

**RESEARCH METHODS II Spring Term 2002:  
three-way and higher order ANOVA**

Imagine you ran an experiment similar to the one you ran in the second module, where you had independent variables of text type (e.g. romance vs football) and gender (males vs females). This time, however, you are interested in the effects of age as well; specifically, in whether the extent to which different genders have different schemata (preferences and knowledge) has changed over the last 20 years. So you run people aged around 40 as well as people aged around 20, and this constitutes a third independent variable (young vs middle-aged). You now have a three way 2 X 2 X 2 between-subjects design.

If you take just the middle-aged people, the extent of the interaction sex by text type gives you a measure of differential gender schemata (of the domains romance and football). For example, if the difference in comprehension between the texts (romantic comprehension score minus football comprehension score) was + 5 for the females and -7 for the men, the extent of the interaction (the difference of differences) is  $5 - (-7) = +12$ . Let us say that this is significant,  $p < .05$ . This is our evidence that middle-aged people have developed systematically different schemata according to gender.

Maybe these days, however, children are not being so strongly socialized into different gender roles as compared to twenty years ago. Anyway, you wish to test the extent to which this is true. If you take just the young people, you might find that the extent of the interaction sex by text type is, say, +5, which is non-significant. What can you conclude? Hopefully, you will resist the temptation to conclude there has been a change in socialization patterns! That is because before you can conclude there is a change you must specifically test for a change. That is, you must specifically test whether the +12 of the middle-aged people is significantly different from the +5 of the young people.

The logic here is just the same as the logic we went over for the two-way interaction. If women understand romantic texts better than football texts to a significant degree, but the difference for men is not significant, this is NOT evidence that men and women have developed different schemata. The men may miss out on significance purely because statistical power is less than 1 (and power will always be less than 1). Power is the probability that you get a significant effect ASSUMING there really is a population effect to detect. A typical power value for psychology experiments is 0.5. So both men and women may understand romance better than football, but it just so happened that the results were significant for the women but not the men. That is why you must test the two way interaction sex by text type if you want to know whether men and women have developed different schemata. If that interaction is significant you have shown that there are differential schemata according to gender.

In just the same way, it may be that middle-aged people and young people have been socialized into the same sex roles to the same degree, and it was just power considerations that meant the two-way interaction was significant for the middle-aged people but not the young people. That is why you must test whether the +12 for the middle-aged people is significantly different from the +5 of the young people. That is, you must test whether there is a difference of difference of differences! This is the three-way interaction text type by gender by age.

In fact, three-way designs come up rather frequently in psychology, and could well come up in your third year project. For this session, I would like you to make up data for an experiment with three independent variables. You could imagine them being the same three we have just discussed, or come up with your own example. For now, make all the variables between-subjects.

Use one column for your dependent variable, and three columns for your three independent variables. Once you have entered your data, click on "Analyze" on the SPSS controls at the top of the screen. Then click on "General Linear Model" on the menu that appears. Then click on "Univariate" on the next menu that appears, and carry on as in the second module, but specifying three independent variables.

Now in your output you will have tests of significance for:

Three main effects (text type, sex, age);  
Three two-way interactions (text type X age; text type X sex; sex by age);  
and one three-way interaction (text type X age X sex).

These effects can be interpreted as follows. A main effect of sex means you average over the other two variables – text type and age – to get one overall average for females and one overall average for males. The main effect of sex is testing whether females have an overall different level of comprehension from males.

A two-way interaction text type by sex means you average over age now, and ask if the difference in comprehension

between the texts for females (taking both ages together) is different from the difference in comprehension between the texts for males (taking both ages together). (Note that if in the population there is a two-way interaction for middle-aged people but not for young people, then there is a population two-way interaction overall.)

Finally the three-way interaction, means that that the two-way interaction text type by sex is different for the different ages. But just as there are two ways of breaking down a two-way interaction, there are three ways of breaking down a three-way interaction! Thus, a three-way interaction also means that the two-way interaction text type by age is different for the different sexes; and also that the two-way interaction age by sex is different for the different text types.

