Mixed Factorial ANOVA

Introduction

The final ANOVA design that we need to look at is one in which you have a mixture of between-group and repeated measures variables. It should be obvious that you need at least two independent variables for this type of design to be possible, but you can have more complex scenarios too (e.g. two between-group and one repeated measures, one between-group and two repeated measures, or even two of each). SPSS allows you to test almost any design you might want to of virtually any degree of complexity. However, interaction terms are difficult enough to interpret with only two variables so imagine how difficult they are if you include four!

An Example

Table 1: Data from **LooksOrPersonality.sav** (Att = Attractive, Av = Average, Ug = Ugly)

	High Charisma			Some Charisma			Dullard		
Looks	**Att**	**Av**	**Ugly**	**Att**	**Av**	**Ug**	**Att**	**Av**	**Ug**
Male	86	84	67	88	69	50	97	48	47
	91	83	53	83	74	48	86	50	46
	89	88	48	99	70	48	90	45	48
	89	69	58	86	77	40	87	47	53
	80	81	57	88	71	50	82	50	45
	80	84	51	96	63	42	92	48	43
	89	85	61	87	79	44	86	50	45
	100	94	56	86	71	54	84	54	47
	90	74	54	92	71	58	78	38	45
	89	86	63	80	73	49	91	48	39
Female	89	91	93	88	65	54	55	48	52
	84	90	85	95	70	60	50	44	45
	99	100	89	80	79	53	51	48	44
	86	89	83	86	74	58	52	48	47
	89	87	80	83	74	43	58	50	48
	80	81	79	86	59	47	51	47	40
	82	92	85	81	66	47	50	45	47
	97	69	87	95	72	51	45	48	46
	95	92	90	98	64	53	54	53	45
	95	93	96	79	66	46	52	39	47

It seems that lots of magazines go on all the time about how men and women want different things from relationships (or perhaps it's just my girlfriend's copies of Marie Clare's, which obviously I don't read, honestly). The big question to which we all want to know the answer is are looks or personality more important. Imagine you wanted to put this to the test. You devised a cunning plan whereby you'd set up a speed-dating night. Little did the people who came along know that you'd got some of your friends to act as the dates. Specifically you found 9 men and 9 women to act as the date. In each of these groups three people were extremely attractive people but differed in their personality: one had tonnes of charisma , one had some charisma, and the third person was as dull as this handout. Another three people were of average attractiveness, and again differed in their personality: one was highly charismatic, one had some charisma and the third was a dullard. The final three were, not wishing to be unkind in any way, butt-ugly and again one was charismatic, one had some charisma and the final poor soul was mind-numbingly tedious. The participants were the people who came to the speed dating night, and over the course of the evening they speed-dated all 9 members of the opposite sex that you'd set up for them. After their 5 minute date, they rated how much they'd like to have a proper date with the person as a percentage (100% = 'I'd pay large sums of money for your phone number', 0% = 'I'd pay a large sum of money for a plane ticket to get me as far away as possible from you'). As such, each participant rated 9 different people who varied in their attractiveness and personality. So, there are two repeated measures variables: **looks** (with three levels because the person could be attractive, average or ugly) and **personality** (again with three levels because the person could have lots of charisma, have some charisma, or be a dullard). Of course the people giving the ratings could be male of female, so we should also include the **gender** of the person making the ratings (male or female), and this, of course, will be a between group variable.

Running the Analysis

Data Entry

To enter these data into SPSS we use the same procedure as the two-way repeated measures ANOVA that we came across in the previous handout. Remember that each row in the data editor represents a single participant's data. If a person participates in all experimental conditions (in this case (s)he dates all of the people who differ in attractiveness and all of the people who differ in their charisma) then each experimental condition must be represented by a column in the data editor. In this experiment there are nine experimental conditions and so the data need to be entered in nine columns. Therefore, create the following nine variables in the data editor with the names as given. For each one, you should also enter a full variable name for clarity in the output.

att_high	Attractive	+	High Charisma
av_high	Average Looks	+	High Charisma
ug_high	Ugly	+	High Charisma
att_some	Attractive	+	Some Charisma
av_some	Average Looks	+	Some Charisma
ug_some	Ugly	+	Some Charisma
att_none	Attractive	+	Dullard
av_none	Average Looks	+	Dullard
ug_none	Ugly	+	Dullard

Once these variables have been created, enter the data as in Table 1. If you have problems entering the data then use the file **LooksOrPersonality.sav** from the course website. First we have to define our repeated measures variables, so access the *define factors* dialog box using the menu path **Analyze⇒General Linear Model⇒Repeated Measures …**. As with two-way repeated measures ANOVA (see the previous handout) we need to give names to our repeated measures variables and specify how many levels they have. In this case there are two within-subject factors: **looks** (attractive, average or ugly) and **charisma** (high charisma, some charisma and dullard). In the *define factors* dialog box replace the word *factor1* with the word *looks*. When you have given this repeated measures factor a name, tell the computer that this variable has 3 levels by typing the number 3 into the box labelled *Number of Levels*. Click on Add to add this variable to the list of repeated measures variables. This variable will now appear in the white box at the bottom of the dialog box and appears as *looks(3)*. Now repeat this process for the second independent variable. Enter the word *charisma* into the space labelled *Within-Subject Factor Name* and then, because there were three levels of this variable, enter the number 3 into the space labelled *Number of Levels*. Click on Add to include this variable in the list of factors; it will appear as *charisma(3)*. The finished dialog box is shown in Figure . When you have entered both of the within-subject factors click on Define to go to the main dialog box.

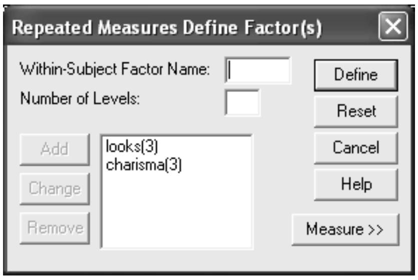


Figure 1: *Define factors* dialog box for factorial repeated measures ANOVA

The main dialog box is the same as when we did a factorial repeated measures ANOVA in the previous handout. At the top of the *Within-Subjects Variables* box, SPSS states that there are two factors: **looks** and **charisma**. In the box below there is a series of question marks followed by bracketed numbers. The numbers in brackets represent the levels of the factors (independent variables)—see the previous chapter for a more detailed explanation.

