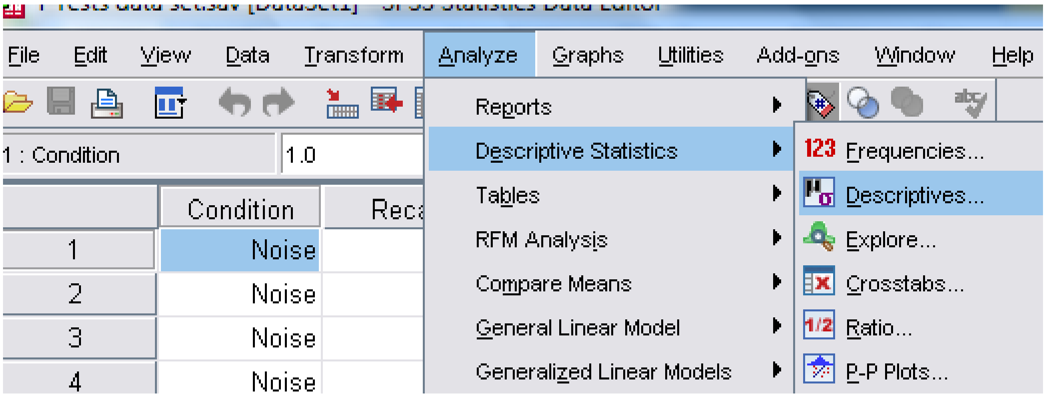
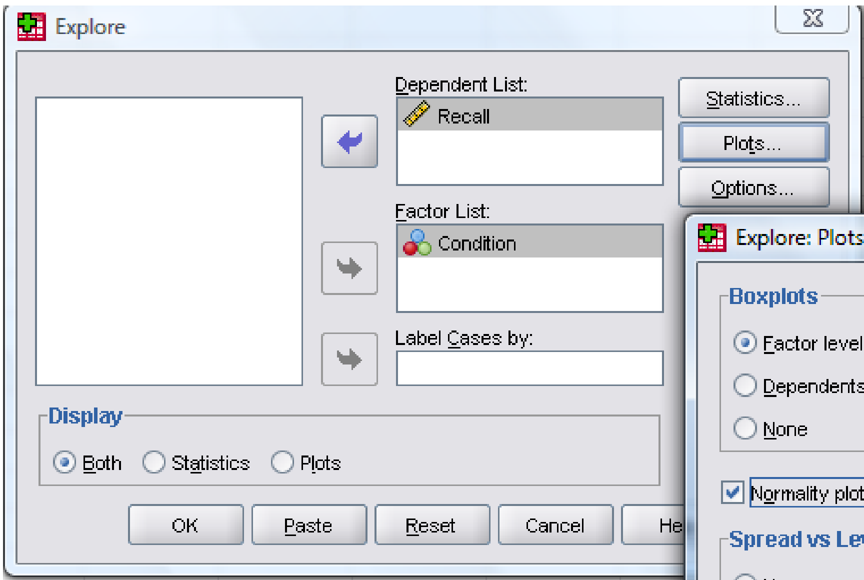
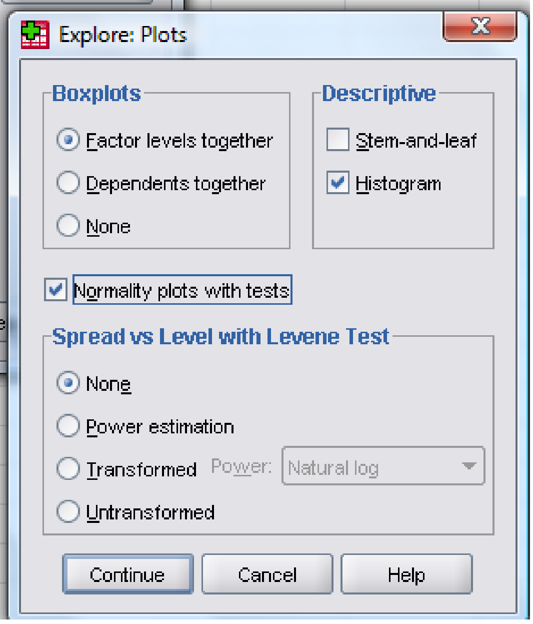
SPSS – *t* tests (and NP Equivalent)

Descriptive Statistics

To get all the descriptive statistics you need: Analyze > Descriptive Statistics>Explore.







