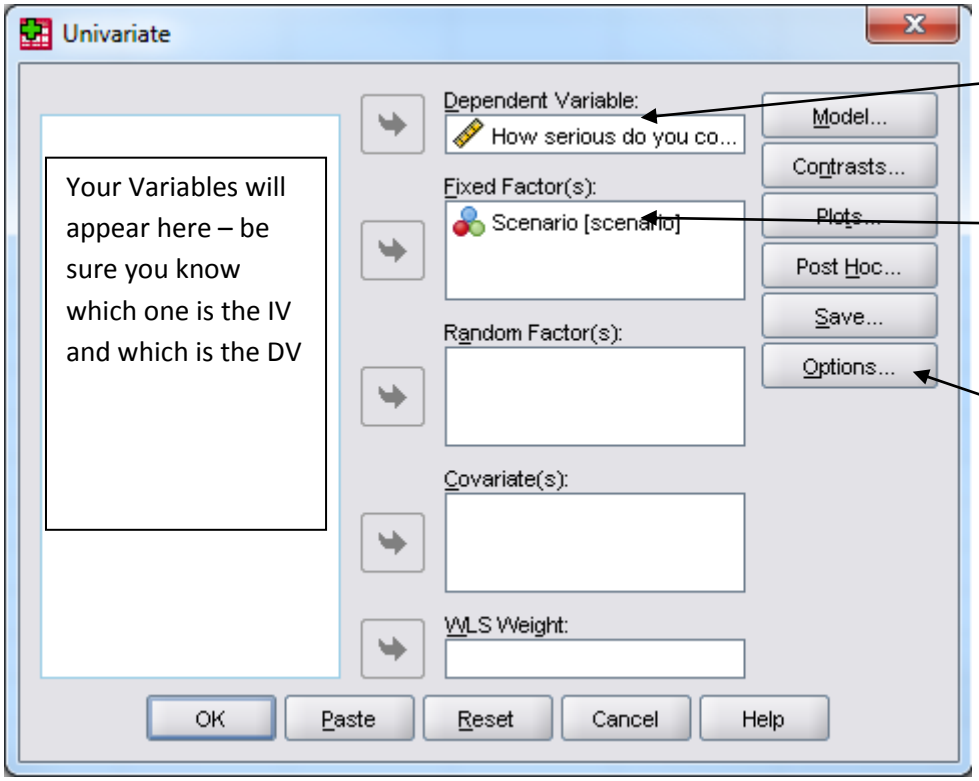
One Way ANOVA (Between Measures)

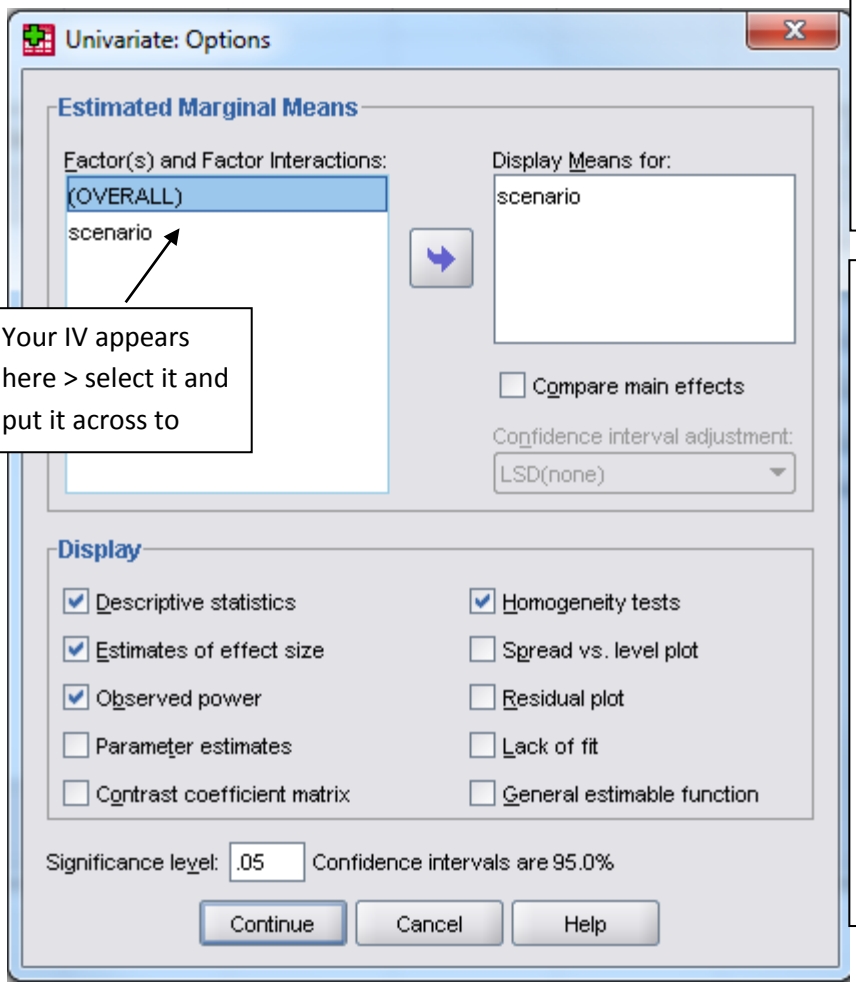
This is a data set from one of my dissertation students last year drastically reduced. This domestic violence data set focuses on attitudes towards domestic violence taking into consideration the gender of the perpetrator and that of the victim. The study used a vignette whereby the gender of the perpetrator and the victim were changed. Everything else remained identical. Each level of the scenario is identified by two sets of gender therefore male/male = male perpetrator and male victim (i.e. a homosexual violence scenario). The IV was therefore Scenario with four levels (Male/ Male, Male/ Female, Female/ Male and Female/ Female). The DV was one question – "how seriously do you rate this case?" The higher number equates to higher levels of perceived seriousness.