In this example, there are two independent variables and so there are two numbers in the brackets. The first number refers to levels of the first factor listed above the box (in this case **looks**). The second number in the bracket refers to levels of the second factor listed above the box (in this case **charisma**). As with the other repeated measures ANOVAs we've come across, we have to replace the question marks with variables from the list on the left-hand side of the dialog box. With between-group designs, in which coding variables are used, the levels of a particular factor are specified by the codes assigned to them in the data editor. However, in repeated measures designs, no such coding scheme is used and so we determine which condition to assign to a level at this stage (again look back to the previous handout for more about this).

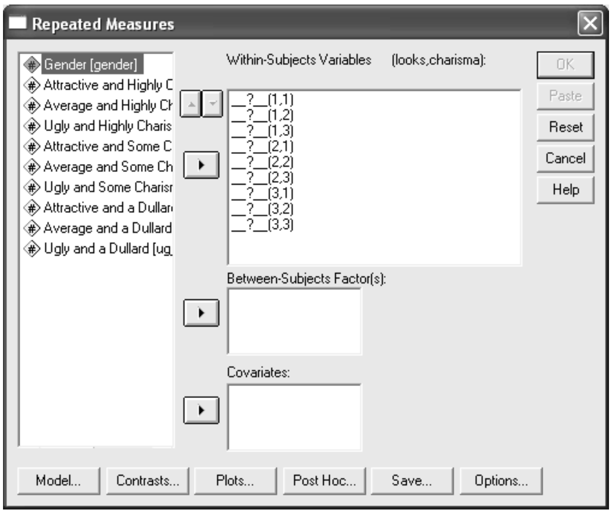


Figure 2

The variables can be entered as follows:

att_high	_?_(1,1)
att_some	_?_(1,2)
att_none	_?_(1,3)
ug_high	_?_(2,1)
ug_some	_?_(2,2)
ug_none	_?_(2,3)
av_high	_?_(3,1)
av_some	_?_(3,2)
av_none	_?_(3,3)

So far the procedure has been similar to other factorial repeated measures designs. However, we have a mixed design here, and so we also need to specify our between-group factor as well. We do this by selecting **gender** in the variables list and clicking to transfer it to the box labelled *Between-Subjects Factors*. The completed dialog box should look exactly like Figure 3. I've already discussed the options for the buttons at the bottom of this dialog box, so I'll talk only about the ones of particular interest for this example.

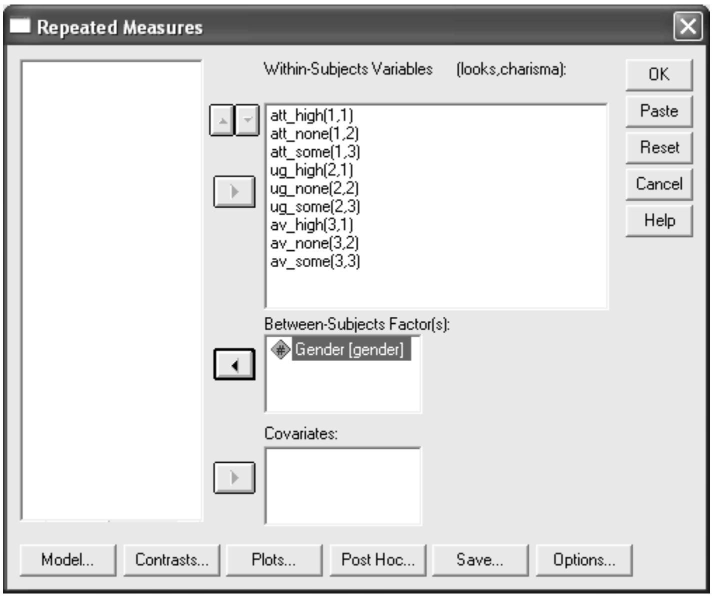


Figure 3

Other Options

The addition of an extra variable makes it necessary to choose a different graph to the one in the previous handout. Click on Plots... to access the dialog box in Figure 4. Place **looks** in the slot labelled *Horizontal Axis:* and **charisma** in slot labelled *Separate Line:*, finally, place **gender** in the slot labelled *Separate Plots*. When all three variables have been specified, don't forget to click on Add to add this combination to the list of plots. By asking SPSS to plot the looks × charisma × gender interaction, we should get the interaction graph for looks and charisma, but a separate version of this graph will be produced for male and female participants. You could also think about plotting graphs for the two way interactions (e.g. looks × charisma, looks × gender, and charisma × gender).

As far as other options are concerned, you should select the same ones that were chosen for the previous handout. It is worth selecting estimated marginal means for all effects (because these values will help you to understand any significant effects). When all of the appropriate options have been selected, run the analysis.

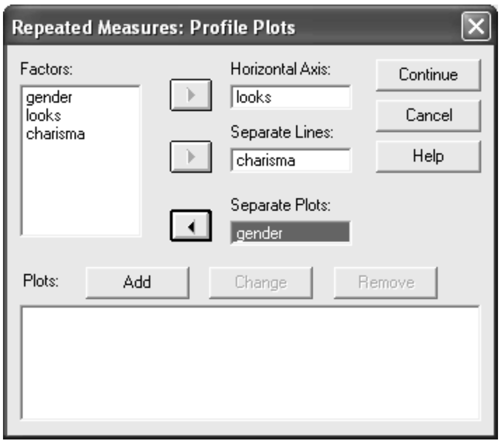


Figure 4: *Plots* dialog box for a three-way mixed ANOVA

Output for Mixed Factorial ANOVA: Main Analysis

The initial output is the same as the two-way ANOVA example: there is a table listing the repeated measures variables from the data editor and the level of each independent variable that they represent. The second table contains descriptive statistics (mean and standard deviation) for each of the nine conditions split according to whether participants were male or female (see SPSS Output 1). The names in this table are the names I gave the variables in the data editor (therefore, your output may differ slightly). These descriptive statistics are interesting because they show us the pattern of means across all experimental conditions (so, we use these means to produce the graphs of the three-way interaction).

