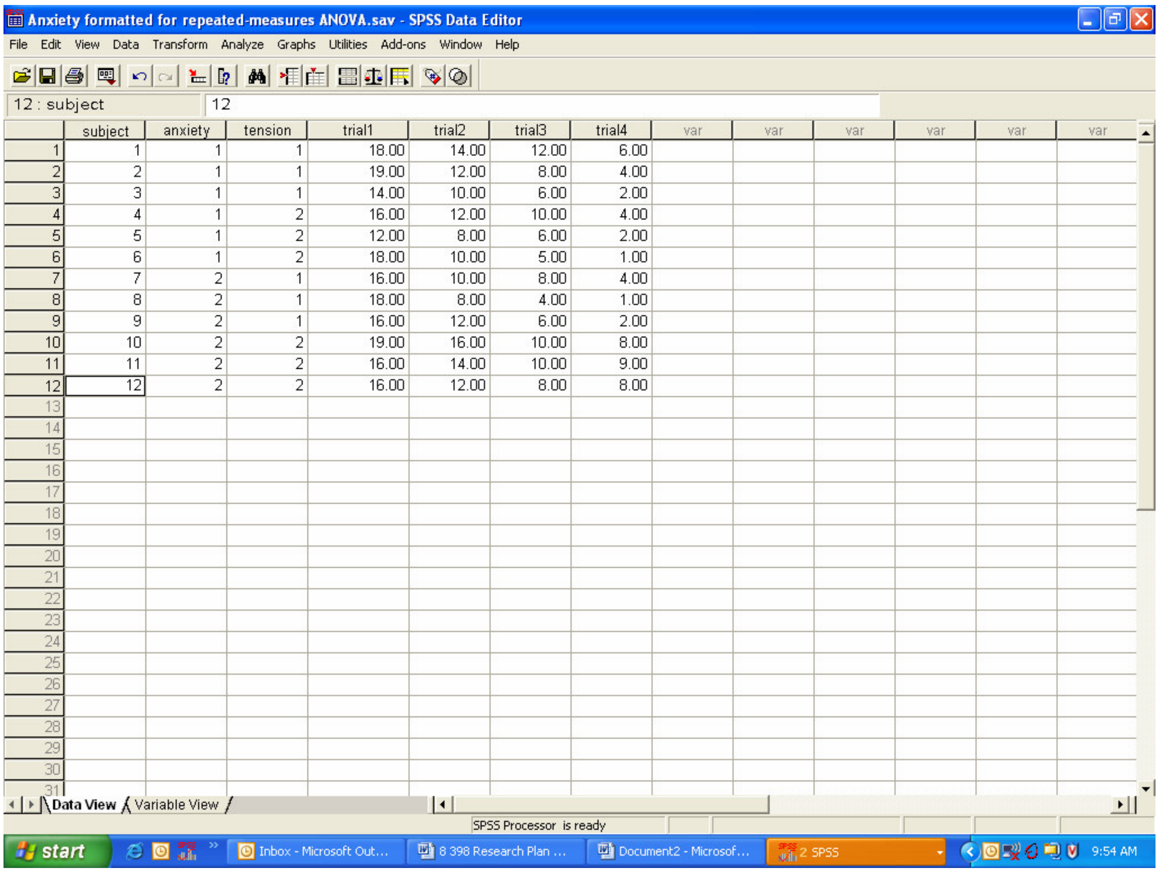
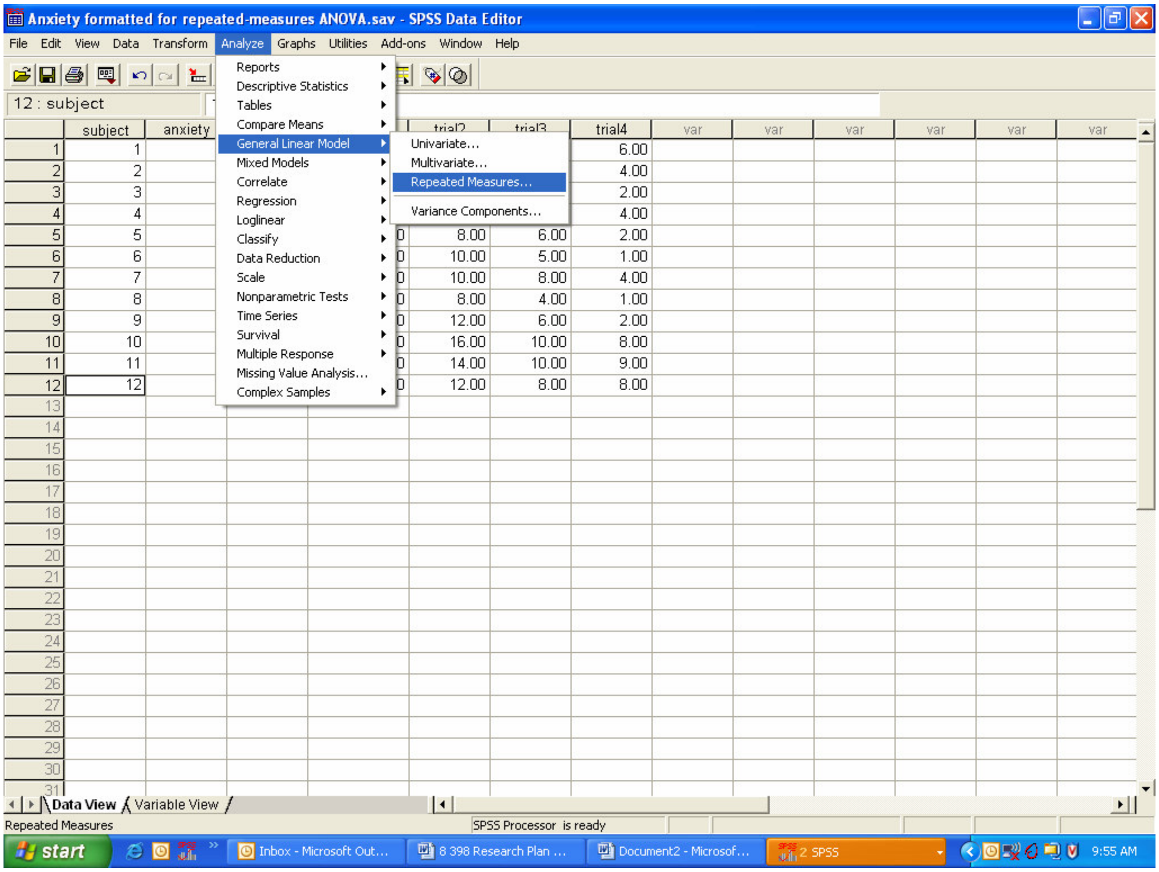
Repeated-Measures ANOVA in SPSS