	scenario	DM1
1	1.00	7.00
2	1.00	3.00
3	1.00	7.00
4	1.00	6.00
5	1.00	6.00
6	1.00	7.00
7	1.00	7.00
8	1.00	6.00
9	1.00	5.00
10	1.00	6.00
11	1.00	7.00
12	1.00	6.00
13	1.00	7.00
14	1.00	7.00
15	1.00	6.00
16	1.00	7.00
17	1.00	7.00

Between subjects data should look like this

Analyze > General Linear Model > Univariate



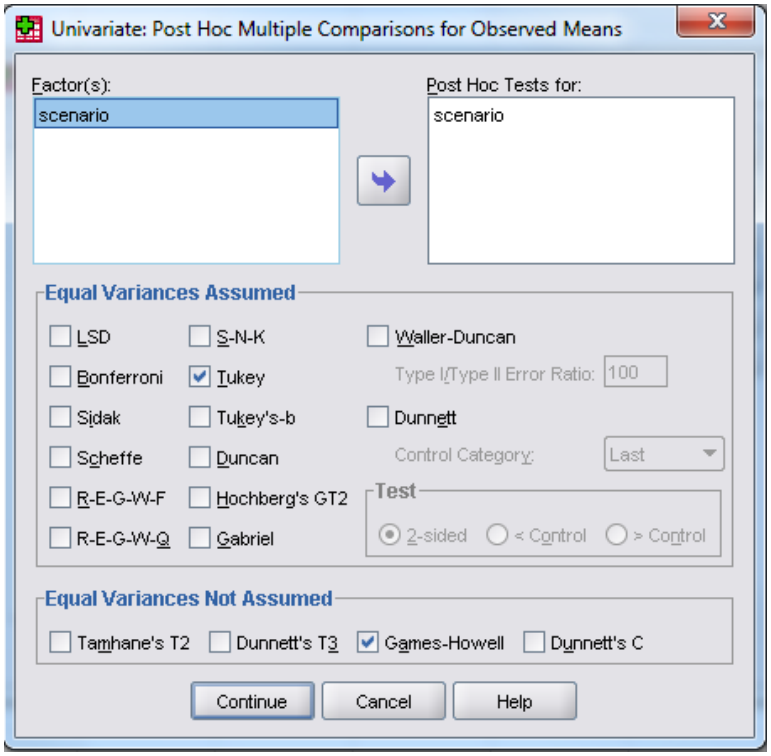


Your IV appears here > select it and put it across to

Within this section you need to tick:

- Descriptive Statistics
- Estimates of Effect Size
- Observed Power
- Homogeniety Tests

The Estimates of effect size should be reported if our factors turn out to be significant, though you should be aware that SPSS provides partial estimates of effect size (the proportion of factor plus error variance accounted for by the factor) and that may not be what you want. If you want full estimates of eta squared (that is the proportion of total variance accounted for by the factor), you will need to divide SS(factor) by SS(total) yourself. The Sums of Squares (SS) are given in SPSS ANOVA summary tables. The Observed power will be potentially useful for planning future experiments and should be reported in order to facilitate any future meta-analyses



Click on Post Hoc on the main screen to get to this box.

You will need to select the most appropriate test here for your study.

Once you have selected your post hoc tests and descriptive information then click OK to bring up your output.

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
Scenario	1.00	Man/woman	43
	2.00	Woman/man	30
	3.00	Man/man	51
	4.00	Woman/woman	33

This box just informs you of the different levels of the IV. It will also tells you how many participants you have in each group.

I have scored it out as it doesn't provide you with any NEW information

Descriptive Statistics

Dependent Variable:How serious do you consider this case to be?