Within-Subjects Factors

Measure: MEASURE_1

LOOKS	CHARISMA	Dependent Variable
1	1	ATT_HIGH
	2	ATT_NONE
	3	ATT_SOME
2	1	UG_HIGH
	2	UG_NONE
	3	UG_SOME
3	1	AV_HIGH
	2	AV_NONE
	3	AV_SOME

Descriptive Statistics

	Gender	Mean	Std. Deviation	N
Attractive and Highly Charismatic	Male	88.30	5.697	10
	Female	89.60	6.637	10
	Total	88.95	6.057	20
Attractive and a Dullard	Male	87.30	5.438	10
	Female	51.80	3.458	10
	Total	69.55	18.743	20
Attractive and Some Charisma	Male	88.50	5.740	10
	Female	87.10	6.806	10
	Total	87.80	6.170	20
Ugly and Highly Charismatic	Male	56.80	5.731	10
	Female	86.70	5.438	10
	Total	71.75	16.274	20
Ugly and a Dullard	Male	45.80	3.584	10
	Female	46.10	3.071	10
	Total	45.95	3.252	20
Ugly and Some Charisma	Male	48.30	5.376	10
	Female	51.20	5.453	10
	Total	49.75	5.476	20
Average and Highly Charismatic	Male	82.80	7.005	10
	Female	88.40	8.329	10
	Total	85.60	8.022	20
Average and a Dullard	Male	47.80	4.185	10
	Female	47.00	3.742	10
	Total	47.40	3.885	20
Average and Some Charisma	Male	71.80	4.417	10
	Female	68.90	5.953	10
	Total	70.35	5.314	20

SPSS Output 1

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
LOOKS	.960	.690	2	.708	.962	1.000	.500
CHARISMA	.929	1.246	2	.536	.934	1.000	.500
LOOKS * CHARISMA	.613	8.025	9	.534	.799	1.000	.250

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+GENDER
Within Subjects Design: LOOKS+CHARISMA+LOOKS*CHARISMA

SPSS Output 2

SPSS Output 2 shows the results of Mauchly's sphericity test for each of the three repeated measures effects in the model. None of the effects violate the assumption of sphericity because all of the values in the column labelled *Sig.* are above 0.05; therefore, we can assume sphericity when we look at our F-statistics.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
LOOKS	Sphericity Assumed	20779.633	2	10389.817	423.733	.000
	Greenhouse-Geisser	20779.633	1.923	10803.275	423.733	.000
	Huynh-Feldt	20779.633	2.000	10389.817	423.733	.000
	Lower-bound	20779.633	1.000	20779.633	423.733	.000
LOOKS * GENDER	Sphericity Assumed	3944.100	2	1972.050	80.427	.000
	Greenhouse-Geisser	3944.100	1.923	2050.527	80.427	.000
	Huynh-Feldt	3944.100	2.000	1972.050	80.427	.000
	Lower-bound	3944.100	1.000	3944.100	80.427	.000
Error(LOOKS)	Sphericity Assumed	882.711	36	24.520		
	Greenhouse-Geisser	882.711	34.622	25.496		
	Huynh-Feldt	882.711	36.000	24.520		
	Lower-bound	882.711	18.000	49.040		
CHARISMA	Sphericity Assumed	23233.600	2	11616.800	328.250	.000
	Greenhouse-Geisser	23233.600	1.868	12437.761	328.250	.000
	Huynh-Feldt	23233.600	2.000	11616.800	328.250	.000
	Lower-bound	23233.600	1.000	23233.600	328.250	.000
CHARISMA * GENDER	Sphericity Assumed	4420.133	2	2210.067	62.449	.000
	Greenhouse-Geisser	4420.133	1.868	2366.252	62.449	.000
	Huynh-Feldt	4420.133	2.000	2210.067	62.449	.000
	Lower-bound	4420.133	1.000	4420.133	62.449	.000
Error(CHARISMA)	Sphericity Assumed	1274.044	36	35.390		
	Greenhouse-Geisser	1274.044	33.624	37.891		
	Huynh-Feldt	1274.044	36.000	35.390		
	Lower-bound	1274.044	18.000	70.780		
LOOKS * CHARISMA	Sphericity Assumed	4055.267	4	1013.817	36.633	.000
	Greenhouse-Geisser	4055.267	3.197	1268.295	36.633	.000
	Huynh-Feldt	4055.267	4.000	1013.817	36.633	.000
	Lower-bound	4055.267	1.000	4055.267	36.633	.000
LOOKS * CHARISMA * GENDER	Sphericity Assumed	2669.667	4	667.417	24.116	.000
	Greenhouse-Geisser	2669.667	3.197	834.945	24.116	.000
	Huynh-Feldt	2669.667	4.000	667.417	24.116	.000
	Lower-bound	2669.667	1.000	2669.667	24.116	.000
Error(LOOKS*CHARISMA)	Sphericity Assumed	1992.622	72	27.675		
	Greenhouse-Geisser	1992.622	57.554	34.622		
	Huynh-Feldt	1992.622	72.000	27.675		
	Lower-bound	1992.622	18.000	110.701		

SPSS Output 3

SPSS Output 3 shows the summary table of the repeated measures effects in the ANOVA with corrected *F* values. As with factorial repeated measures ANOVA the output is split into sections for each of the effects in the model and their associated error terms. The only difference is that the interactions between our between-group variable of gender and the repeated measures effects are included also.

Again, we need to look at the column labelled *Sig.* and if the values in this column are less than 0.05 for a particular effect then it is statistically significant. Working down from the top of the table we find a significant effect of looks, which means that if we ignore whether the date was charismatic, and whether the rating was from a man or a woman, then the attractiveness of a person significantly affected the ratings they received. The looks × gender interaction is also significant, which means that although the ratings were affected by whether the date was attractive, average or ugly, the way in which ratings were affected by attractiveness was different in male and female raters.

Next we find a significant effect of charisma, which means that if we ignore whether the date was attractive, and whether the rating was from a man or a woman, then the charisma of a person significantly affected the ratings they received. The charisma × gender interaction is also significant, so although the ratings were affected by whether the date had high charisma, some charisma or was a dullard, the way in which ratings were affected by charisma was different in male and female raters.