Enter the IV into the Factor list and the DV into the Dependent list.

Select plots and tick Histogram and Normality plots

Descriptives

	Condition			Statistic	Std. Error
Recall	Noise	Mean		7.25	.719
		95% Confidence Interval for Mean	Lower Bound	5.67	
			Upper Bound	8.83	
		5% Trimmed Mean		7.28	
		Median		6.50	
		Variance		6.205	
		Std. Deviation		2.491	
		Minimum		3	
		Maximum		11	
		Range		8	
		Interquartile Range		5	
		Skewness		-.003	.637
		Kurtosis		-1.104	1.232
	No Noise	Mean		13.83	.796
		95% Confidence Interval for Mean	Lower Bound	12.08	
			Upper Bound	15.59	
		5% Trimmed Mean		13.87	
		Median		14.00	
		Variance		7.606	
		Std. Deviation		2.758	
		Minimum		9	
		Maximum		18	
		Range		9	
		Interquartile Range		5	
		Skewness		-.198	.637
		Kurtosis		-.937	1.232

This table can be used to find the mean, 95% CI around the mean, and SD. Additionally you can find information on the distribution of data: skewness and kurtosis.

Tests of Normality

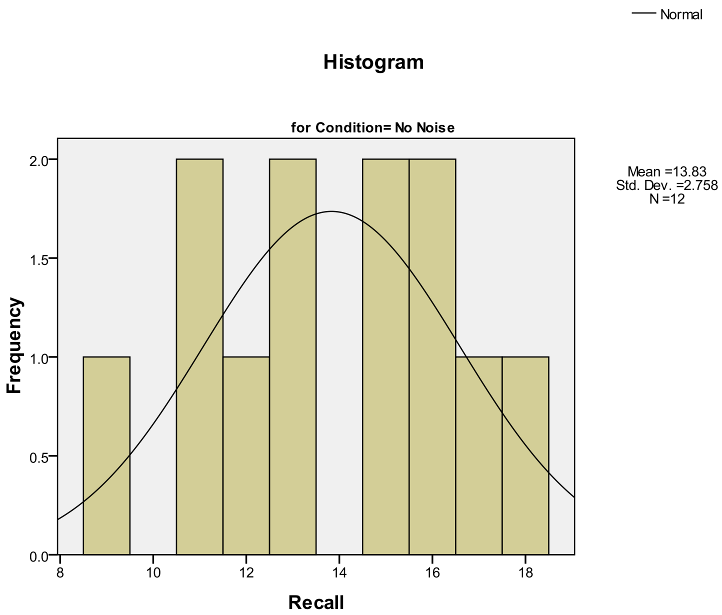
		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Condition	Statistic	df	Sig.	Statistic	df	Sig.
Recall	Noise	.192	12	.200*	.938	12	.467
	No Noise	.164	12	.200*	.963	12	.825

a. Lilliefors Significance Correction

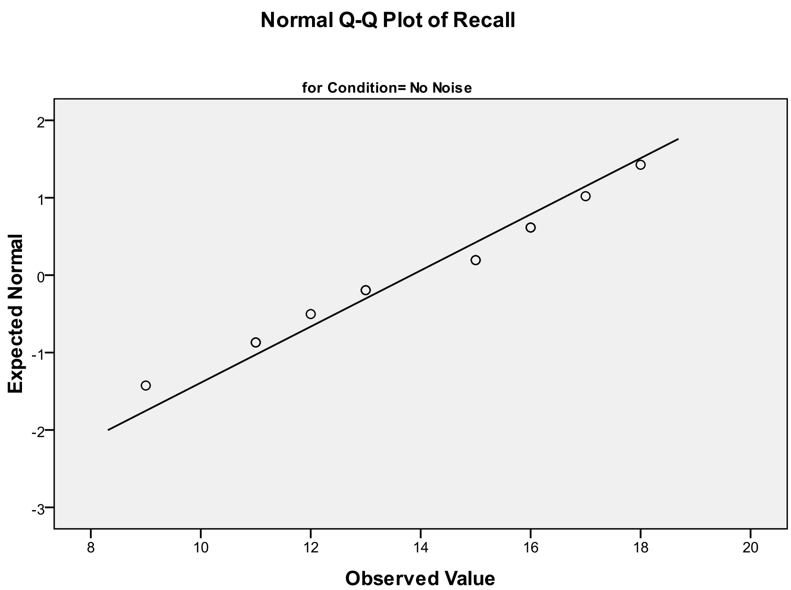
*. This is a lower bound of the true significance.

