

```
data STDp391; *Subsamples in a factorial RCBD;
input D R Block plot number @@;
cards;
```

3	0	1	1	14.7	3	4	1	1	8.8	3	8	1	1	6.9	3	0	1	1	16.7	3	4	1	1	10.8	3	8	1	1	8.9
3	0	2	1	13.6	3	4	2	1	13.6	3	8	2	1	9.3	3	0	2	1	15.6	3	4	2	1	15.6	3	8	2	1	11.3
3	0	3	1	15.5	3	4	3	1	10.9	3	8	3	1	8.7	3	0	3	1	17.5	3	4	3	1	12.9	3	8	3	1	10.7
3	0	4	1	13.7	3	4	4	1	11.4	3	8	4	1	8.6	3	0	4	1	15.7	3	4	4	1	13.4	3	8	4	1	10.6
10	0	1	1	17.0	10	4	1	1	12.6	10	8	1	1	7.8	10	0	1	1	19.0	10	4	1	1	14.6	10	8	1	1	9.8
10	0	2	1	16.4	10	4	2	1	9.6	10	8	2	1	7.2	10	0	2	1	18.4	10	4	2	1	11.6	10	8	2	1	9.2
10	0	3	1	14.1	10	4	3	1	10.8	10	8	3	1	10.3	10	0	3	1	16.1	10	4	3	1	12.8	10	8	3	1	12.3
10	0	4	1	13.4	10	4	4	1	12.3	10	8	4	1	10.2	10	0	4	1	15.4	10	4	4	1	14.3	10	8	4	1	12.2

```
;
proc GLM;
class D R Block plot;
model number= Block D R D*R plot(D*R*Block);
random plot(D*R*Block);
test h= D e= plot(D*R*Block);
test h= R e= plot(D*R*Block);
test h= D*R e= plot(D*R*Block);
proc varcomp Method= Typel;
class D R Block plot;
model number= Block D R D*R plot(D*R*Block);
run; quit;
```

Dependent Variable: number

Source	DF	SS	MS	F	Pr > F
Model	23	391.2	17.0	8.51	<.0001
Error	24	48.0	2.0		
Corrected Total	47	439.2			

Source	DF	SS	MS	F	Pr > F
Block	3	1.16	0.39	0.19	0.90
D	1	3.00	3.00	1.50	0.23
R	2	307.33	153.66	76.83	<.0001
D*R	2	0.98	0.49	0.24	0.7846
plot(D*R*Block)	15	78.77	5.25	2.63	0.0170

Tests of Hypotheses Using MS for plot(D*R*Block) as Error Term

Source	DF	SS	MS	F Value	Pr > F
D	1	3.00	3.00	0.57	0.4614
R	2	307.33	153.66	29.26	<.0001
D*R	2	0.98	0.49	0.09	0.9114

Variance Component	Estimate	%	Plot= \$50 Subsample= \$5
Var(Block)	-0.40528	0	Optimum allocation
Var(D)	0.10458	1	SQRT[(50*1.62)/(5*2)]= 2.84
Var(R)	9.57333	72	Use 3 subsamples
Var(D*R)	-0.59514	0	
Var(plot(D*R*Block))	1.62556	12	
Var(Error)	2.00000	15	

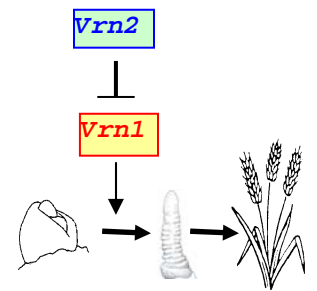
$$N_s = \sqrt{\frac{C_{e.u.} * S_{SUB}^2}{C_{SUB} * S_{e.u.}^2}}$$

Partition of interaction example: effect of *Vrn1* and *Vrn2* genes on flowering.

Each plant from a segregating population from a cross between parents A and B (N=102) was characterized with molecular markers and the number of alleles of parent A indicated (BB= 0, AB=1, AA=2).

