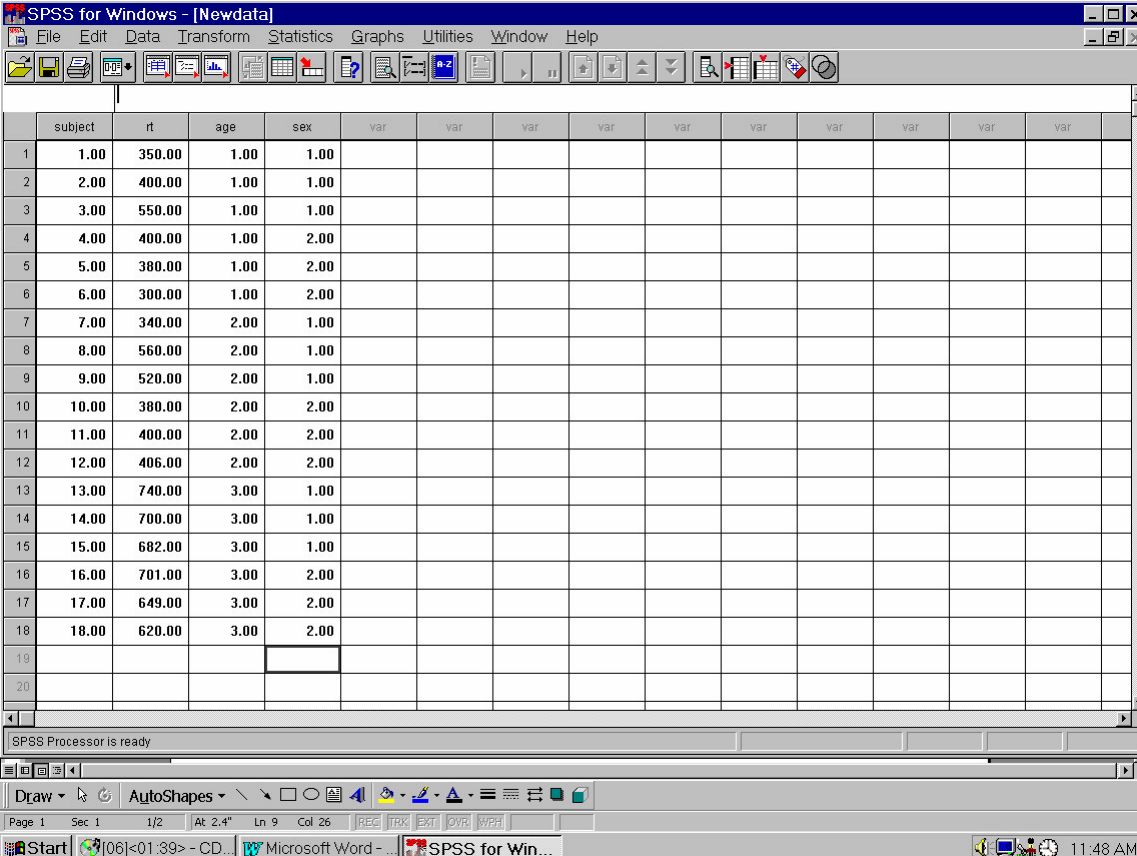# Research Methods II: Spring Term 2004

Using SPSS: Two-way Between-Subjects ANOVA

1. Entering the data:

In a two-way between subjects ANOVA, we are looking at the effect of two independent variables, for example age (in this case, with three levels: young, middle-aged, and old) and sex (two levels), on a dependent variable, for example reaction time. The completed data set might look like this.

	subject	rt	age	sex	var	var	var	var	var	var	var	var	var	var
1	1.00	350.00	1.00	1.00										
2	2.00	400.00	1.00	1.00										
3	3.00	550.00	1.00	1.00										
4	4.00	400.00	1.00	2.00										
5	5.00	380.00	1.00	2.00										
6	6.00	300.00	1.00	2.00										
7	7.00	340.00	2.00	1.00										
8	8.00	560.00	2.00	1.00										
9	9.00	520.00	2.00	1.00										
10	10.00	380.00	2.00	2.00										
11	11.00	400.00	2.00	2.00										
12	12.00	406.00	2.00	2.00										
13	13.00	740.00	3.00	1.00										
14	14.00	700.00	3.00	1.00										
15	15.00	682.00	3.00	1.00										
16	16.00	701.00	3.00	2.00										
17	17.00	649.00	3.00	2.00										
18	18.00	620.00	3.00	2.00										

2. The ANOVA commands:

(a) Click on "Analyze" on the SPSS controls at the top of the screen. Then click on "General linear model" on the menu that appears. Click on "univariate" on the next menu. You will now be presented with a dialog box.