Scenario	Mean	Std. Deviation	N
Man/woman	6.4884	.79798	43
Woman/man	5.8333	1.08543	30
Man/man	6.4314	.98499	51
Woman/woman	5.6061	1.47774	33
Total	6.1592	1.13516	157

This box provides you with the descriptive statistics that you will need to report. Note that this box provides you with mean and SD but no 95%CI.

The estimated marginal means box provides you with the mean, SE (remember the difference) and the 95%CI (which you will need)

Levene's Test of Equality of Error Variances[a]

Dependent Variable:How serious do you consider this case to be?

F	df1	df2	Sig.
4.303	3	153	.006

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + scenario

This is your Homogeneity of Variances test.

Remember this tests the differences in variances across the three different groups. If there are significant differences in variation this would be below 0.05 therefore indicating a violation of the assumption. In this case there is a violation of homogeneity of variances.

Recall that in large sample sizes this test becomes unreliable so you may want to use the following rule of thumb: if the biggest variance is 3x bigger than the smallest variance then you have a problem.

Tests of Between-Subjects Effects

Dependent Variable:How serious do you consider this case to be?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	21.720[a]	3	7.240	6.178	.001	.108	18.534	.960
Intercept	5571.526	1	5571.526	4754.301	.000	.969	4754.301	1.000
scenario	21.720	3	7.240	6.178	.001	.108	18.534	.960
Error	179.299	153	1.172					
Total	6157.000	157						
Corrected Total	201.019	156						

a. R Squared = .108 (Adjusted R Squared = .091)

b. Computed using alpha = .05

The above box is the main table needed for your write up...

You should be reading the line that contains the name of the IV only with addition of the error df for write up.

One way Between Measures ANOVA indicated a significant difference between two or more groups: F(3, 153) = 6.178, p < 0.001, $p\eta^2$ = 0.108, observed power = 0.960

Remember Partial eta squared or $p\eta^2$ is the effect size calculation for ANOVA
Observed Power may be useful for future research so report it.

Estimated Marginal Means

Scenario

Dependent Variable:How serious do you consider this case to be?

Scenario	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Man/woman	6.488	.165	6.162	6.815
Woman/man	5.833	.198	5.443	6.224
Man/man	6.431	.152	6.132	6.731
Woman/woman	5.606	.188	5.234	5.978

The 95% CI within this table should be reported with your descriptive statistics. Very briefly they indicate where the true population mean is.

Post Hoc Tests

Scenario

Remember an ANOVA only tells you whether there is a difference somewhere between the groups but it doesn't tell you where. You use your post hoc tests to highlight where these differences are.

Multiple Comparisons

Dependent Variable:How serious do you consider this case to be?

	(I) Scenario	(J) Scenario	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Man/woman	Woman/man	.6550	.25752	.057	-.0139	1.3239
		Man/man	.0570	.22412	.994	-.5252	.6392
		Woman/woman	.8823*	.25053	.003	.2316	1.5331
	Woman/man	Man/woman	-.6550	.25752	.057	-1.3239	.0139
		Man/man	-.5980	.24908	.081	-1.2450	.0489
		Woman/woman	.2273	.27308	.839	-.4821	.9366
	Man/man	Man/woman	-.0570	.22412	.994	-.6392	.5252
		Woman/man	.5980	.24908	.081	-.0489	1.2450
		Woman/woman	.8253*	.24185	.005	.1971	1.4535
	Woman/woman	Man/woman	-.8823*	.25053	.003	-1.5331	-.2316
		Woman/man	-.2273	.27308	.839	-.9366	.4821
		Man/man	-.8253*	.24185	.005	-1.4535	-.1971
Games-Howell	Man/woman	Woman/man	.6550*	.23255	.034	.0370	1.2730
		Man/man	.0570	.18393	.990	-.4243	.5383
		Woman/woman	.8823*	.28457	.017	.1239	1.6407
	Woman/man	Man/woman	-.6550*	.23255	.034	-1.2730	-.0370
		Man/man	-.5980	.24144	.075	-1.2373	.0412
		Woman/woman	.2273	.32472	.897	-.6314	1.0860
	Man/man	Man/woman	-.0570	.18393	.990	-.5383	.4243
		Woman/man	.5980	.24144	.075	-.0412	1.2373
		Woman/woman	.8253*	.29188	.033	.0498	1.6008
	Woman/woman	Man/woman	-.8823*	.28457	.017	-1.6407	-.1239
		Woman/man	-.2273	.32472	.897	-1.0860	.6314
		Man/man	-.8253*	.29188	.033	-1.6008	-.0498

Based on observed means.
The error term is Mean Square(Error) = 1.172.

*. The mean difference is significant at the .05 level.

You need to write up ALL the results even if they are insignificant.