These are statistical tests for normality of data. I read the Shapiro -Wilk and if this is significant then we have a problem with our distribution. However, in large sample sizes these become overly sensitive when small deviations produce significant results, therefore you should also visually inspect the histograms.

Because the above test is unreliable in larger sample sizes I often assess for normality using a z score calculation instead. This is based on figures found in the explore table – the skewness and kurtosis and their respective standard error terms. To test for normality using this you calculate: z = s/se (i.e. skew / standard error of skew). In the above example this would be: - 0.003 / 0.343 = 0.0175; thus z = 0.0175. The general rules are: for small sample size (which is what you will be dealing with) a score of z = 1.96 or above indicates a problem.



In Psychology you rarely get a perfectly normal distribution. I would be happy with the above histogram. Further investigation can be made from the normality plots: see Figure 2.

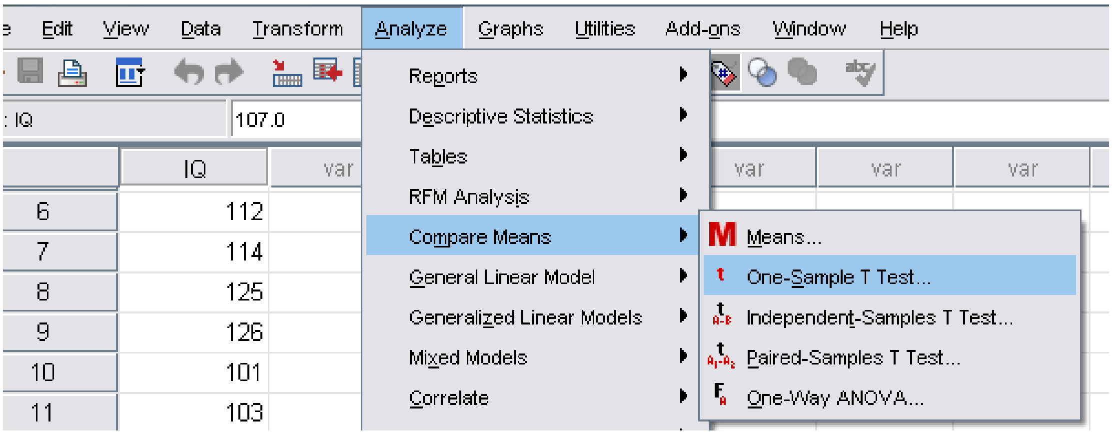


The Q-Q plots the observed values against those that are predicted if your data were normal (a bit like a scatterplot). Therefore your line of best fit indicates a perfect fit with normality. Any deviations from that line indicate some form of abnormality. Your aim is to have a no deviation (rarely happens) or a gentle s shape around the line without too greater deviation.