(b) Highlight the dependent variable on which you want to perform the ANOVA: "rt", the column containing the reaction-time scores. Click on the upper arrow-button to move "rt" into the box entitled "dependent variable". We have now told SPSS where our data, our scores, are located.

(c) We now need to tell SPSS about each of the independent variables in our experiment. One independent variable is "age". Highlight "age" and use the arrow-button to move it into the box labelled "Fixed factor(s)". Then highlight "sex" and use the same arrow-button to move it into the box labelled "Fixed factor(s)".

(d) Click on "Options" which will open another dialogue box. Click on the box for "Descriptive statistics" and then "Continue". This is to produce means and standard deviations for all our conditions.

(e) Click on "Plots...". Highlight one of your factors and use the arrow-button to move it into the "Horizontal axis" box, and move the other factor into the "separate lines" box. Click on "Add" and then "Continue". This will produce a line graph of your data, and give you an visual impression of any main effects or interaction in your data (remember that the size of these effects have to be to be large enough relative to your mean square error to be significant - for example, it might look like the lines are somewhat non-parallel even though the interaction is non-significant).

(f) click on "Post hoc...". Highlight "age" and use the arrow-button to put it in the "Post hoc tests for" box. Then click on "S-N-K" (for Student Newman-Keuls"), or whichever post hoc test you prefer, and "Continue". Note that we only ask for post hoc tests for age; we do not ask for post hoc tests for "sex". Why? Because age has three levels and sex has only two. If we get a significant main effect of age we do not know exactly which ages are significantly different from which others: that is what the post hoc test tells us. If we get a significant main effect of sex, we already know everything there is to know about that main effect: men and women differ. There is nothing more to find out about the main effect of sex.

(g) Now click on "OK", and the results of our ANOVA appear in the "output" window, looking something like this:

Univariate Analysis of Variance

Between-Subjects Factors

		N
AGE	1.00	6
	2.00	6
	3.00	6
SEX	1.00	9
	2.00	9

[Below is presented the mean and standard deviation for each combination of age and sex]

Descriptive Statistics
Dependent Variable: RT

AGE	SEX	Mean	Std. Deviation	N
1.00	1.00	433.3333	104.0833	3
	2.00	360.0000	52.9150	3
	Total	396.6667	84.0635	6
2.00	1.00	473.3333	117.1893	3
	2.00	395.3333	13.6137	3
	Total	434.3333	85.9806	6
3.00	1.00	707.3333	29.6873	3
	2.00	656.6667	41.0406	3
	Total	682.0000	42.3840	6
Total	1.00	538.0000	150.9669	9
	2.00	470.6667	144.4360	9
	Total	504.3333	147.4537	18

Note: Ignore the rows called "corrected model" and "intercept". The "error" refers to the within-group variability, and the mean square (MS) associated with it is sometimes called "mean square error", or MSe. It is just the average of the variances for each of your six groups. We now have three treatment mean squares, one for each main effect ("age", "sex") and also for the interaction ("age by sex"). Each MS is just the SS divided by its DF. For each of the three effects, the F ratio is the mean square treatment divided by the MSe – i.e. a comparison of the between group variability to the within group variability.
Note that the DF for each main effect is (as in the one-way case) the number of levels of that independent variable – 1.
The degrees of freedom for the variance within each group is the number of subjects – 1; as there are 3 subjects in a group, the DF are 2. Because there are 6 such groups, the DF for the MSe is 6*2 = 12.
There is a general rule for determining the DF for an interaction: it is the product of the degrees of freedom for each independent variable in the interaction. The row labelled "Total" just gives the variance for all your data taken as a whole, ignoring the independent variables.