Post Hoc Analyses indicated these differences to be between Man/ Woman and Woman/ Woman (p = 0.003, 95%CI [0.232, 1.533]) and Man/ Man and Woman/ Woman (p = 0.005, 95%CI [0.197, 1.454]. The difference between Man/Woman and Woman/Man was approaching significance (p = 0.057, 95%CI [-0.014, 1.324]. There was no significant difference between Man/Woman and Man/Man (p = .994, 95%CI [-0.525, 0.639]

Homogeneous Subsets

How serious do you consider this case to be?

	Scenario	N	Subset 1	Subset 2	Subset 3
Tukey HSD[a,,b,,c]	Woman/woman	33	5.6061		
	Woman/man	30	5.8333	5.8333	
	Man/man	51		6.4314	6.4314
	Man/woman	43			6.4884
	Sig.		.800	.083	.996

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = 1.172.

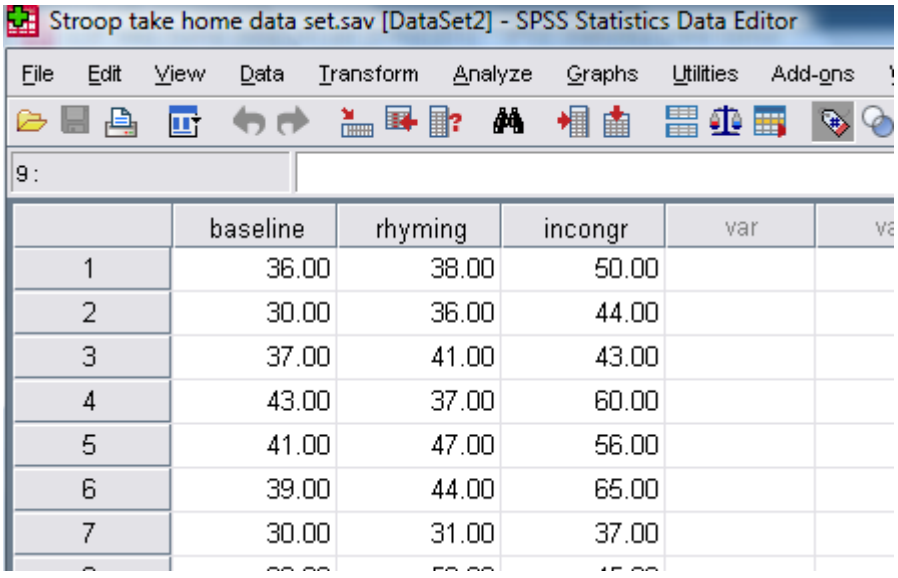
a. Uses Harmonic Mean Sample Size = 37.559.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

This box simply indicates if some of your levels of the IV can be grouped together as one. In this case, due to all being significantly different from one another, there are no homogenous subsets.

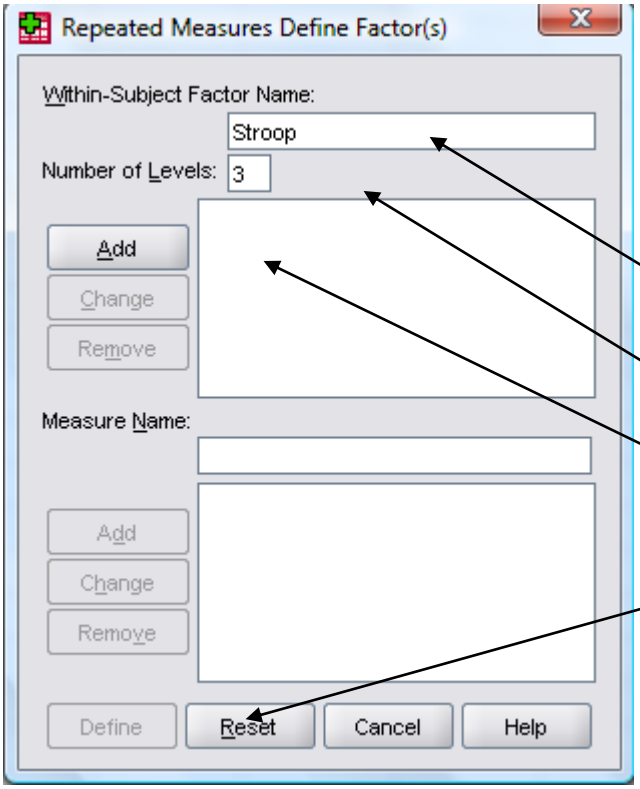
One Way Repeated Measures ANOVA



	baseline	rhyming	incongr	var
1	36.00	38.00	50.00	
2	30.00	36.00	44.00	
3	37.00	41.00	43.00	
4	43.00	37.00	60.00	
5	41.00	47.00	56.00	
6	39.00	44.00	65.00	
7	30.00	31.00	37.00	

Within subjects data set should look like this.

Analyze > General Linear Model > Repeated Measures



This is the first box that will appear.

This is where you tell the computer how many levels of your IV you have.