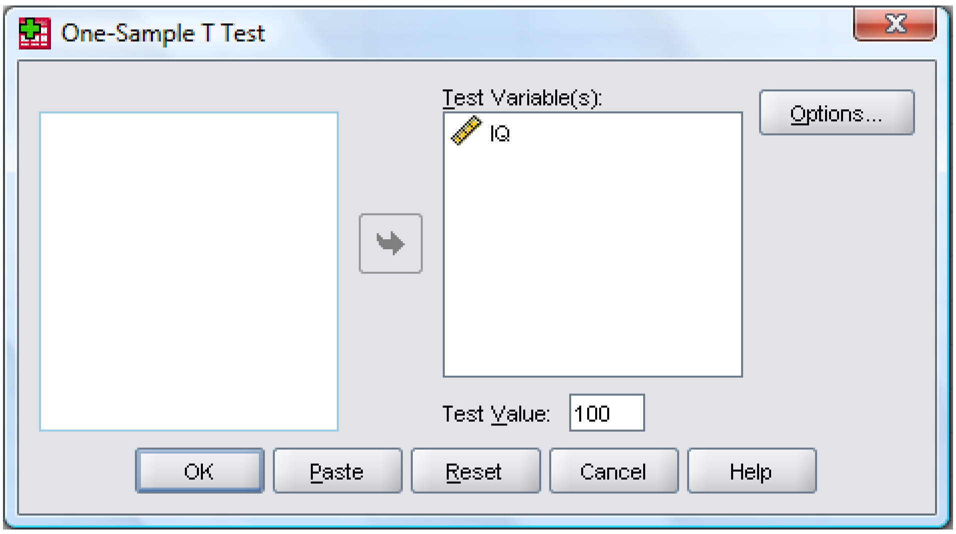
One Sample *t* – test

We know that the population IQ is 100 with a SD of 15 but is the university population significantly different in IQ? In order to assess this you would run a one sample T Test comparing your samples IQ score against the 'norm'

Analyze > compare means > one sample T Test



Place the variable of interest into the 'Test Variable' box. Make sure you state the 'Test Value' as well (in this case 100)



Below is the output from the one sample T Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
IQ	30	115.63	11.868	2.167

One-Sample Test

	Test Value = 100					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
IQ	7.215	29	.000	15.633	11.20	20.07

Writing up: One sample T Test indicate that a student samples IQ is significantly higher than the ‘norm’: $t(29) = 7.215$, $p < 0.001$, 95% CI = 11.20 > 20.07.

Independent Samples *t* – tests

	Condition	Recall	var
1	Noise	5	
2	Noise	10	
3	Noise	6	
4	Noise	6	
5	Noise	7	
6	Noise	3	
7	Noise	6	
8	Noise	9	
9	Noise	5	
10	Noise	10	
11	Noise	11	
12	Noise	9	
13	No Noise	15	
14	No Noise	9	
15	No Noise	16	
16	No Noise	15	
17	No Noise	16	
18	No Noise	18	
19	No Noise	17	
20	No Noise	13	
21	No Noise	11	
22	No Noise	12	
23	No Noise	13	
24	No Noise	11	

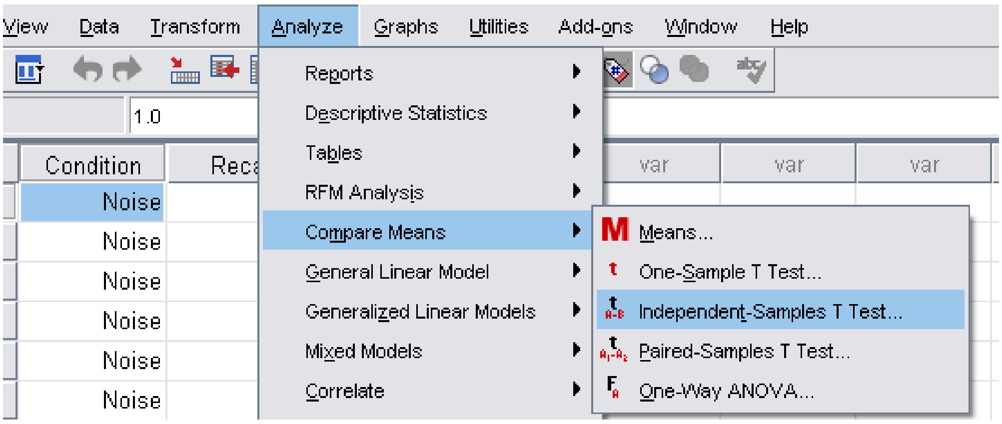
For running a between subjects t test your data should look like this.