Next we find a significant interaction between looks and charisma, which means that if we ignore the gender of the rater, the profile of ratings across different levels of attractiveness was different for highly charismatic dates, charismatic dates and dullards. (It is equally true to say this the opposite way around: the profile of ratings across different levels of charisma was different for attractive, average and ugly dates). Just to add to the mounting confusion, the looks × charisma × gender interaction is also significant, meaning that the looks × charisma interaction was significantly different in men and women participants!

This is all a lot to take in so we'll look at each of these effects in turn in subsequent sections. First though, we need to see what has happened to our main effect of gender.

The Effect of Gender

The main effect of gender is listed separately from the repeated measure effects in a table labelled *Tests of Between-Subjects Effects*. Before looking at this table it is important to check the assumption of homogeneity of variance using Levene's test SPSS produces a table listing Levene's test for each of the repeated measures variables in the data editor, and we need to look for any variable that has a significant value. SPSS Output 4 shows both tables. The table showing Levene's test indicates that variances are homogeneous for all levels of the repeated measures variables (because all significance values are greater than 0.05). If any values were significant, then this would compromise the accuracy of the *F*-test for gender, and we would have to consider transforming all of our data to stabilize the variances between groups (see Field, 2004, Chapter 3). Fortunately, in this example a transformation is unnecessary. The second table shows the ANOVA summary table for the main effect of gender, and this reveals a non-significant effect (because the significance of 0.946 is greater than the standard cut-off point of 0.05).

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

	F	df1	df2	Sig.
Attractive and Highly Charismatic	1.131	1	18	.302
Attractive and a Dullard	1.949	1	18	.180
Attractive and Some Charisma	.599	1	18	.449
Ugly and Highly Charismatic	.005	1	18	.945
Ugly and a Dullard	.082	1	18	.778
Ugly and Some Charisma	.124	1	18	.729
Average and Highly Charismatic	.102	1	18	.753
Average and a Dullard	.004	1	18	.950
Average and Some Charisma	1.763	1	18	.201

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

a. Design: Intercept+GENDER
Within Subjects Design: LOOKS+CHARISMA+LOOKS*CHARISMA

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	94027.756	1	94027.756	20036.900	.000
GENDER	.022	1	.022	.005	.946
Error	84.469	18	4.693		

SPSS Output 4

We can report that there was a non-significant main effect of gender, $F(1, 18) < 1$, *ns*. This effect tells us that if we ignore all other variables, male participants' ratings were basically the same as females'. If you requested that SPSS display means for the gender effect you should scan through your output and find the table in a section headed *Estimated Marginal Means*. SPSS Output 5 is a table of means for the main effect of gender with the associated standard errors. This information is plotted in Figure 5. It is clear from this graph that men and women's ratings were generally the same.

1. Gender

Measure: MEASURE_1

Gender	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Male	68.600	.685	67.161	70.039
Female	68.533	.685	67.094	69.973

SPSS Output 5

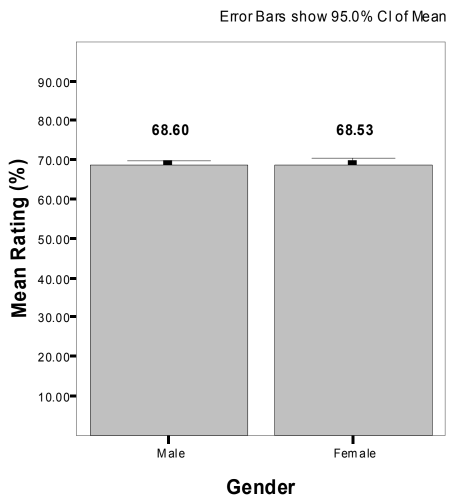


Figure 5

The Effect of Looks

We came across the main effect of looks in SPSS Output 3. Now we're going to have a look at what this effect means. We can report that there was a significant main effect of looks, $F(2, 36) = 423.73$, $p < .001$. This effect tells us that if we ignore all other variables, ratings were different for attractive, average and unattractive dates. If you requested that SPSS display means for the looks effect (I'll assume you did from now on) you will find the table in a section headed *Estimated Marginal Means*. SPSS Output 6 is a table of means for the main effect of looks with the associated standard errors. The levels of looks are labelled simply 1, 2 and 3, and it's down to you to remember how you entered the variables (or you can look at the summary table that SPSS produces at the beginning of the output—see SPSS Output 1). If you followed what I did then level 1 is attractive, level 2 is ugly and level 3 is average. To make things easier, this information is plotted in Figure 6. You can see that as attractiveness falls, the mean rating falls too. So this main effect seems to reflect that the raters were more likely to express a greater interest in going out with attractive people than average or ugly people. However, we really need to look at some contrasts to find out exactly what's going on (see Field, 2000 or 2004 if you're interested).

2. LOOKS

Measure: MEASURE_1

LOOKS	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	82.100	.652	80.729	83.471
2	55.817	.651	54.449	57.184
3	67.783	.820	66.061	69.505

SPSS Output 6

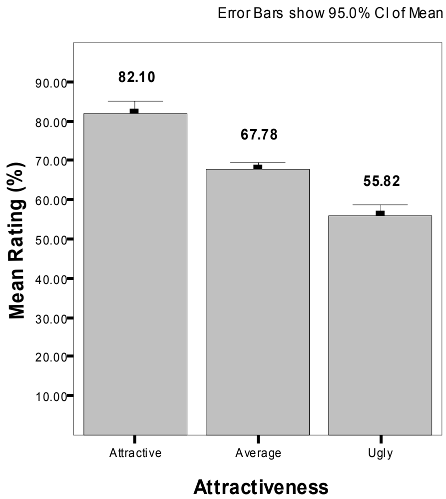


Figure 6

The Effect of Charisma

The main effect of charisma is in SPSS Output 3. We can report that there was a significant main effect of charisma, $F(2, 36) = 328.25$, $p < .001$. This effect tells us that if we ignore all other variables, ratings were different for highly charismatic, a bit charismatic and dullard people. The table labelled *CHARISMA* in the section headed *Estimated Marginal Means* tells us what this effect means (SPSS Output 7). Again, the levels of charisma are labelled simply 1, 2 and 3. If you followed what I did then level 1 is high charisma, level 2 is no charisma and level 3 is some charisma. This information is plotted in Figure 7: As charisma declines, the mean rating falls too. So this main effect seems to reflect that the raters were more likely to express a greater interest in going out with charismatic people than average people or dullards.