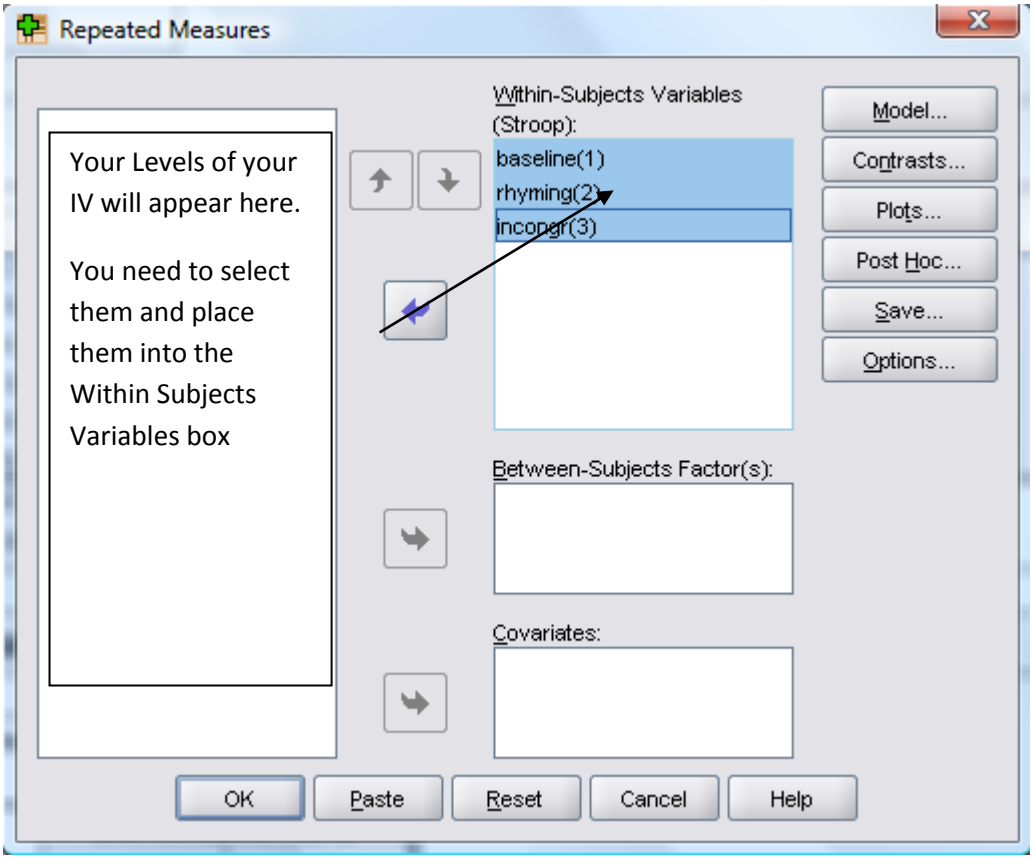
In this case we have 3

Therefore we name the Within Subjects Factor

And state the Number of Levels

Click Add.