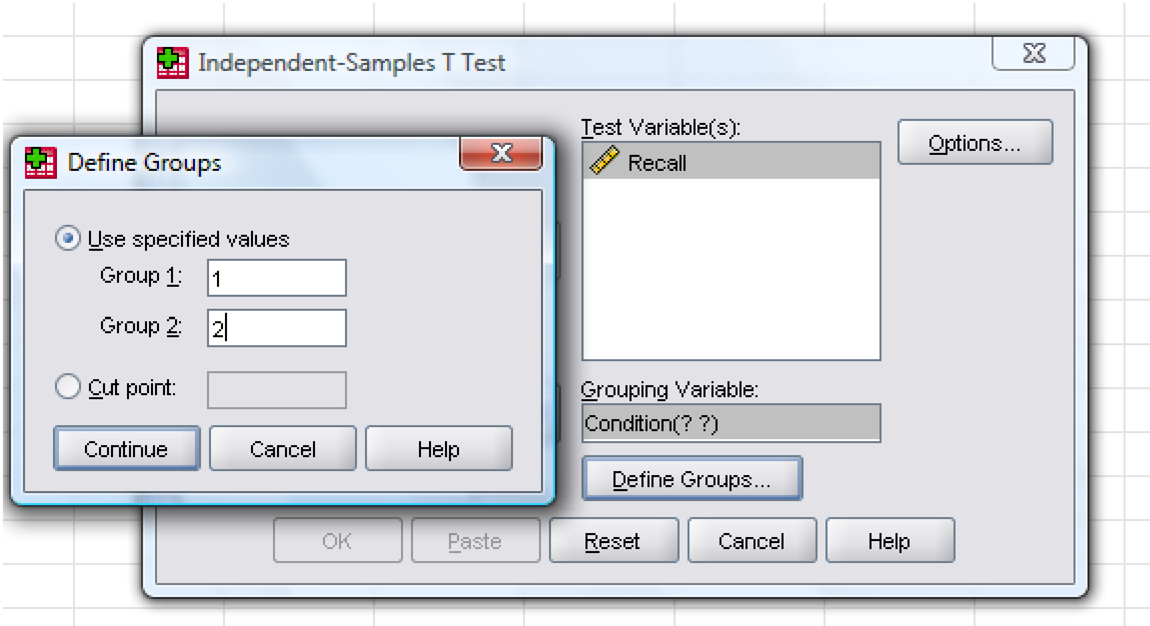
The IV is one column and the levels are given values

The DV makes up the second column and remains continuous.

Running the Independent Samples T Test

Analyze > compare means > independent samples T tests





Place the IV in the ‘Grouping Variable’. Select ‘Define Groups’ – specify the values of the IVs.

Select Continue and then select OK to run the analysis.

Below you can find the output for the independent samples T Test.

Group Statistics

	Condition	N	Mean	Std. Deviation	Std. Error Mean
Recall	Noise	12	7.25	2.491	.719
	No Noise	12	13.83	2.758	.796

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Recall	Equal variances assumed	.177	.678	-6.137	22	.000	-6.583	1.073	-8.808	-4.358
	Equal variances not assumed			-6.137	21.776	.000	-6.583	1.073	-8.809	-4.357

Levenes test for equality of variances assess the assumption of homogeneity of variances. If this is significant you will need to read the bottom line.

In this case it is not so I will go

The bits needed for a write up

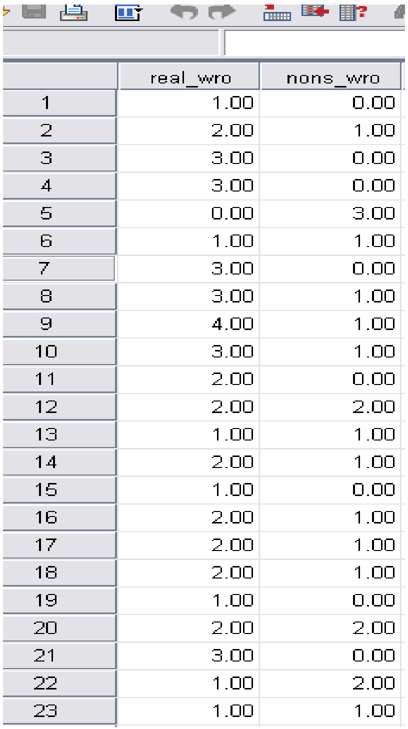
95% CI of the difference between the 2 means. If a 0 appeared amongst these figures you would not have a significant result. Some journal articles are now insisting on having these figures reported. If you are reporting in a table you could add these for some brownie points

Writing up:

Independent Samples T Test indicated a significant difference in the Noise and No Noise conditions on recall levels: t(22) = -6.137, p < 0.001, 95% CI -8.808 > -4.358.

Repeated Measures *t* test

In this study the IV is real or nonsense words and the DV is the recognition of said symbols.