3. CHARISMA

Measure: MEASURE_1

CHARISMA	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	82.100	1.010	79.978	84.222
2	54.300	.573	53.096	55.504
3	69.300	.732	67.763	70.837

SPSS Output 7

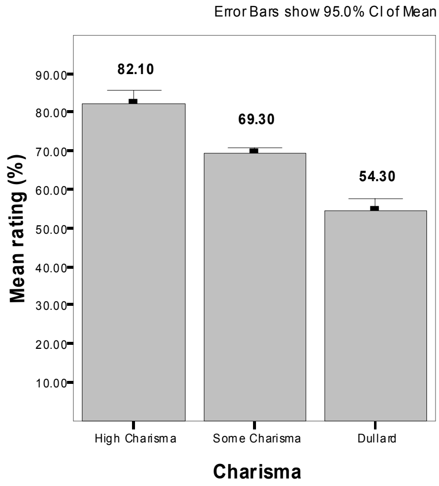


Figure 7

The Interaction between Gender and Looks

SPSS Output 3 indicated that gender interacted in some way with the attractiveness of the date. From the summary table we should report that there was a significant interaction between the attractiveness of the date and the gender of the participant, $F(2, 36) = 80.43$, $p < 0.001$. This effect tells us that the profile of ratings across dates of different attractiveness was different for men and women. We can use the estimated marginal means to determine the nature of this interaction (or we could have asked SPSS for a plot of gender $\times$ look using the dialog box in Figure 4). The means and interaction graph (Figure 8 and SPSS Output 8) show the meaning of this result. The graph shows the average male ratings of dates of different attractiveness ignoring how charismatic the date was (circles). The women's scores are shown as squares. The graph clearly shows that male and female ratings are very similar for average looking dates, but men give higher ratings (i.e. they're really keen to go out with these people) than women for attractive dates, but women express more interest in going out with ugly people than men. In general this interaction seems to suggest than men's interest in dating a person is more influenced by their looks than for females. Although both male's and female's interest decreases as attractiveness decreases, this decrease is more pronounced for men.

4. Gender * LOOKS

Measure: MEASURE_1

Gender	LOOKS	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Male	1	88.033	.923	86.095	89.972
	2	50.300	.921	48.366	52.234
	3	67.467	1.159	65.031	69.902
Female	1	76.167	.923	74.228	78.105
	2	61.333	.921	59.399	63.267
	3	68.100	1.159	65.665	70.535

SPSS Output 8

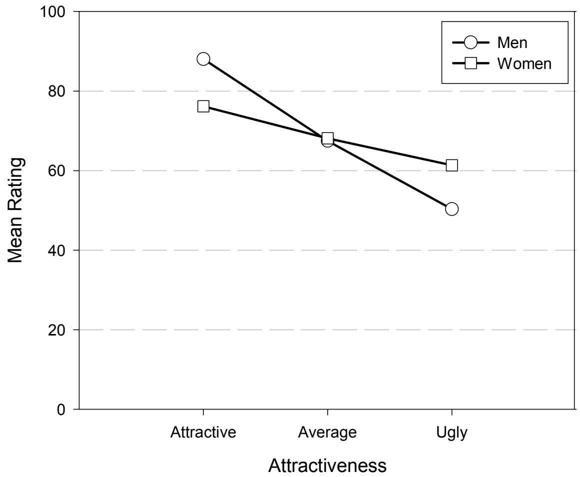


Figure 8

The Interaction between Gender and Charisma

SPSS Output 3 indicated that gender interacted in some way with how charismatic the date was. From the summary table we should report that there was a significant interaction between the attractiveness of the date and the gender of the participant, $F(2, 36) = 62.45$, $p < 0.001$. This effect tells us that the profile of ratings across dates of different levels of charisma was different for men and women. The estimated marginal means (or a plot of gender × charisma using the dialog box in Figure 4) tell us the meaning of this interaction (see Figure 9 and SPSS Output 9) show the meaning of this result. The graph shows the average male ratings of dates of different levels of charisma ignoring how attractive they were (circles). The women's scores are shown as squares. The graph shows almost the reverse pattern as for the attractiveness data; again male and female ratings are very similar for dates with normal amounts of charisma, but this time men show more interest in dates who are dullards than women do, and women show slightly more interest in very charismatic dates than men do. In general this interaction seems to suggest than women's interest in dating a person is more influenced by their charisma than for men. Although both male's and female's interest decreases as charisma decreases, this decreases is more pronounced for females.

5. Gender * CHARISMA

Measure: MEASURE_1

Gender	CHARISMA	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Male	1	75.967	1.428	72.966	78.967
	2	60.300	.810	58.598	62.002
	3	69.533	1.035	67.360	71.707
Female	1	88.233	1.428	85.233	91.234
	2	48.300	.810	46.598	50.002
	3	69.067	1.035	66.893	71.240

SPSS Output 9

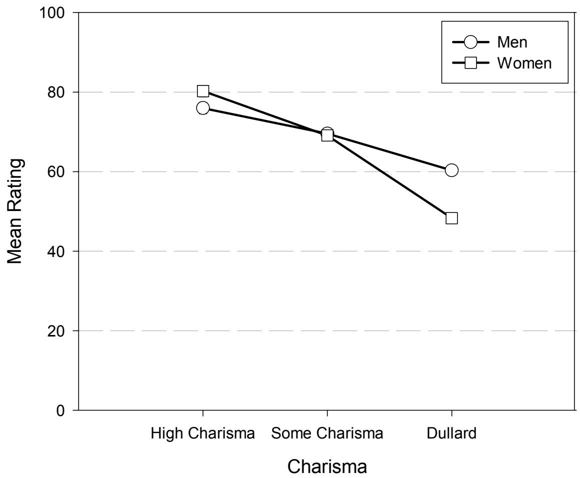


Figure 9

The Interaction between Attractiveness and Charisma

SPSS Output 3 indicated that the attractiveness of the date interacted in some way with how charismatic the date was. From the summary table we should report that there was a significant interaction between the attractiveness of the date and the charisma of the date, $F(4, 72) = 36.63$, $p < 0.001$. This effect tells us that the profile of ratings across dates of different levels of charisma was different attractive, average and ugly dates. The estimated marginal means (or a plot of looks × charisma using the dialog box in Figure 4) tell us the meaning of this interaction (see Figure 10 and SPSS Output 10) show the meaning of this result.