The auxiliary variable “**type**” = each combination of *Vrn1* and *Vrn2* classes.

```
data interpart;
input type Vrn1 Vrn2 days;
cards;
1 0 0 89      1 0 0 97      1 0 0 101     1 0 0 100
1 0 0 98      2 0 1 133     2 0 1 144     2 0 1 148
2 0 1 148     2 0 1 138     2 0 1 130     2 0 1 133
2 0 1 128     2 0 1 130     2 0 1 137     2 0 1 141
2 0 1 134     2 0 1 133     2 0 1 138     2 0 1 131
2 0 1 148     3 0 2 163     3 0 2 153     3 0 2 161
3 0 2 153     3 0 2 156     3 0 2 148     4 1 0 109
4 1 0 83      4 1 0 87      4 1 0 103     4 1 0 110
4 1 0 81      4 1 0 99      4 1 0 98      4 1 0 83
4 1 0 78      4 1 0 92      4 1 0 92      4 1 0 91
4 1 0 85      4 1 0 83      4 1 0 66      5 1 1 122
5 1 1 121     5 1 1 121     5 1 1 122     5 1 1 125
5 1 1 118     5 1 1 123     5 1 1 124     5 1 1 125
5 1 1 108     5 1 1 112     5 1 1 126     5 1 1 118
5 1 1 98      5 1 1 116     5 1 1 106     5 1 1 117
5 1 1 110     5 1 1 113     5 1 1 129     5 1 1 116
6 1 2 140     6 1 2 125     6 1 2 178     6 1 2 136
6 1 2 132     6 1 2 133     6 1 2 135     6 1 2 134
6 1 2 125     6 1 2 125     6 1 2 128     6 1 2 121
6 1 2 128     6 1 2 135     7 2 0 91      7 2 0 103
7 2 0 81      7 2 0 99      7 2 0 88      7 2 0 99
7 2 0 73      8 2 1 137     8 2 1 118     8 2 1 120
8 2 1 153     8 2 1 86      8 2 1 114     8 2 1 126
8 2 1 120     8 2 1 120     8 2 1 118     8 2 1 119
8 2 1 106     8 2 1 112     8 2 1 111     8 2 1 117
9 2 2 124     9 2 2 124
```



```
;
proc glm order=data;
class vrn1 vrn2;
model days= vrn1|vrn2;
contrast 'Lineal Vrn1'      vrn1 -1 0 1;
contrast 'Quadratic Vrn1'  vrn1 1 -2 1;
```

1	2	3	4	5	6	7	8	9	Type
0	0	0	1	1	1	2	2	2	Vrn1
0	1	2	0	1	2	0	1	2	Vrn2

```
proc glm order=data;
class type;
model days= type;
contrast 'Lineal Vrn1' Type -1 -1 -1 0 0 0 1 1 1;
contrast 'Quadrat Vrn1' Type 1 1 1 -2 -2 -2 1 1 1;
contrast 'Lineal Vrn2' Type -1 0 1 -1 0 1 -1 0 1;
contrast 'Quadrat Vrn2' Type 1 -2 1 1 -2 1 1 -2 1;
contrast 'Int l by l' Type 1 0 -1 0 0 0 -1 0 1;
contrast 'Int l by q' Type -1 2 -1 0 0 0 1 -2 1;
contrast 'Int q by l' Type -1 0 1 2 0 -2 -1 0 1;
contrast 'Int q by q' Type 1 -2 1 -2 4 -2 1 -2 1;
run; quit;
```

3x3 Factorial

Class	Levels	Values
Vrn1	3	0 1 2
Vrn2	3	0 1 2

Source	DF	SS	MS	F Value	Pr > F
Model	8	38006	4751	42.97	<.0001
Error	93	10282	111		
Corrected Total	101	48288			

Source	DF	Type III SS	MS	F Value	Pr > F
Vrn1	2	4435	2217	20.06	<.0001
Vrn2	2	21310	10655	96.37	<.0001
Vrn1*Vrn2	4	808	202	1.83	0.1303 NS

Contrast	DF	SS	MS	F Value	Pr > F
Lineal Vrn1	1	2829	2829	25.58	<.0001
Quadrat Vrn1	1	847	847	7.66	0.0068

Partition of interaction using one way ANOVA and contrasts

Class	Levels	Values
type	9	1 2 3 4 5 6 7 8 9

Source	DF	SS	MS	F Value	Pr > F
Type	8	38006	4751	42.97	<.0001
Error	93	10282	111		
Corrected Total	101	48288			

Contrast	DF	SS	MS	F Value	Pr > F
Lineal Vrn1	1	2829	2829	25.58	<.0001
Quadrat Vrn1	1	847	847	7.66	0.0068
Lineal Vrn2	1	16181	16181	146.35	<.0001
Quadrat Vrn2	1	1650	1650	14.92	0.0002
Int 1 by 1	1	631	631	5.71	0.0189
Int 1 by q	1	0	0	0.00	0.9523
Int q by 1	1	12	12	0.11	0.7465
Int q by q	1	161	161	1.46	0.2305

Note that even though the interaction in the 3x3 factorial is not significant, **the lineal by lineal interaction is significant.**

Note also that the Lineal and Quadratic contrast for the **main Vrn1** are identical in both analyses.