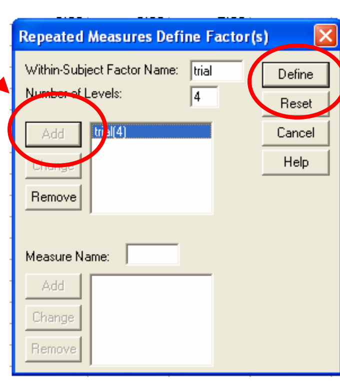
Correct data formatting for a repeated-measures ANOVA in SPSS involves having a *single line of data* for each participant, with the repeated measures entered as separate variables on that same line (in this example, they are called “trial1,” “trial2,” “trial3,” and “trial4”). Note that this is *different from* the setup for a Hierarchical Linear Model, where you would have a separate line of data for each observation and repeat the subject number on each separate line, with a variable named “trial” that goes from 1 to 4 for each participant, and a separate variable called “score” that just has the number in it that corresponds to that trial for that participant.



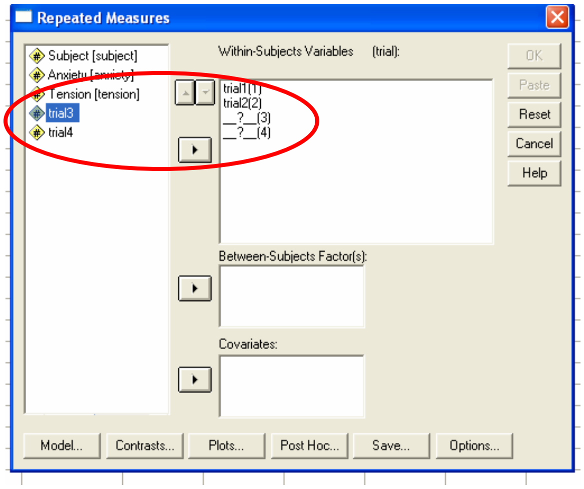
The procedure is found under “Analyze”/“General Linear Model”/“Repeated Measures.”