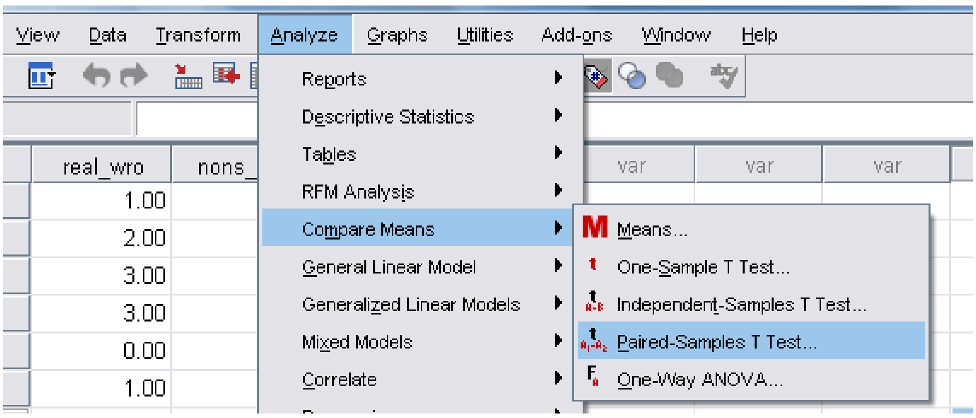
	real_wro	nons_wro
1	1.00	0.00
2	2.00	1.00
3	3.00	0.00
4	3.00	0.00
5	0.00	3.00
6	1.00	1.00
7	3.00	0.00
8	3.00	1.00
9	4.00	1.00
10	3.00	1.00
11	2.00	0.00
12	2.00	2.00
13	1.00	1.00
14	2.00	1.00
15	1.00	0.00
16	2.00	1.00
17	2.00	1.00
18	2.00	1.00
19	1.00	0.00
20	2.00	2.00
21	3.00	0.00
22	1.00	2.00
23	1.00	1.00

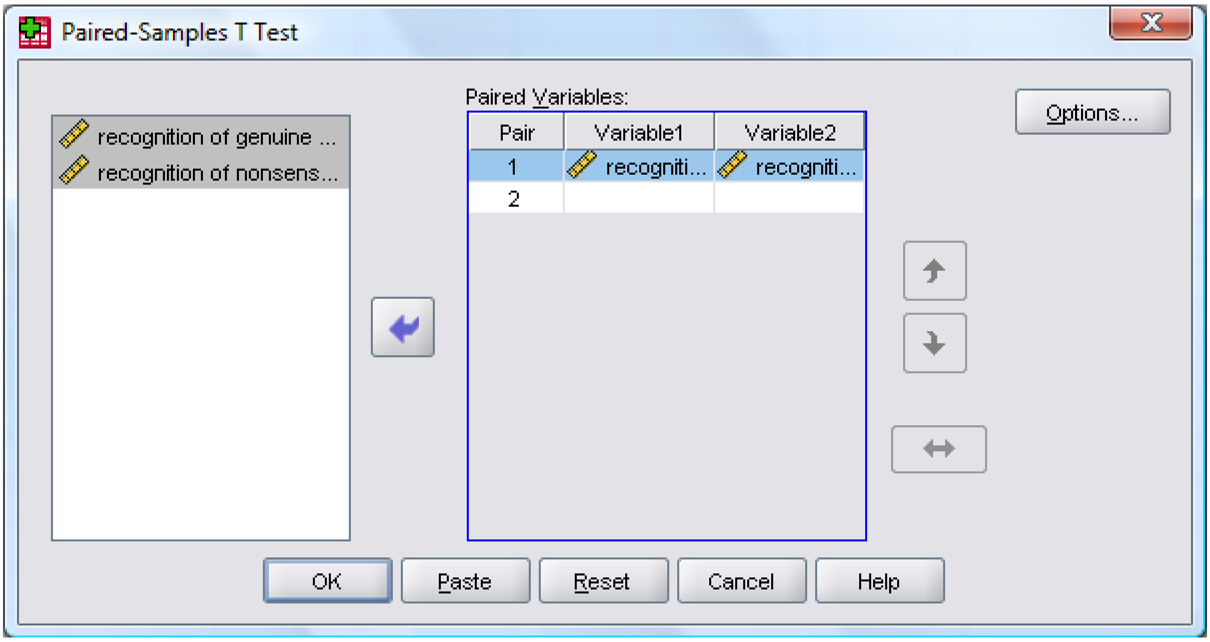
For running a repeated measures t test your data should look like this.