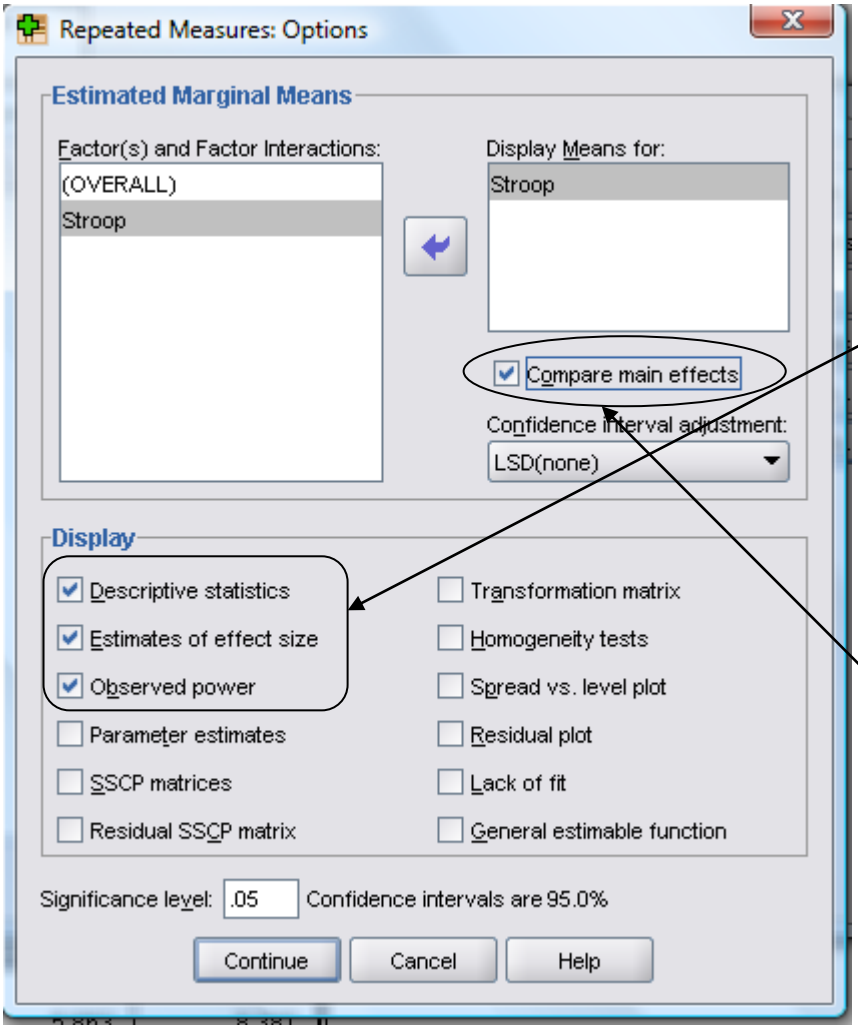
Click Define once it is highlighted



Your Levels of your IV will appear here.

You need to select them and place them into the Within Subjects Variables box

Once you have done this you will need to click on Options.



This is the Options Box

You will need to select your IV and place it into Display Means for

Tick on the usual

- Descriptive statistics
- Estimates of effect size
- Observed power

You do not need to tick the 'Homogeneity Tests' box as this only relates to between subjects IV.

For post hoc tests in repeated measures you select 'Compare Main Effects' and select the appropriate measure from the dropdown menu.

See information for between subject ANOVA to learn about Partial eta and observed power

Click Continue once finished

This is what your output should look like... be warned that there is a lot of 'useless' information in this output.

General Linear Model

Within-Subjects Factors

Measure:MEASURE_1

Stroop	Dependent Variable
1	baseline
2	rhyming
3	incongr

This box only informs you of the levels of your IV.

You do not need to report this box

Descriptive Statistics

	Mean	Std. Deviation	N
baseline	41.1296	12.45858	108
rhyming	48.9722	12.08031	108
incongr	66.2222	16.17159	108

This box provides you with the descriptive statistics of your data that you will need to report.

Remember you can get the 95% CI from the marginal means further down in the output

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Observed Power[b]
Stroop	Pillai's Trace	.681	113.191[a]	2.000	106.000	.000	.681	1.000
	Wilks' Lambda	.319	113.191[a]	2.000	106.000	.000	.681	1.000
	Hotelling's Trace	2.136	113.191[a]	2.000	106.000	.000	.681	1.000
	Roy's Largest Root	2.136	113.191[a]	2.000	106.000	.000	.681	1.000

The above box provides the multivariate tests. For ANOVA you do not need to concern yourself with this.

Mauchly's Test of Sphericity[b]

Measure:MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Stroop	.705	37.013	2	.000	.772	.782	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept

Within Subjects Design: Stroop

Mauchly's test of sphericity: we are now aware that the Homogeneity of variances is important within between group statistics many people assume this isn't an issue in repeated measures. This is not the case therefore the assumption of sphericity can be likened to the assumption of homogeneity of variance. Sphericity is more a general condition of compound symmetry which holds true when both variances across conditions are equal and the covariance's between pairs of conditions are equal. Sphericity thus refers to the equality of variances of the differences between treatment levels. So if you take each pair of treatment levels, and calculate the differences between each pair of scores, then it is necessary that these differences have equal variances. **You need to have at least 3 conditions for sphericity to e an issue.**

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power[a]
Stroop	Sphericity Assumed	35593.521	2	17796.760	170.888	.000	.615	1.000
	Greenhouse-Geisser	35593.521	1.545	23042.091	170.888	.000	.615	1.000
	Huynh-Feldt	35593.521	1.563	22772.168	170.888	.000	.615	1.000
	Lower-bound	35593.521	1.000	35593.521	170.888	.000	.615	1.000
Error(Stroop)	Sphericity Assumed	22286.530	214	104.143				
	Greenhouse-Geisser	22286.530	165.285	134.837				
	Huynh-Feldt	22286.530	167.244	133.258				
	Lower-bound	22286.530	107.000	208.285				

The highlighted aspects are those that need writing up.

Why use the Greenhouse-Geisser? The Greenhouse-Geisser is used when Sphericity cannot be assumed. However, authors often recommend that this is used all the time. When reporting the Greenhouse-Geiser df round the figures up.

Source	Stroop	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power[a]
Stroop	Linear	34000.543	1	34000.543	228.356	.000	.681	1.000
	Quadratic	1592.978	1	1592.978	26.821	.000	.200	.999
Error(Stroop)	Linear	15931.540	107	148.893				
	Quadratic	6354.990	107	59.392				

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power[a]
Intercept	879739.447	1	879739.447	2482.440	.000	.959	1.000
Error	37919.194	107	354.385				

The above two boxes can be ignored – you do not need to write them up or understand what they tell you

Estimated Marginal Means – these provide you with the 95% CI of the mean that you will need for writing up.