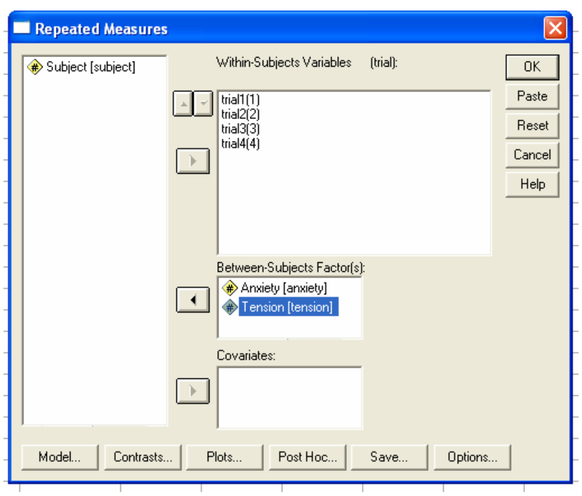
Use the first pop-up dialog box to define your repeated-measures factor – this is how you tell SPSS that the four different “trial” variables are really all a single person’s scores over time on one variable. Give your variable a name (like “trial”), and specify how many different levels it has (i.e., how many times the observation was repeated). Be sure to then hit the “Add” button so that the new variable appears on a list. You can create multiple repeated-measures variables at this step if you need to. Then click on “Define.”

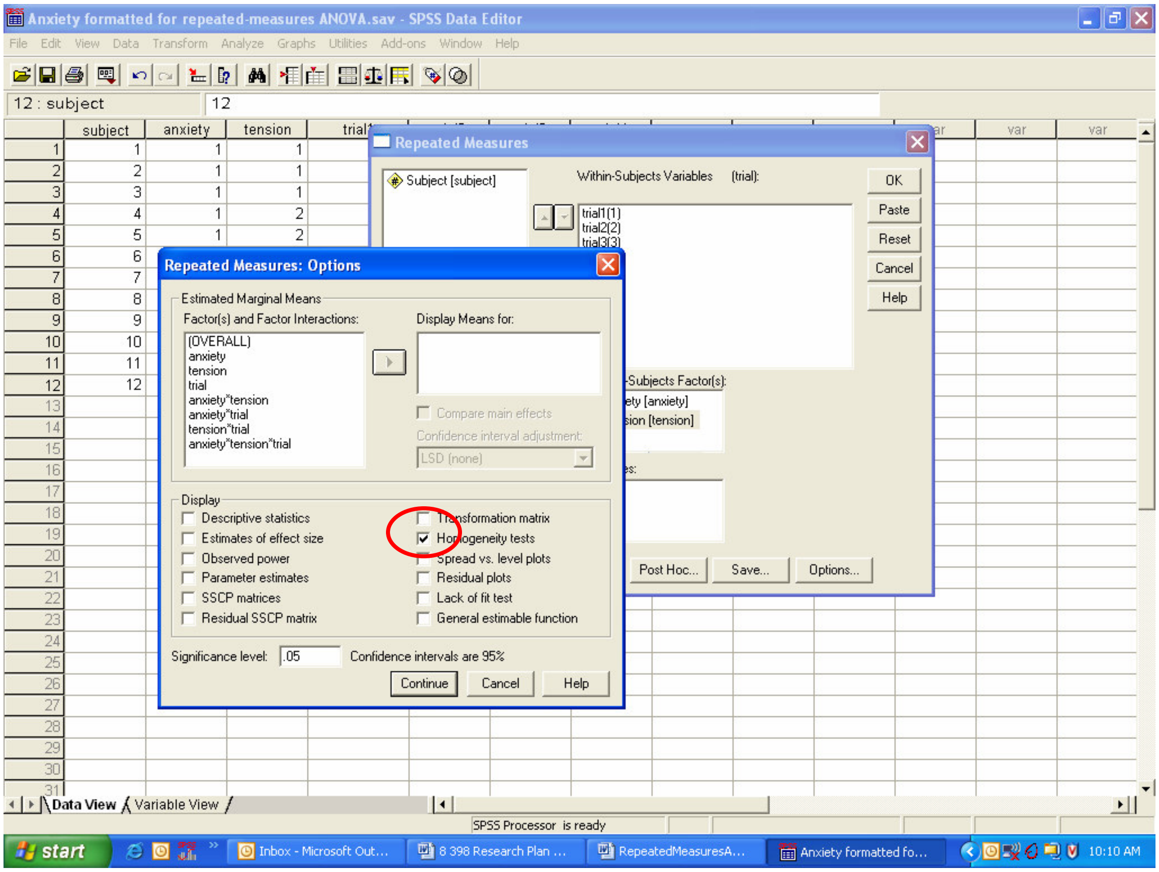


Select the columns in your dataset that represent the various levels of the repeated-measures variable, and use the arrow button to move them into the “blanks” in the right-hand column.