The two levels of the IV are spread across two columns

The DV appears within the rows – in this case recall of words.

Analyze > Compare Means > Paired –Samples T Test





Select the two levels of the IV and place them over into the Paired Variables Box by clicking the middle arrow. Then click OK to get the output of the results.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	recognition of genuine characters not seen before	1.8667	60	.94719	.12228
	recognition of nonsense characters not seen before	1.1500	60	.89868	.11602

Descriptive statistics

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	recognition of genuine characters not seen before & recognition of nonsense characters not seen before	60	-.275	.034

You do not need to report the Paired Samples Correlations... this just tells you how much the two are related to one another

Paired Samples Test

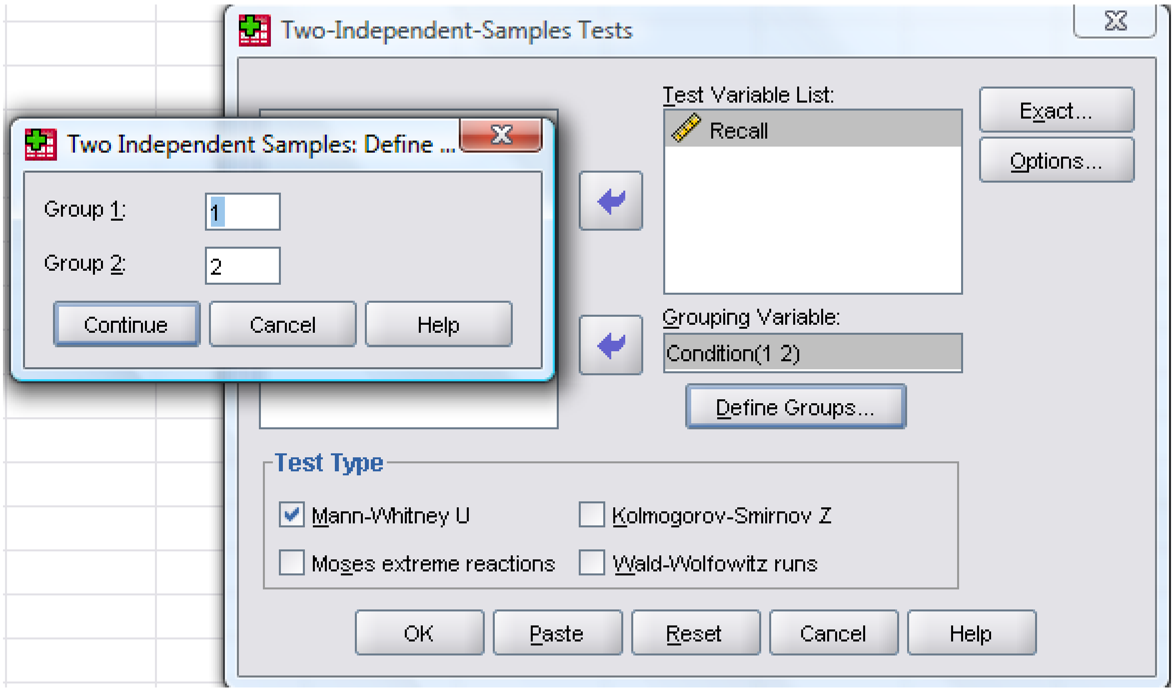
		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	recognition of genuine characters not seen before - recognition of nonsense characters not seen before	.71667	1.47397	.19029	.33590	1.09743	3.766	59	.000

Writing up: A paired samples T Test indicated a significant difference in recognition of genuine characters and recognition of nonsense words: $t(59) = 3.766$, $p < 0.001$, 95%CI 0.34 > 1.10.