The graph shows the average ratings of dates of different levels of attractiveness when the date also had high levels of charisma (circles), some charisma (squares) and no charisma (triangles). Look first at the difference between attractive and average-looking dates. The interest in highly charismatic dates doesn't change (the line is more or less flat between these two points), but for dates with some charisma or no charisma interest levels decline. So, if you have lots of charisma you can get away with being average looking and people will still want to date you. Now, if we look at the difference between average looking and ugly dates, a different pattern is observed. For dates with no charisma (triangles) there is no difference between ugly and average people (so if you're a dullard you have to be really attractive before people want to date you). However, for those with charisma, there is a decline in interest if you're ugly (so, if you're ugly, having charisma won't help you much). This interaction is very complex!

6. LOOKS * CHARISMA

Measure: MEASURE_1

LOOKS	CHARISMA	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	88.950	1.383	86.045	91.855
	2	69.550	1.019	67.409	71.691
	3	87.800	1.408	84.842	90.758
2	1	71.750	1.249	69.126	74.374
	2	45.950	.746	44.382	47.518
	3	49.750	1.211	47.206	52.294
3	1	85.600	1.721	81.985	89.215
	2	47.400	.888	45.535	49.265
	3	70.350	1.172	67.888	72.812

SPSS Output 10

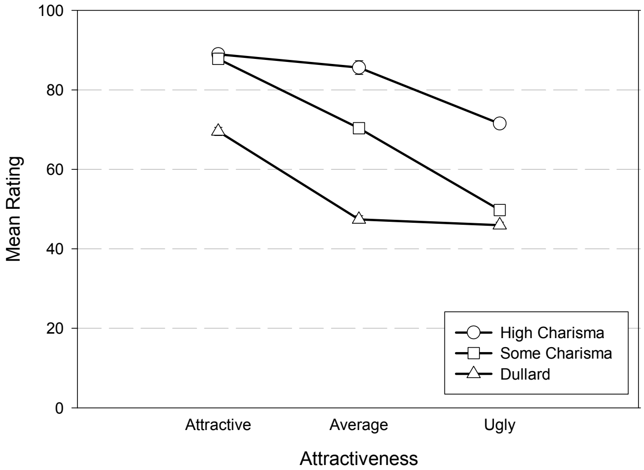
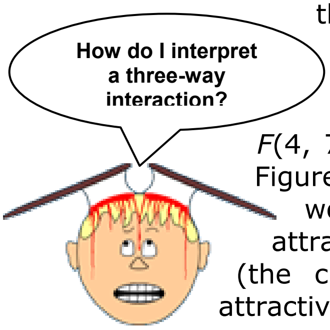


Figure 10

The Interaction between Looks, Charisma and Gender

The three-way interaction tells us whether the looks × charisma interaction described above is the same for men and women (i.e. whether the combined effect of attractiveness of the date and their level of charisma is the same for male participants as for female subjects). SPSS Output 3 tells us that there is a significant three-way looks × charisma × gender interaction, $F(4, 72) = 24.12$, $p < .001$. The nature of this interaction is shown up in Figure 11, which shows the looks by charisma interaction for men and women separately. The male graph shows that when dates are attractive, men will express a high interest regardless of charisma levels (the circle, square and dot all overlap). At the opposite end of the attractiveness scale, when a date is ugly, regardless of charisma men will express very little interest (ratings are all low). The only time charisma makes any difference to a man is if the date is average looking, in which case high charisma boosts interest, being a dullard reduces interest, and having a bit of charisma leaves things somewhere in between. The take home message being that men are superficial cretins who are more interested in physical attributes. The picture for women is very different. If someone has high levels of charisma then it doesn't really matter what they look like, women will express an interest in them (the line of circles is relatively flat. At the other extreme, if the date is a dullard, then they will express no interest in them, regardless of how attractive they are (the line of triangles is relatively flat). The only time attractiveness makes a difference is when someone has an average amount of charisma, in which case being attractive boosts interest, and being ugly reduces it. Put another way, women prioritise charisma over physical appearance. Again, we can look at some contrasts to further break this interaction down. These contrasts are similar to those for the looks × charisma interaction, but they now also take into account the effect of gender as well!



7. Gender * LOOKS * CHARISMA

Measure: MEASURE_1

Gender	LOOKS	CHARISMA	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Male	1	1	88.300	1.956	84.191	92.409
		2	87.300	1.441	84.273	90.327
		3	88.500	1.991	84.317	92.683
	2	1	56.800	1.767	53.089	60.511
		2	45.800	1.055	43.583	48.017
		3	48.300	1.712	44.703	51.897
	3	1	82.800	2.434	77.687	87.913
		2	47.800	1.255	45.163	50.437
		3	71.800	1.657	68.318	75.282
Female	1	1	89.600	1.956	85.491	93.709
		2	51.800	1.441	48.773	54.827
		3	87.100	1.991	82.917	91.283
	2	1	86.700	1.767	82.989	90.411
		2	46.100	1.055	43.883	48.317
		3	51.200	1.712	47.603	54.797
	3	1	88.400	2.434	83.287	93.513
		2	47.000	1.255	44.363	49.637
		3	68.900	1.657	65.418	72.382

