Explaining Interaction

February 4, 2008

For example: $a = 4, b = 3$.

- Factor A: levels $1, 2, \ldots, a$
- Factor B: levels $1, 2, \ldots, b$

Most complicated model: Main Effects plus Interaction

$$\text{A} + \text{B} + \text{A.B}$$

that is, we have

- a baseline mean: β_0
- an effect for each level of Factor A: $\beta_i^{(A)}$
- an effect for each level of Factor B: $\beta_j^{(B)}$
- an interaction that modifies the effect of changing levels of Factor A at each level of Factor B: $\gamma_{ij}^{(AB)}$

In SPSS, the baseline group is the one where Factor A has level *a* and Factor B has level *b*, but this choice is **arbitrary**; changing this assumption should have no effect on the results we obtain.

Thus we adopt the following modelling strategy:

- Establish a baseline
- Look for changes from baseline introduced by Factor A
- Look for changes from baseline introduced by Factor B
- Look for changes from baseline introduced by Factor A and Factor B **additively**, so that the effect of changing the level of Factor A is **identical** in each level of Factor B, and *vice versa*).
- Look for changes from baseline introduced by Factor A and Factor B **additively with interaction**, so that the effect of changing the level of Factor A is **different** in each level of Factor B, and *vice versa*).

Two-way table: 4×3

Factor B

Factor A	1	2	3
1			
2			
3			
4			

Null Model : Baseline Mean Only

Null Model: cell entries are means for data for each treatment.

Factor B

Factor A	1	2	3
1	β_0	β_0	β_0
2	β_0	β_0	β_0
3	β_0	β_0	β_0
4	β_0	β_0	β_0

Effect of Factor A only

Main Effect Only: A

Factor B

Factor A

	1	2	3
1	$\beta_0 + \beta_1^{(A)}$	$\beta_0 + \beta_1^{(A)}$	$\beta_0 + \beta_1^{(A)}$
2	$\beta_0 + \beta_2^{(A)}$	$\beta_0 + \beta_2^{(A)}$	$\beta_0 + \beta_2^{(A)}$
3	$\beta_0 + \beta_3^{(A)}$	$\beta_0 + \beta_3^{(A)}$	$\beta_0 + \beta_3^{(A)}$
4	β_0	β_0	β_0

Effect of Factor B only

Main Effect Only: B

Factor B

Factor A	1	2	3
1	$\beta_0 + \beta_1^{(B)}$	$\beta_0 + \beta_2^{(B)}$	β_0
2	$\beta_0 + \beta_1^{(B)}$	$\beta_0 + \beta_2^{(B)}$	β_0
3	$\beta_0 + \beta_1^{(B)}$	$\beta_0 + \beta_2^{(B)}$	β_0
4	$\beta_0 + \beta_1^{(B)}$	$\beta_0 + \beta_2^{(B)}$	β_0

Effect of Factor A plus Effect of Factor B

Main Effects Only: A + B

Factor B

Factor A

	1	2	3
1	$\beta_0 + \beta_1^{(A)} + \beta_1^{(B)}$	$\beta_0 + \beta_1^{(A)} + \beta_2^{(B)}$	$\beta_0 + \beta_1^{(A)}$
2	$\beta_0 + \beta_2^{(A)} + \beta_1^{(B)}$	$\beta_0 + \beta_2^{(A)} + \beta_2^{(B)}$	$\beta_0 + \beta_2^{(A)}$
3	$\beta_0 + \beta_3^{(A)} + \beta_1^{(B)}$	$\beta_0 + \beta_3^{(A)} + \beta_2^{(B)}$	$\beta_0 + \beta_3^{(A)}$
4	$\beta_0 + \beta_1^{(B)}$	$\beta_0 + \beta_2^{(B)}$	β_0

Main effects plus Interaction between A and B

Main Effects Plus Interaction: A + B + A.B

Factor B

Factor A

	1	2	3
1	$\beta_0 + \beta_1^{(A)} + \beta_1^{(B)} + \gamma_{11}^{(AB)}$	$\beta_0 + \beta_1^{(A)} + \beta_2^{(B)} + \gamma_{12}^{(AB)}$	$\beta_0 + \beta_1^{(A)}$
2	$\beta_0 + \beta_2^{(A)} + \beta_1^{(B)} + \gamma_{21}^{(AB)}$	$\beta_0 + \beta_2^{(A)} + \beta_2^{(B)} + \gamma_{22}^{(AB)}$	$\beta_0 + \beta_2^{(A)}$
3	$\beta_0 + \beta_3^{(A)} + \beta_1^{(B)} + \gamma_{31}^{(AB)}$	$\beta_0 + \beta_3^{(A)} + \beta_2^{(B)} + \gamma_{32}^{(AB)}$	$\beta_0 + \beta_3^{(A)}$
4	$\beta_0 + \beta_1^{(B)}$	$\beta_0 + \beta_2^{(B)}$	β_0

Q. Why are the following models

- A.B
- A + A.B
- B + A.B

not considered ?

A. Because they make specific and perhaps **unrealistic** assumptions about the data, and they imply that the levels of the factors are **not arbitrarily labelled**.

Q. Why are the following models

- A.B
- A + A.B
- B + A.B

not considered ?

A. Because they make specific and perhaps **unrealistic** assumptions about the data, and they imply that the levels of the factors are **not arbitrarily labelled**.