Non Parametric Alternatives

Independent sample T Test > Mann Whitney U Test

Analyze > non parametric tests> two independent samples



Mann-Whitney Test

Ranks

	Condition	N	Mean Rank	Sum of Ranks
Recall	Noise	12	6.92	83.00
	No Noise	12	18.08	217.00
	Total	24		

Test Statistics[b]

	Recall
Mann-Whitney U	5.000
Wilcoxon W	83.000
Z	-3.883
Asymp. Sig. (2-tailed)	.000
Exact Sig. [2*(1-tailed Sig.)]	.000[a]

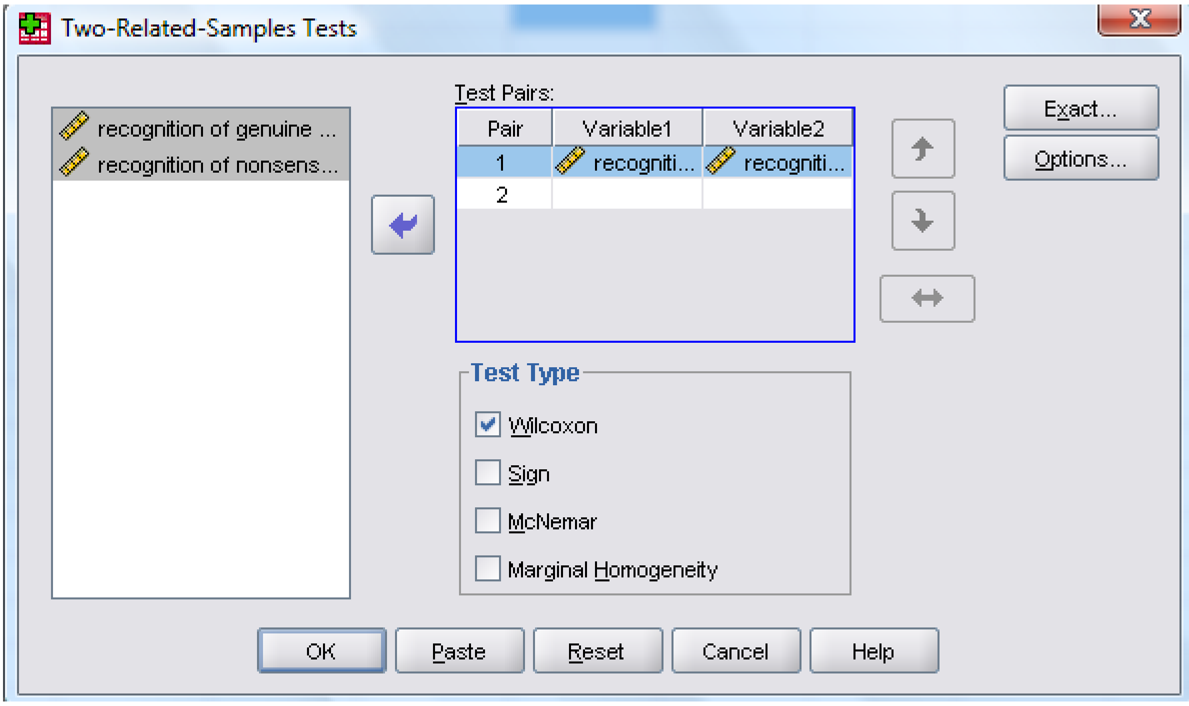
a. Not corrected for ties.

b. Grouping Variable: Condition

Write up: Mann-Whitney U Test indicated a significant difference between Noise and No Noise conditions: $z = -3.883$, $p < 0.001$, $N = 24$

Repeated Measures T Test > Wilcoxen

Analyze > Non Parametric Tests > two related samples



Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
recognition of nonsense characters not seen before - recognition of genuine characters not seen before	Negative Ranks	32[a]	24.48	783.50
	Positive Ranks	12[b]	17.21	206.50
	Ties	16[c]		
	Total	60		

a. recognition of nonsense characters not seen before < recognition of genuine characters not seen before

b. recognition of nonsense characters not seen before > recognition of genuine characters not seen before

c. recognition of nonsense characters not seen before = recognition of genuine characters not seen before

Test Statistics[b]

	recognition of nonsense characters not seen before - recognition of genuine characters not seen before
Z	-3.438[a]
Asymp. Sig. (2-tailed)	.001

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

Write up: Wilcoxen test indicated a significant difference between the recognition of genuine and nonsense characters: z = -3.438, p = 0.001, N = 60