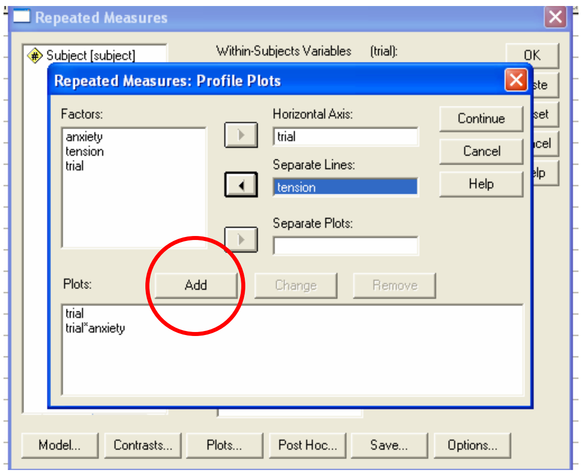
Next, specify your model. The within-subjects factors represent *scores on the DV at each trial*. Therefore, they will be treated as the dependent (criterion) variable for this analysis. For predictors, enter grouping variables (e.g., treatment vs. control) as “between-subjects factors,” and other interval/ratio-level predictors as “covariates.” In this example, participants’ anxiety level (“Anxiety”) and tension level (“Tension”) were both manipulated experimentally, so these are entered as grouping variables (“between-subjects factors”). Therefore, this is a 2 (two levels of anxiety) x 2 (two levels of tension) factorial repeated-measures ANOVA design.





Use the “Options” button to open this window, and click on the check-box to select “Homogeneity Tests.” This will allow you to test the *homogeneity of variance assumption* for the repeated-measures dependent variable. This window also shows you all the different interactions that will be tested as part of your analysis. If you *don’t* want all of these results, you can select just specific main effects and interactions by using the “Model” button in the main dialog window. Click “Continue” to go on.

One other type of output you might want is a graph showing the average change in scores on the dependent variable for individuals over time, maybe with the results sorted based on the experimental group the person is in. The "Plots" button on the main dialog box lets you do this. In this dialog box, the DV (repeated-measures variable) goes on the "horizontal axis" of the graph. If you just want to see that variable's change over time, just enter that for the horizontal axis and leave the other fields blank. If you want to see separate lines on the graph for people in different groups, put the grouping variable into the "separate lines" box. Remember to hit the "Add" button to add your graph to the list that SPSS will provide.



Now return to the main dialog box and hit "OK" to see the output.

SPSS uses a multivariate analysis to detect repeated-measures effects. This approach *assumes that there is some change from each time period to the next* on the repeated measure. If you expect more of a "stepwise" pattern in your results, it may be advisable to test only the pre-post difference in the repeated measure. Alternately, you may want to consider using a different procedure like ANCOVA or HLM.

Here are the results of the multivariate test:

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
trial	Pillai's Trace	.985	127.686[a]	3.000	6.000	.000
	Wilks' Lambda	.015	127.686[a]	3.000	6.000	.000
	Hotelling's Trace	63.843	127.686[a]	3.000	6.000	.000
	Roy's Largest Root	63.843	127.686[a]	3.000	6.000	.000
trial * anxiety	Pillai's Trace	.756	6.183[a]	3.000	6.000	.029
	Wilks' Lambda	.244	6.183[a]	3.000	6.000	.029
	Hotelling's Trace	3.091	6.183[a]	3.000	6.000	.029
	Roy's Largest Root	3.091	6.183[a]	3.000	6.000	.029
trial * tension	Pillai's Trace	.639	3.546[a]	3.000	6.000	.088
	Wilks' Lambda	.361	3.546[a]	3.000	6.000	.088
	Hotelling's Trace	1.773	3.546[a]	3.000	6.000	.088
	Roy's Largest Root	1.773	3.546[a]	3.000	6.000	.088
trial * anxiety * tension	Pillai's Trace	.672	4.099[a]	3.000	6.000	.067
	Wilks' Lambda	.328	4.099[a]	3.000	6.000	.067
	Hotelling's Trace	2.050	4.099[a]	3.000	6.000	.067
	Roy's Largest Root	2.050	4.099[a]	3.000	6.000	.067

a. Exact statistic

b.

Design: Intercept+anxiety+tension+anxiety * tension
Within Subjects Design: trial

The significant *p*-values show an effect of time (trial) on the dependent variable – a within-subjects effect reflected by the repeated measures. All four multivariate tests also show a significant interaction between trial and anxiety level, meaning that the level of anxiety induced in the participant had a significant effect on their performance over time. The effect of tension, by contrast, did not reach conventional levels of statistical significance. Because tension had a non-significant effect, it probably doesn't make sense to look at the results for the three-way interaction between tension, anxiety, and time.