SPSS Output 11

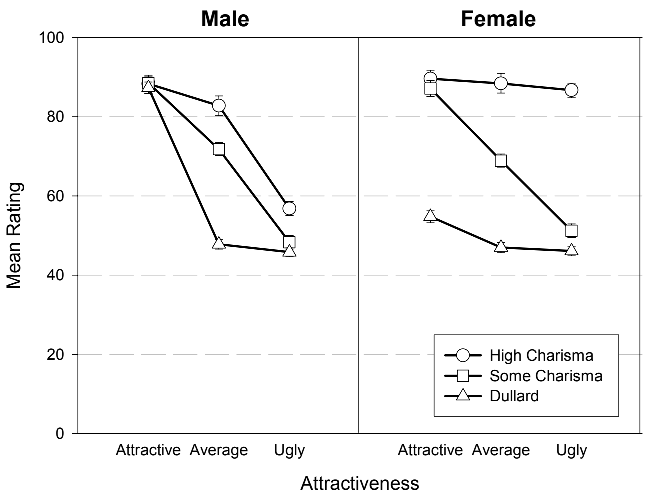


Figure 11: Graphs showing the looks by charisma interaction for men and women. Lines represent high charisma (circles), some charisma (squares) and no charisma (triangles)

Conclusions

What should be clear from this handout is that when more than two independent variables are used in an ANOVA, it yields complex interaction effects that require a great deal of concentration to interpret (imagine interpreting a four-way interaction!). Therefore, it is essential to take a systematic approach to interpretation and plotting graphs is a particularly useful way to proceed.

Example 2:

There is evidence that attitudes towards stimuli can be changed using positive and negative imagery (e.g. Stuart, Shimp and Engle, 1987, but see Field and Davey, 1999) and these researchers were interested in answering two questions. On the one hand, the government had funded them to look at whether negative imagery in advertising could be used to change attitudes towards alcohol. Conversely, an alcohol company had provided funding to see whether positive imagery could be used to improve attitudes towards alcohol. The scientists designed *two* studies to address both issues.

In the first study, participants viewed a total of 3 mock adverts over three sessions. In one session, they saw three adverts containing three different products with a negative image (a dead body with the slogan 'drinking this product makes your liver explode': The products were: (1) a brand of beer (Brain Death)); (2) a brand of wine (Dangleberry); and (3) a brand of water (Puritan). The gender of the participants was noted. Table 1 contains the data (each row represents a single subject). After each advert subjects were asked to rate the drinks on a scale ranging from –100 (dislike very much) through 0 (neutral) to 100 (like very much). The order of adverts was randomised. There are two independent variables in each experiment: the type of drink (beer, wine or water) and the gender of the participant (male or female).

Enter the data and analyse with a mixed ANOVA (this is an easier example as there are only 2 independent variables).

Table 2: Data for Experiment 1

Experiment	**Experiment 1: Negative Imagery**		
Drink	Beer	Wine	Water
Male	6	–5	–14
	30	–12	–10
	15	–15	–16
	30	–4	–10
	12	–2	5
	17	–6	–6
	21	–2	–20
	23	–7	–12
	20	–10	–9
	27	–15	–6
Female	–19	–13	–2
	–18	–16	–17
	–8	–23	–19
	–6	–22	–11
	–6	–9	–10
	–9	–18	–17
	–17	–17	–4
	–12	–15	–4
	–11	–14	–1
	–6	–15	–1

Answer

Mauchly's sphericity test for the repeated measures variable is shown below. The main effect of drink does not significantly violate the sphericity assumption because the significance value is greater than 0.05 [W = .847, $p > 0.05$]. Therefore, the F-value for the main effect of drink (and its interaction with the between-group variable **gender**) does not need to be corrected for violations of sphericity (see last handout).

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
DRINK	.847	2.827	2	.243	.867	1.000	.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 layers (by default) of the Tests of Within Subjects Effects table.

b.
Design: Intercept+GENDER
Within Subjects Design: DRINK

The summary table of the repeated measures effects in the ANOVA with corrected *F*-values is below. The output is split into sections for each of the effects in the model and their associated error terms. The table format is the same as for other examples we have seen, except that the interactions between gender and the repeated-measures effects are included also. By looking at the significance values it is clear that there are significant effects of the type of drink used and the interaction of this and the gender of the subject.

The fact that gender interacts significantly with the type of drink used tells us that men and women respond differently to the adverts for Beer, Wine and Water.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
DRINK	Sphericity Assumed	3098.433	2	1549.217	40.060	.000
	Greenhouse-Geisser	3098.433	1.734	1786.545	40.060	.000
	Huynh-Feldt	3098.433	2.000	1549.217	40.060	.000
	Lower-bound	3098.433	1.000	3098.433	40.060	.000
DRINK * GENDER	Sphericity Assumed	2788.033	2	1394.017	36.047	.000
	Greenhouse-Geisser	2788.033	1.734	1607.570	36.047	.000
	Huynh-Feldt	2788.033	2.000	1394.017	36.047	.000
	Lower-bound	2788.033	1.000	2788.033	36.047	.000
Error(DRINK)	Sphericity Assumed	1392.200	36	38.672		
	Greenhouse-Geisser	1392.200	31.218	44.597		
	Huynh-Feldt	1392.200	36.000	38.672		
	Lower-bound	1392.200	18.000	77.344		

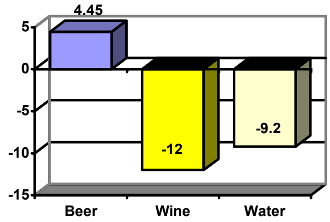
The Effect of Drink

For the effect of drink noted above we can look at how to interpret this result. Initially, we should report that there was a significant main effect of drink [$F(2, 36) = 40.06$, $p < 0.001$]. This effect tells us that if we ignore the gender of participants, some types of drink were still rated significantly differently to others. In the section on contrasts we requested that SPSS display means for all of the effects in the model (before conducting post hoc tests) and if you scan through your output you should find the table in a section headed *Estimated Marginal Means*. This is a table of means for the main effect of drink with the associated standard errors. The levels of this variable are labelled, 1, 2 and 3 and so we must think back to how we entered the variable to see which row of the table relates to which condition. We entered this variable with the beer condition first and the water condition last. The graph displays this information. It is clear from this graph that beer is naturally rated higher than wine and water (with beer being rated most highly). To see the nature of this effect we can look at the post hoc tests (see below) and the contrasts (see section below).