SPSS will not fit such models, although it appears that it does !

Recall the definition of interaction:

- Variation in the effect of changing levels of one factor at the different levels of the other factor.
- For example, the effect on the response mean of moving from level 1 to level 2 for Factor B is **different** at different levels of Factor A.

Consider the model

$$A.B$$

this model implies that all parameters apart from the **baseline** and the **interaction** parameters are zero.

Interaction between A and B only

Interaction only: A.B

Factor B

Factor A

	1	2	3
1	$\beta_0 + 0 + 0 + \gamma_{11}^{(AB)}$	$\beta_0 + 0 + 0 + \gamma_{12}^{(AB)}$	$\beta_0 + 0$
2	$\beta_0 + 0 + 0 + \gamma_{21}^{(AB)}$	$\beta_0 + 0 + 0 + \gamma_{22}^{(AB)}$	$\beta_0 + 0$
3	$\beta_0 + 0 + 0 + \gamma_{31}^{(AB)}$	$\beta_0 + 0 + 0 + \gamma_{32}^{(AB)}$	$\beta_0 + 0$
4	$\beta_0 \quad + 0$	$\beta_0 \quad + 0$	β_0

In this set-up,

- for Factor A, Level 4: the effect of moving from Level 3 to Level 2 of factor B is **zero**
- for Factor A, Level 3: the effect of moving from Level 3 to Level 2 of factor B is $\gamma_{32}^{(AB)}$.

Therefore, there is a **fundamental difference** between the way that we regard the levels of Factor A.

Main Effect of A plus Interaction between A and B only

Interaction only: A + A.B

Factor B

Factor A	1	2	3
1	$\beta_0 + \beta_1^{(A)} + 0 + \gamma_{11}^{(AB)}$	$\beta_0 + \beta_1^{(A)} + 0 + \gamma_{12}^{(AB)}$	$\beta_0 + \beta_1^{(A)}$
2	$\beta_0 + \beta_2^{(A)} + 0 + \gamma_{21}^{(AB)}$	$\beta_0 + \beta_2^{(A)} + 0 + \gamma_{22}^{(AB)}$	$\beta_0 + \beta_2^{(A)}$
3	$\beta_0 + \beta_3^{(A)} + 0 + \gamma_{31}^{(AB)}$	$\beta_0 + \beta_3^{(A)} + 0 + \gamma_{32}^{(AB)}$	$\beta_0 + \beta_3^{(A)}$
4	$\beta_0 + 0$	$\beta_0 + 0$	β_0

In this set-up,

- for Factor A, Level 4: the effect of moving from Level 3 to Level 2 of factor B is **zero**
- for Factor A, Level 3: the effect of moving from Level 3 to Level 2 of factor B is $\gamma_{32}^{(AB)}$.

Therefore, there is a **fundamental difference** between the way that we regard the levels of Factor A. If we rearrange the labels of the levels of Factor A

we may get a different result.

Therefore, although it is possible **in general** to fit such models, it is no longer possible to talk of the effect of "Factor A".

How does SPSS Handle Such Models ?

It is possible to fit the models

$$A + A.B \qquad B + A.B \qquad A.B$$

in SPSS. For example, for the model A+A.B

- *Analyze* $\longrightarrow$ *General Linear Model* $\longrightarrow$ *Univariate*
- Select the *Dependent Variable* and *Fixed Factor(s)*
- Click *Model* to bring up the *Univariate: Model* dialog box.
- Select Factor A as a **Main Effect** using the *Build* pull-down list, click the selection arrow,
- highlight Factor A and Factor B simultaneously, and select **Interaction** from the *Build* pull-down list, and click the selection arrow.
- Click *Continue*, and then *OK*.

How does SPSS Handle Such Models ?

This produces the usual ANOVA table, with terms including

Factor A

and

Factor A * Factor B

However, in fact the model

$$A + B + A.B$$

has been fitted !

- The results are just reported differently
- The terms B and A.B are reported together !

Example: Batteries Data

A - Material
B - Temperature
Model A + B + A.B

Dependent Variable: Battery Life					
Source	Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	59154.000	8	7394.250	11.103	0.000
Intercept	398792.250	1	398792.250	598.829	0.000
material	10633.167	2	5316.583	7.983	0.002
temp	39083.167	2	19541.583	29.344	0.000
material * temp	9437.667	4	2359.417	3.543	0.019
Error	17980.750	27	665.954		
Total	475927.000	36			
Corrected Total	77134.750	35			
R Squared = .767 (Adjusted R Squared = .698)					

$$SS = SST_A + SST_B + SSI_{AB} + SSE$$

Example: Batteries Data

A - Material

B - Temperature

Model A + A.B

Dependent Variable: Battery Life					
Source	Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	59154.000	8	7394.250	11.103	0.000
Intercept	398792.250	1	398792.250	598.829	0.000
material	10633.167	2	5316.583	7.983	0.002
material * temp	48520.833	6	8086.806	12.143	0.000
Error	17980.750	27	665.954		
Total	475927.000	36			
Corrected Total	77134.750	35			
R Squared = .767 (Adjusted R Squared = .698)					

$$SS = SST_A + SSI_{B:AB} + SSE$$

where

$$SSI_{B:AB} = SST_B + SSI_{AB}$$