Stroop

Estimates

Measure:MEASURE_1

Stroop	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	41.130	1.199	38.753	43.506
2	48.972	1.162	46.668	51.277
3	66.222	1.556	63.137	69.307

Pairwise Comparisons

Measure:MEASURE_1

(I) Stroop	(J) Stroop	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-7.843*	.980	.000	-9.785	-5.900
	3	-25.093*	1.661	.000	-28.384	-21.801
2	1	7.843*	.980	.000	5.900	9.785
	3	-17.250*	1.438	.000	-20.101	-14.399
3	1	25.093*	1.661	.000	21.801	28.384
	2	17.250*	1.438	.000	14.399	20.101

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

When writing up the post hoc tests ensure you write up ALL tests including those that are not significant. You will also need to report the 95% CI. Notice that these are now the 95% CI of the DIFFERENCE. If these intervals include or go through 0 then this will be reflected in the significance value.

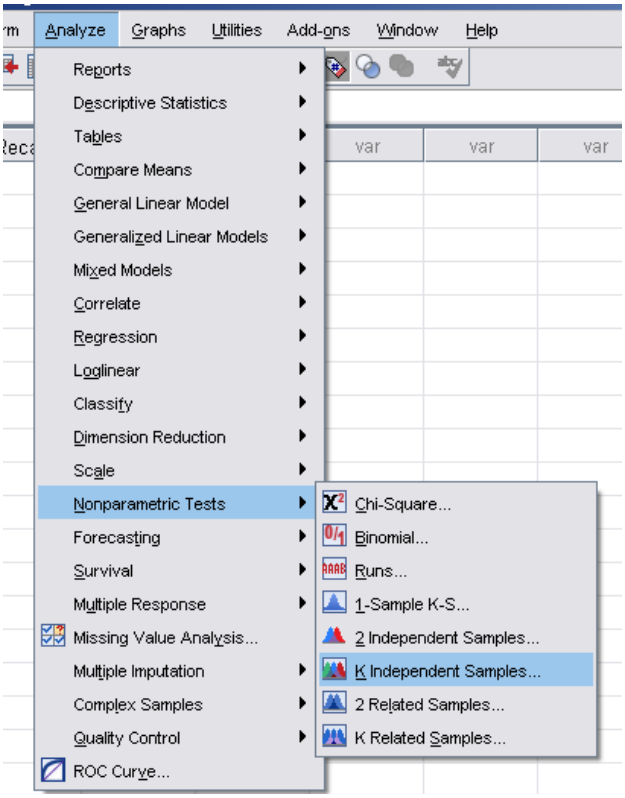
Another table of the multivariate tests which is not needed for the write up.

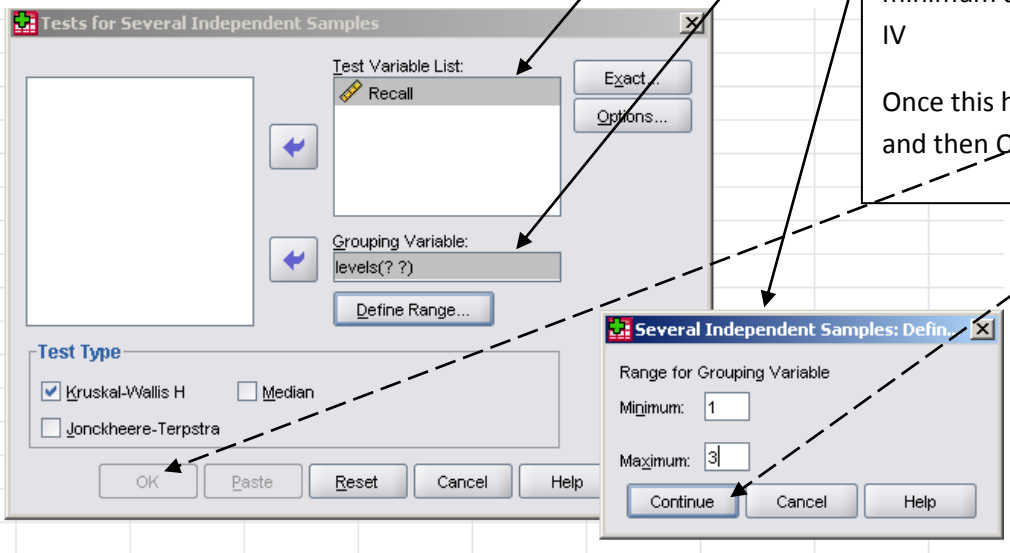
	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Observed Power[b]
Pillai's trace	.681	113.191[a]	2.000	106.000	.000	.681	1.000
Wilks' lambda	.319	113.191[a]	2.000	106.000	.000	.681	1.000
Hotelling's trace	2.136	113.191[a]	2.000	106.000	.000	.681	1.000
Roy's largest root	2.136	113.191[a]	2.000	106.000	.000	.681	1.000