Next is a test for the assumption of *sphericity*, which is important in multivariate tests.

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
trial	.187	11.254	5	.049	.536	.902	.333

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+anxiety+tension+anxiety * tension
Within Subjects Design: trial

In this case, the assumption of sphericity was *not met* (because the *p*-value of the test was significant, indicating a significant *difference* from the conditions under which the assumption holds true). Unfortunately, this means that we cannot rely on the multivariate tests examined above. The "epsilon" values on the right-hand side of this table are three different ways to calculate an appropriate adjustment to the degrees of freedom of the *F*-test. The next table shows revised results using each of these corrections. The Lower-Bound test is the most conservative; the Huynh-Feldt test is generally least conservative.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
trial	Sphericity Assumed	991.500	3	330.500	152.051	.000
	Greenhouse-Geisser	991.500	1.608	616.432	152.051	.000
	Huynh-Feldt	991.500	2.707	366.284	152.051	.000
	Lower-bound	991.500	1.000	991.500	152.051	.000
trial * anxiety	Sphericity Assumed	8.417	3	2.806	1.291	.300
	Greenhouse-Geisser	8.417	1.608	5.233	1.291	.300
	Huynh-Feldt	8.417	2.707	3.109	1.291	.301
	Lower-bound	8.417	1.000	8.417	1.291	.289
trial * tension	Sphericity Assumed	12.167	3	4.056	1.866	.162
	Greenhouse-Geisser	12.167	1.608	7.564	1.866	.197
	Huynh-Feldt	12.167	2.707	4.495	1.866	.169
	Lower-bound	12.167	1.000	12.167	1.866	.209
trial * anxiety * tension	Sphericity Assumed	12.750	3	4.250	1.955	.148
	Greenhouse-Geisser	12.750	1.608	7.927	1.955	.185
	Huynh-Feldt	12.750	2.707	4.710	1.955	.155
	Lower-bound	12.750	1.000	12.750	1.955	.200
Error(trial)	Sphericity Assumed	52.167	24	2.174		
	Greenhouse-Geisser	52.167	12.868	4.054		
	Huynh-Feldt	52.167	21.655	2.409		
	Lower-bound	52.167	8.000	6.521		

The revised tests: This time, there is a significant effect of time (trial), but no interaction with either of the experimental variables.

SPSS also automatically tests for nonlinear trends. This can be helpful if you think that an effect will be curvilinear with respect to time (for example, you might expect an effect strong at first, diminishing over time). This table shows the same significant linear effect of time on the DV as we saw in the previous analyses, plus a possible quadratic effect for the interaction between anxiety and time. This effect should be interpreted cautiously given the overall nonsignificant effect of anxiety and the fact that it only barely achieved the .05 cut-off for significance. Nevertheless, it may be worth further investigation.

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	trial	Type III Sum of Squares	df	Mean Square	F	Sig.
trial	Linear	984.150	1	984.150	247.845	.000
	Quadratic	6.750	1	6.750	3.411	.102
	Cubic	.600	1	.600	1.051	.335
trial * anxiety	Linear	1.667	1	1.667	.420	.535
	Quadratic	3.000	1	3.000	1.516	.253
	Cubic	3.750	1	3.750	6.569	.033
trial * tension	Linear	10.417	1	10.417	2.623	.144
	Quadratic	.083	1	.083	.042	.843
	Cubic	1.667	1	1.667	2.920	.126
trial * anxiety * tension	Linear	9.600	1	9.600	2.418	.159
	Quadratic	.333	1	.333	.168	.692
	Cubic	2.817	1	2.817	4.934	.057
Error(trial)	Linear	31.767	8	3.971		
	Quadratic	15.833	8	1.979		
	Cubic	4.567	8	.571		

SPSS concludes the analysis with a *univariate* test, looking at the main effect of each predictor on the *average* of the four time intervals for the DV. I don't recommend using this particular analysis, as it by definition ignores effects of time. If you really didn't expect changes over time, you probably wouldn't have used a repeated-measures ANOVA. But, this analysis may help you to see if there are any main effects of your IVs on your DV, if time did *not* have a significant effect in the analyses examined previously. If you found nonsignificant effects for time, this could at least help to suggest alternative analytic strategies for further research.

Finally, we have the graphs that were specified in the “Plots” command. These clearly show the change over time (trial) in the DV. They also show the hint of a nonlinear effect of anxiety, as there are slight differences between the two anxiety groups on the type of change observed in the DV—you can see how the two groups seem to diverge at Trial 4. Again, this is not a firm conclusion (unless that was the effect we expected and that we were specifically testing for), but it might be worth further investigation.