### **Interpreting interactions:**

#### **1. Significant three way interaction**

If the three way interaction is significant, this means that the two-way interaction sex by text type was different for the different ages. To try to get some information on what this difference might be, conduct two separate two-way ANOVAs. First select just the young people, by clicking on "Data", "Select cases", "If condition is satisfied", "If...". Enter "age = 1" in the dialogue box, or whatever your code is. Then, as before, click on "Analyze", "General Linear Model" and then "Univariate". Now tell SPSS you have two factors, sex and text type. In the output you are ONLY interested in one effect: The two way interaction. Ignore the main effects.

If this two way interaction is significant, you can try to obtain information about its interpretation (just as you did in module 2) by analyzing the effect of e.g. gender separately for each text type (for just the young people). Now to select the relevant cases you would have to set the selection condition to e.g. "(age=1) and (text=1)".

Having analyzed the two-way interaction for young people, you would test the two-way interaction for middle-aged people. If that interaction was significant, you would look at simple effects to interpret it.

Notice the logic of the procedure here. To interpret a three-way interaction, you break it down into several two-way interactions. These are called partial interactions, because they are interactions calculated on only part of the data. To interpret a two-way interaction, you break it down into several one-way effects (i.e. simple effects). In general, an n-way interaction can be interpreted by breaking it down into its component (n - 1)-way partial interactions. These partial interactions can be interpreted in turn by breaking them down into (n - 2)-way partial interactions, and so on.

#### **2. Significant two-way interaction**

Let us say that in the three-way ANOVA the interaction sex by text type was significant. You may not wish to analyze this interaction further if you have a significant three-way interaction (because the three-way interaction tells that this two-way interaction is different for different ages: So maybe you don't want to try to interpret the average of two quite different things). However, in case you decide it would be interesting to interpret the two-way interaction.

You break down the two-way into (1) a simple effect of text type for women (averaging over both ages) and (2) a simple effect of text type for men (averaging over both ages). To test (1) use select cases to select just the women. Then perform a two-way ANOVA age by text type, but in the output you are ONLY interested in the main effect of text type. To test (2), use select cases to select just the men, and repeat the procedure.

Note what you have done: You have interpreted a two-way interaction by breaking it down into two one-ways.

### **Planned vs post hoc tests**

For your third-year project bear the following in mind: A rule of thumb in analyzing n-way designs is that you are allowed to analyze (n - 1)-way partial interactions ONLY IF the n-way interaction is significant. The analyses that you perform only when some other analysis is significant are called **post-hoc analyses**. Newman-Keuls is an example of a post hoc analysis. For example, following this rule for the current data set, you could only test the two-way interaction sex by text type for just young people if the three-way interaction was significant. In a similar way, you can only test for simple effects if the two-way interaction was significant.

But this is just a rule of thumb. You can violate it when you wish to perform just one or two tests that are strongly predicted in advance of collecting the data and which provided the whole *raison d'être* of the experiment (such tests are called **planned comparisons**). (For example, if the whole point of the experiment was to investigate the two-way interaction text-type by sex for middle-aged people, and you just threw in some young people out of curiosity, you could test the partial two-way interaction text-type by sex on just the middle-aged people, regardless of whether the three way interaction was significant.)

**Repeated measures**

Now as an exercise, make up a new data set with three independent variables of which all or some are within-subjects variables, and repeat the above procedures, with necessary changes given that its now repeated measures. Remember all the data for one subject should appear on one row.

**Discussion**

In discussing results, ONLY discuss effects that you have explicitly tested. If you obtained a two way interaction text type by gender, and you broke it down into simple effects of gender for different text types, then in the discussion DO NOT talk about why men may understand football differently from romance. You could only talk about why men may understand romance differently from women. Similarly, if text type is significant for men but not women (say in different experiments reported in the literature) but the two-way interaction has not been tested, or is not significant, DO NOT talk about why men may have different schemata from women. You could only talk about why women may understand football differently from romance.