Tests of Between-Subjects Effects
Dependent Variable: RT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	309388.667	5	61877.733	12.327	.000
Intercept	4578338.000	1	4578338.000	912.090	.000
AGE	288345.333	2	144172.667	28.722	.000
SEX	20402.000	1	20402.000	4.064	.067
AGE * SEX	641.333	2	320.667	.064	.938
Error	60235.333	12	5019.611		
Total	4947962.000	18			
Corrected Total	369624.000	17			

a R Squared = .837 (Adjusted R Squared = .769)

[Note that there was a significant main effect of age. This significant effect entitles us to look at the results of the post hoc tests. If the main effect had not been significant we could not have looked at the post hoc tests: This is what "post hoc" means. It is Latin for "after the fact", i.e. a test used only after the fact of getting a significant overall effect.
The Newman Keuls below shows that ages 1 and 2 were not significantly different from each other, but age 3 was significantly different from each of the other ages.]

Post Hoc Tests

AGE

Homogeneous Subsets

RT
Student-Newman-Keuls

AGE	N	Subset 1	Subset 2
1.00	6	396.6667	
2.00	6	434.3333	
3.00	6		682.0000
Sig.		.375	1.000

Means for groups in homogeneous subsets are displayed. Based on Type III Sum of Squares The error term is Mean Square(Error) = 5019.611.
a Uses Harmonic Mean Sample Size = 6.000.
b Alpha = .05.

(c) Checking assumptions
The same assumptions are made for this analysis as for the one-way between-subjects analysis of variance. That is, the distribution of scores within any cell should be roughly normal. A cell is a particular combination of levels of the independent variables; in this case, a particular sex and age (e.g. young males). There are six cells in our example. Enter a new variable in the data spreadsheet (let's call it "cell"), where you give a different number for each cell (1 to 6). To produce a histogram for each level of "cell" click on "Analyze" -> "Descriptive statistics" -> "Explore", then click on "Plots…" and choosing either or both of histogram and stem-and-leaf plots.
Each cell should have approximately the same variance (the homogeneity of variance assumption). If you have equal numbers of subjects in each cell, then so long as the largest variance is no more than five times the smallest, then your variances are equal enough.

Interpreting the results:

Now we can start to make sense of what is going on in these data. The ANOVA tells us we have a statistically significant effect of age, $F(2, 12) = 28.72$, $p<0.001$. There is no significant effect of sex, $F(1, 12) = 4.06$, $p = 0.067$. The interaction between sex and age is not significant either, $F(2, 12) = 0.06$, $p = 0.938$. To interpret the effect of age, we need only consider the means for the different ages, averaged over males and females. Inspection of the means for the different ages suggests that as people get older, their reaction times get longer (the means are 397, 434, and 682, respectively). But exactly which ages are different from which others? The Newman-Keuls shows the pattern: The oldest group is significantly slower than the younger two groups, which not differ significantly amongst themselves.

3. What to do with a significant interaction

If your interaction is significant it means that the effect of age is different for the different sexes (equivalently: the effect of sex - i.e. the difference between males and females – is different for different ages). But a significant F for the interaction does not tell you any more than this: You don't know what effect, if any, age has for the different sexes. Maybe there is no effect on age for the females but there is for the males. Maybe there is an effect of age for both males and females, but a different pattern for each: Maybe a larger slowing of rt with age for females than males, or maybe a speeding of age for females and a slowing for males, or something else.

Thus, if you have a significant interaction, and you regard the interaction as potentially theoretically interesting, you need to analyze your data further. One procedure would be to consider first just the males and perform a one-way analysis of variance on the male data, determining the effect of age on rt scores (this is called the "simple effect" of age – it is not the main effect of age, because main effects apply to the whole data; a simple effect is a "main effect" for part of the data; in this case, just the male data). If this one-way is significant, you would interpret it by looking at the pattern given by Newman-Keuls. Then look at just the female data, perform a one-way on the rt scores (simple effect of age for the females), and look at the Newman-Keuls if that one-way is significant.

Another procedure would be consider first just the young people, and perform a t-test contrasting the two sexes. Then consider the middle-aged people and use a t-test to test for the effect of sex, and then again for the old people.

In general, you should not need to perform both procedures. Think about which way of considering the interaction (either: a different effect of age for each sex; or: a different effect of sex for each age) best captures what is interesting about the data and follow the corresponding procedure. Maybe your theory makes clear predictions about both ways of considering the interaction, in which case you may analyse both ways.