Estimates

Measure: MEASURE_1

DRINK	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	4.450	1.482	1.337	7.563
2	-12.000	1.018	-14.139	-9.861
3	-9.200	1.556	-12.470	-5.930



The pairwise comparisons for the main effect of drink corrected using a Bonferroni adjustments are below. This table indicates that the significant main effect reflects a significant difference ($p < 0.01$) between levels 1 and 2 (beer and wine) and 1 and 3 (beer and water) but not between levels 2 and 3 (wine and water). This seems to indicate that negative imagery had an effect on ratings of both wine and water but not on beer.

Pairwise Comparisons

Measure: MEASURE_1

(I) DRINK	(J) DRINK	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
1	2	16.450*	1.996	.000	11.181	21.719
	3	13.650*	2.262	.000	7.681	19.619
2	1	-16.450*	1.996	.000	-21.719	-11.181
	3	-2.800	1.581	.281	-6.974	1.374
3	1	-13.650*	2.262	.000	-19.619	-7.681
	2	2.800	1.581	.281	-1.374	6.974

Based on estimated marginal means

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

a. Adjustment for multiple comparisons: Bonferroni.

Between-Group Variables: Gender

The main effect of gender is listed separately from the repeated measure effects in a table labelled *tests of between-subjects effects*. Before looking at this table it is important to check the assumption of homogeneity of variance using Levene's test (see Field, 2000 chapter 6). SPSS produces a table listing Levene's test for each level of the repeated-measures variables in the Data Editor, and we need to look for any variable that has a significant value. SPSS Output 6 shows both tables. The table showing Levene's test indicates that variances are homogenous for all levels of the repeated measures variables (because all significance values are greater than 0.05). If any values were significant, then this would compromise the accuracy of the F-test for gender, and we would have to consider transforming all of our data to stabilise the variances between groups (one popular transformation is to take the square root of all values). Fortunately, in this example a transformation is unnecessary. The second table shows the ANOVA summary table for the main effect of gender, and this reveals a significant effect (because the significance of 0.000, is less than the standard cut-off point of 0.05).

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

	F	df1	df2	Sig.
Beer + Corpse	1.305	1	18	.268
Wine+ Corpse	1.048	1	18	.320
Water+ Corpse	.804	1	18	.382

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

a.
Design: Intercept+GENDER
Within Subjects Design: DRINK

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	623.472	1	623.472	52.319	.000
GENDER	823.472	1	823.472	69.103	.000
Error	214.500	18	11.917		

We report that there was a significant main effect of gender [F (1, 18) = 69.10, $p < 0.001$]. This effect tells us that if we ignore all other variables, male subjects' ratings were significantly different to females. If you requested that SPSS display means for the gender effect you should find the table in the output headed *Estimated Marginal Means*. The table of means for the main effect of gender with the associated standard errors is below. This information is plotted in the graph. It is clear from this graph that men's ratings were generally significantly more positive than females. Therefore, men gave more positive ratings than women regardless of the drink being advertised.

Estimates

Measure: MEASURE_1

Gender	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Male	.833	1.092	-1.460	3.127
Female	-12.000	1.092	-14.293	-9.707

SPSS Output 7

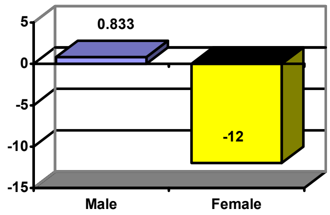


Figure 6

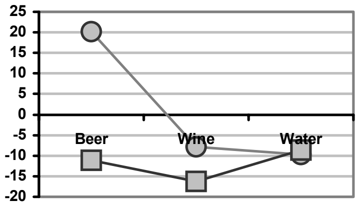
The Interaction between Gender and Drink

We know that gender interacted in some way with the type of drink used as a stimulus. From the summary table we should report that there was a significant interaction between the type of drink used and the gender of the subject [F (2, 36) = 36.05, $p < 0.001$]. This effect tells us that the type of drink being advertised had a different effect on men and women. We can use the estimated marginal means to determine the nature of this interaction (or we could have asked SPSS for a plot of gender × drink using the dialog box in Figure 4). The means and interaction graph shows the meaning of this result. The graph shows the average male ratings of each drink (circles) and the women's scores are shown as squares. The graph clearly shows that male and female ratings are very similar for wine and water, but men seem to rate beer more highly than women—regardless of the imagery in the advert. We could interpret this interaction as meaning that the type of drink being advertised influenced ratings differently in men and women. Specifically, ratings were similar for wine and water but males rated beer higher than women. Therefore, negative advertising seems to work, except that males cannot be put off of their beer by seeing this kind of advert! This interaction can be clarified using contrasts (see Field 2000, Chapter 9).

3. Gender * DRINK

Measure: MEASURE_1

Gender	DRINK	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Male	1	20.100	2.096	15.697	24.503
	2	-7.800	1.440	-10.825	-4.775
	3	-9.800	2.201	-14.424	-5.176
Female	1	-11.200	2.096	-15.603	-6.797
	2	-16.200	1.440	-19.225	-13.175
	3	-8.600	2.201	-13.224	-3.976



Example 3:

In a second experiment (a week later), the subjects saw the same three brands, but this time presented with positive images (a sexy naked man or women—depending on the subject's gender—and the slogan 'drinking this product makes you a horny stud-muffin'). After each advert subjects were asked to rate the drinks on the same scale. Run the same analysis on this second set of data. What conclusions do you reach? The really adventurous might also want to try running this analysis as a three-way ANOVA in which the type of imagery (positive or negative) is a second repeated measures variable (for help see Field, 2000, chapter 9).

Experiment	Experiment 2: Positive Imagery		
Drink	**Beer**	**Wine**	**Water**
Male	1	38	10
	43	20	9
	15	20	6
	10	28	20
	8	11	27
	17	17	9
	30	15	19
	34	27	12
	34	24	12
	26	23	21
Female	1	28	33
	7	26	23
	22	34	21
	30	32	17
	40	24	15
	15	29	13
	20	30	16
	9	24	17
	14	34	19
	15	23	29

This handout contains large excerpts of the following texts (so copyright exists!)

Field, A. P. (2000). *Discovering statistics using SPSS for Windows: advanced techniques for the beginner*. London: Sage.

Field, A. P. (2004). *Discovering statistics using SPSS: advanced techniques for the beginner (second edition)*. London: Sage.