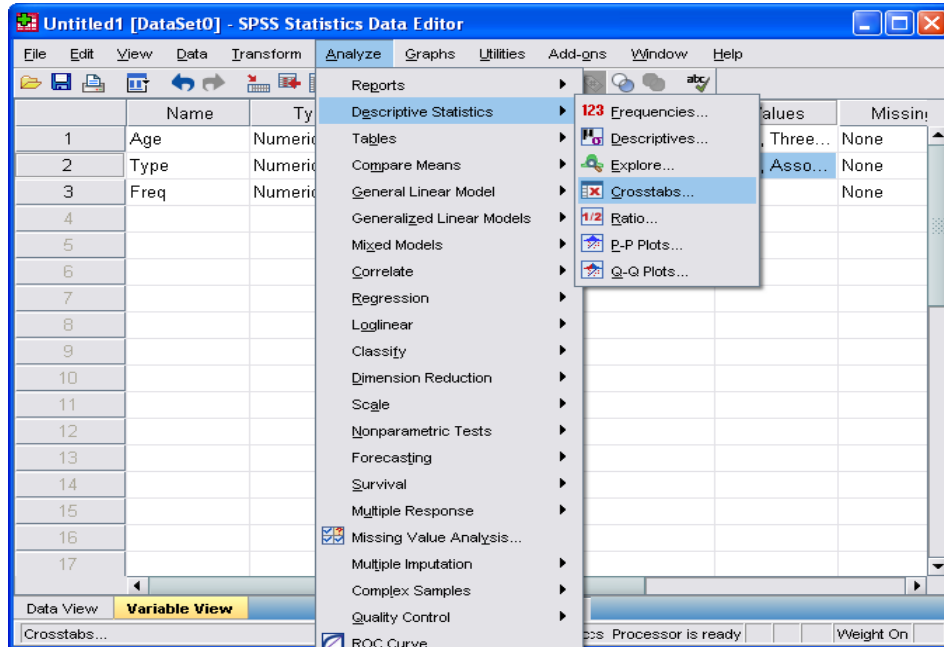
You will obtain the following dialogue box:



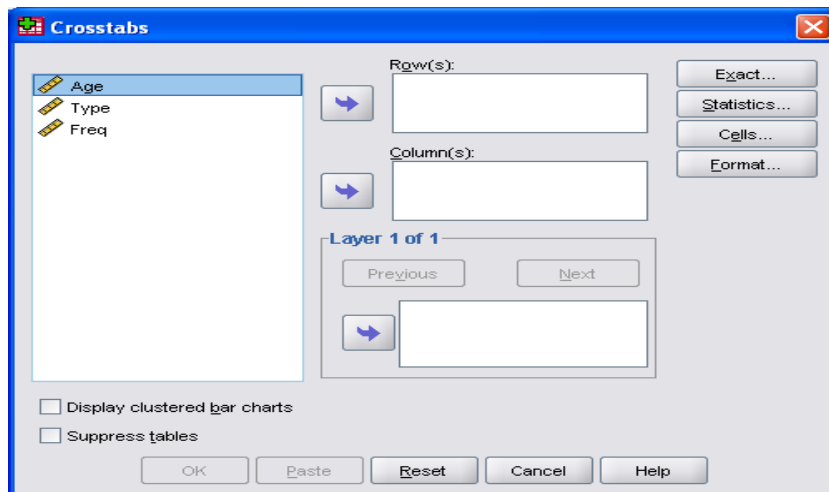
Click the mouse in the '**Weight cases by**' button. The select 'freq' by clicking the mouse on this variable name and click the arrow button which will have the effect of placing this in the '**Frequency Variable**' box:



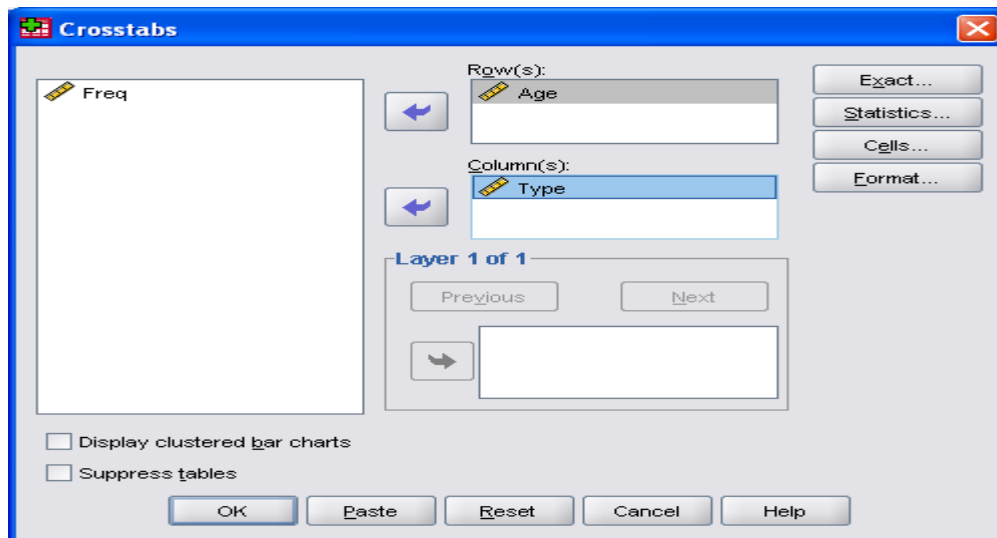
Click the **OK** button. The dialogue box will then disappear and you will get an output window telling you the command has been executed. Now go to the '**Analyse**' menu. From this select '**Descriptive Statistics**' and from this sub-menu, select '**Crosstabs...**'.



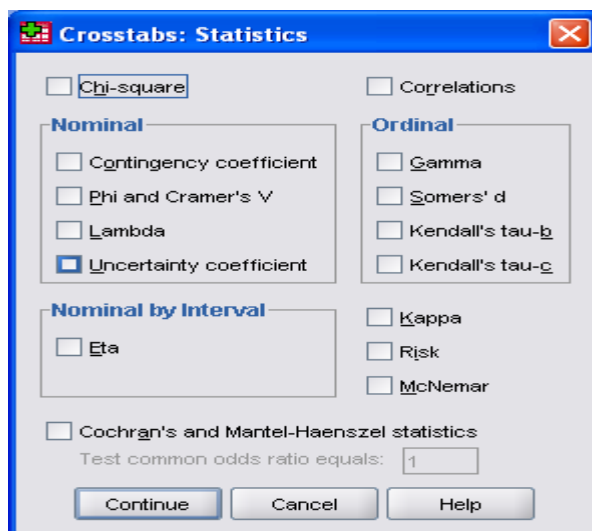
You will obtain the following dialogue box:



Select 'Age' and place this in the '**Rows**' box by clicking the arrow button next to it. Similarly, select 'Type' and place this in the '**Columns**' box:



Next, click the '**Statistics**' button:



In the resulting dialogue box, click the '**Chi Square**' box at the top left hand corner of this. Click '**Continue**' which will return you to the main dialogue box, and here click **OK**. SPSS will then do the analysis and provide the following output:

After a 'case processing summary' SPSS reproduces the contingency table on which you did the analysis:

Age * Type Crosstabulation

Count		Type			Total
		Associative	Cooperative	Parallel	
Age	Three year olds	30	16	14	60
	Four year olds	10	30	10	50
Total		40	46	24	110

The next table provides the results of the Chi-Squared test:

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	14.135 ^a	2	.001
Likelihood Ratio	14.554	2	.001
Linear-by-Linear Association	3.428	1	.064
N of Valid Cases	110		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.91.

You can see that you have a range of test statistics for chi-squared. The first of these – Pearson is the 'standard' way of calculating the statistic and in general this is the one you will use most of the time. You can see that the resulting value of χ^2 of 14.135 has a significance value of .001, indicating a statistically significant result.

To report this result:

Differences in play patterns between the two age groups were apparent - the majority of the three year old children engaged in associative play, whereas the majority of the four year olds engaged in cooperative play. A chi-squared test indicated a significant association between age and type of play ($\chi^2 = 14.135$, $df = 1$, $p < .05$).

Important note about Chi-squared Statistic

As mentioned above, most of the time we would use the **Pearson** value as our chi-squared statistic. The one exception to this is cases where we have a 2x2 cross tabulation (i.e. two rows and two columns). If you do a chi-squared test on this type of table, you will see that in the results you get an additional value of chi-squared called a '**continuity correction**'. In the case of a 2x2 table, this provides the most accurate value for the test statistic, so this is the value you should use, the interpretation is the same as any other test statistic: The significance value should be less than or equal to .05.