In order to conduct these simple effects you need to tell SPSS to consider only part of the data. Go to the "Data" menu on the main menu bar at the top of the screen. Click on "Select cases". Click on "If condition is satisfied". The "If…" below it will then turn from faded to bold. Click on "If…". Enter "sex = 1" in the dialog box. Click "Continue", then "OK". Now any further analyses by SPSS will only be on the data where sex = 1; i.e. only for the males. Now if you click on "Analyze", "Compare means" and perform a one-way with rt as dependent variable and age as independent variable, you can determine the simple effect of age for the males. Then return to the "select cases" box and set the if condition as "if sex = 2", and you can determine the simple effect of age for the females.

Let us say you wanted to analyse the interaction as a different effect of sex at different levels of age. Then use "Select Cases" to select all cases where "age = 1". Having set up this if condition, click on "Statistics", then "Compare Means", then "Independent Samples T-test…". Enter "rt" as the "Test Variable" (i.e. the dependent variable) and "Sex" as the "Grouping Variable" (i.e. the independent variable). You then need to click on "Define Groups", entering "1" for Group One and "2" for Group Two, and then click "Continue". Click on "OK". The t-test will then be performed. SPSS also tests whether the variances are equal in each group (Levene's test for "equality of variances"). If Levene's test is significant, it means the variances do differ significantly. In looking up your t-test in SPSS' table, you should then use the row entitled "Equal variances not assumed". If Levene's test is non-significant, you can use the row entitled "Equal variances assumed". Remember when reporting a t-test, you report it as, for example: "An unrelated t-test on the reaction times of the young subjects indicated men were significantly slower than women, $t(27) = 4.56$, $p = .0012$." The number in brackets after the t is its degrees of freedom, labelled "df" in the printout. Repeat the procedure selecting "age=2" and finally "age=3".

You would only perform these simple effects analyses if your interaction was significant. Do not conduct them if you have a non-significant interaction.

An example Results section
This is based on different data to the previous example.

Table 1 displays the means and standard deviations of the reaction times (in milliseconds) of males and females of different ages.

Table 1.

	Males	Females
Young	554 (19)	535 (7)
Middle-aged	605 (7)	530 (14)
Old	723 (16)	755 (7)

NB: Standard deviations appear in parentheses.

Inspection of stem-and-leaf plots indicated that the data were roughly normally distributed in each group. The variances were not entirely homogenous.

A 3 X 2 (age(young vs middle-aged vs old) by sex(male vs female)) between-subjects ANOVA was conducted on the reaction time scores. There were significant main effects of age, $F_{(2,6)} = 283.87$, $p < .0001$, and of sex, $F_{(1,6)} = 7.91$, $p = .031$. The main effect of age was not analysed by further post hoc tests, because it was qualified by a significant interaction, $F_{(2,6)} = 17.97$, $p = .003$.

Note: the sentence starts with a "3 X 2" read as a "three by two". The number of numbers indicates the number of independent variables. The first number refers to the number of levels of the first independent variable, and the second number to the number of levels of the second independent variable. The names of the independent variables and their levels then follows in brackets.
Report the statistics (F etc) for all effects, even if nonsignificant.
Newman-Keuls is one of a number of test that can be conducted only after a significant F, and so is called a post hoc test. In this case, the main effect of age was not analysed by a Newman-Keuls test, or any other post hoc test. It would not be wrong to analyse it, but it may not be particularly meaningful to do so: the interaction says that the effect of age is different for males and females, and in this case the researcher was not interested in the average of two quite different effects. Whether you are interested or not just depends on the purposes of your particular research.
Note also, as stated previously, Newman-Keuls is only useful if your independent variable has MORE than two levels. If it has only two levels then you have already found out by the significant main effect everything you need to know – there is a difference!
Finally, a point of terminology – a main effect can only be "qualified" by an interaction if the main effect is significant.

The interaction was analysed further by considering the simple effect of age for males and females separately. For the males, the effect of age was significant, $F_{(2,3)} = 69.01$, $p = .0031$. A Newman-Keuls test indicated that all groups were significantly different from each other. In general, reaction time increased with age. For the females, the effect of age was also significant, $F_{(2,3)} = 330.17$, $p = .0003$. A Newman-Keuls indicated that the old subjects were slower than the other two ages, which did not differ amongst themselves. In summary, the age by sex interaction arose because reaction times began to slow down at a later age in women rather than men.