Write up: One way repeated ANOVA indicated a significant difference in stroop tasks: $F(2, 165) = 170.89$, $p < 0.001$, $p\eta^2 = 0.615$, observed power = 1.00. Post hoc analysis indicated these differences to be between conditions 1 & 2 ($p < 0.001$, 95%CI [-9.785, -5.900]), conditions 1 & 3 ($p < 0.001$, 95% CI [-28.384, -21.801]) and conditions 2 & 3 ($p < 0.001$, 95%CI [-20.101, -14.399]).

Non- Parametric Equivalents: Between Measures – Kruskal Wallis

Analyze > K Independent Samples





Place your DV into the 'Test Variable List'

Place your IV into the 'Grouping Variable' and then select 'Define Groups'

This will bring up a separate dialogue box. Here you need to state the minimum and maximum range for the IV

Once this has been done click continue and then OK

NPar Tests

```
[DataSet5] E:\northampton PG stats teaching\LOP (Between) ANOVA.sav
```

Kruskal-Wallis Test

Ranks

	levels	N	Mean Rank
Recall	capitals	48	45.60
	rhyme	48	67.82
	category	48	104.07
	Total	144	

Test Statistics[a,b]

	Recall
Chi-Square	48.303
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: levels

This is the output

The first box give you the rankings of your results. I would recommend reporting median and IQR rather than mean ranks. However these mean ranks show that category has the highest recall rate (as the rank is higher)

The second box is the inferential statistics that need to write up

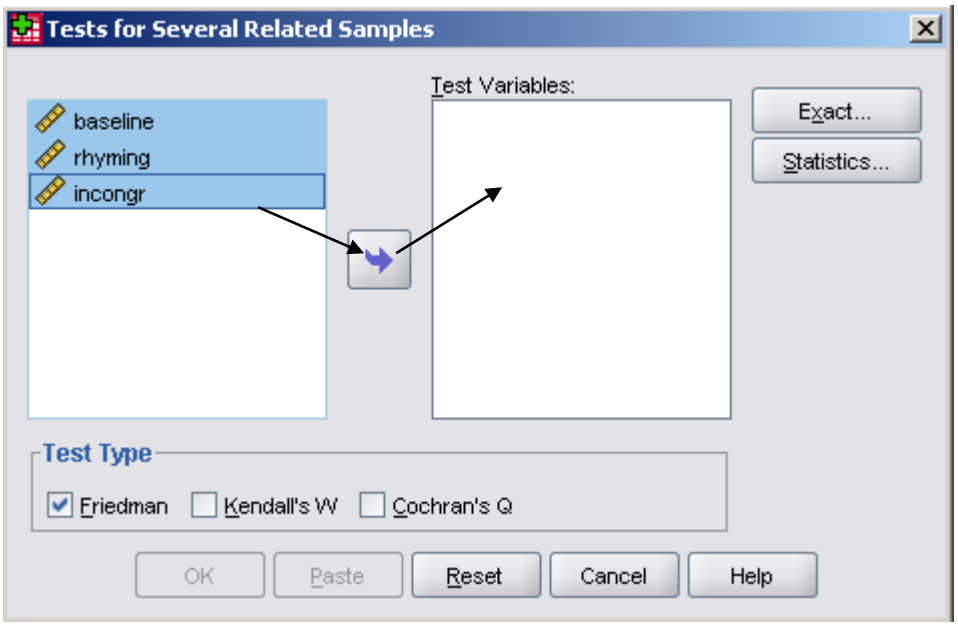
Write up for this case:

Kruskal Wallis analysis indicated a significant effect between groups on level of recall: χ^2 (2) 48.303, $p < 0.001$

If you would like to follow this up run non-parametric Mann-Whitney test correcting for number of tests used.

Non- Parametric Equivalents: Within Measures – Freedman

Analyze> Non Parametric tests> K Related Tests



Select all the levels of the variables of interest and place them over into 'test variables'

Then select OK

NPar Tests

```
[DataSet6] E:\northampton PG stats teaching\Stroop (within) ANOVA.sav
```

Friedman Test

Ranks

	Mean Rank
baseline	1.23
rhyming	1.91
incongr	2.86

Test Statistics[a]

N	108
Chi-Square	145.814
df	2
Asymp. Sig.	.000

a. Friedman Test

Mean ranks as before. This indicates that the incongruent group has the longest time to read through the list.

Test statistics are needed for your write up

Thus for this study

Freedman test indicates a significant difference between the groups on speed of reading: χ^2 (2) 145.814, $p < 0.001$

If you would like to follow this up run non-parametric wilcoxen test correcting for number of tests used.