Testing simple effects after significant interaction (ST&D, p. 358)

Factor A in this experiment is time of bleeding of a lamb, and Factor B is treatment vs. no treatment with estrogen. Variable: phosphorous.

Here are the treatment totals of the 5 replications

Factor	A= time			
	Level	(a1)= A.M.	(a2)= P.M.	Total
B= estrogen	(b1)= control	Mean of 5 obs.: 66.39	Mean of 5 obs.: 182.67	249.06
	(b2)= treated	Mean of 5 obs.: 96.80	Mean of 5 obs.: 139.06	235.86
	Total	163.19	321.73	484.92

SAS analysis

```

data fact1;
input id time $ estgn $ phos @@;
cards;
1 am c 8.53 2 am t 17.53 3 pm c 39.14 4 pm t 32.00
1 am c 20.53 2 am t 21.07 3 pm c 26.20 4 pm t 23.80
1 am c 12.53 2 am t 20.80 3 pm c 31.33 4 pm t 28.87
1 am c 14.00 2 am t 17.33 3 pm c 45.80 4 pm t 25.06
1 am c 10.80 2 am t 20.07 3 pm c 40.20 4 pm t 29.33
;
proc glm;
class time estgn;
model phos=time|estgn;
run; quit;

```

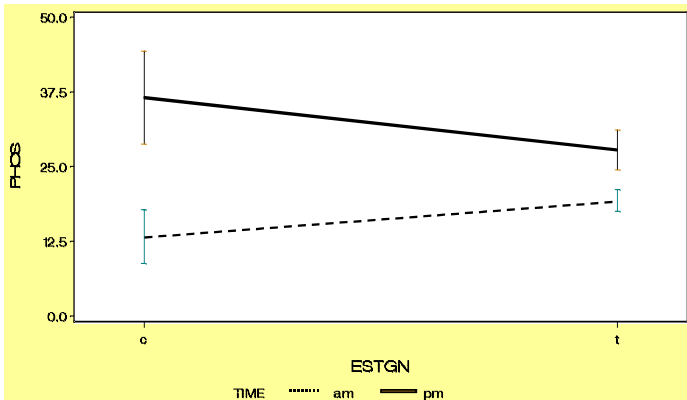
OUTPUT

First PROC GLM

Dependent Variable: PHOS

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TIME	1	1256.74658	1256.74658	52.93	0.0001
ESTGN	1	8.71200	8.71200	0.37	0.5532
TIME*ESTGN	1	273.94802	273.94802	11.54	0.0037
ERROR	16	379.92000	23.75000		

A significant interaction ($P=0.0037$) indicates **heterogeneous** simple effects



Non-parallel lines, as those observed in this graphic indicate **interaction**

If interactions are present in a fixed-effects model the next step is the analysis of the simple effects.

One general way of testing the simple effects is using the **by** statement
Use **proc sort** before, to sort by the variable used in the by statement

```
proc sort;
  by time;
proc glm;
  class estgn;
  model phos= estgn;
  means estgn / Hovtest= Levene;
  by time;
proc sort;
  by estgn;
proc glm;
  class time;
  model phos=time;
  means time / Hovtest= Levene;
  by estgn;
run; quit;
```

You need to test the assumptions for each one way ANOVA!

One Way ANOVAS	DF	Contrast SS	Pr > F
Between time within Control	1	1352.1	0.0004
Between time within treated	1	178.6	0.0011
Between estrogen level am	1	92.5	0.0237
Between estrogen level pm	1	190.2	0.0495

You can use 4 **CONTRAST** but you need to **adjust α** to control the MEER (3 df)

With $\alpha=0.05$ MEER= $1-(1-\alpha)^4=0.18$ With $\alpha=0.01$ MEER= $1-(1-\alpha)^4=0.04$

9. 8. Three way ANOVA (fixed-effects model)

3 or more factors may be analyzed simultaneously at different levels.

As the number of factors increases, the e.u. becomes very large.

Logistic difficulties with such large experiments: it may not be possible to run all the tests in one day or all the material in a chamber.

Treatments may be confounded with undesired effects if different treatments are applied under not quite the same experimental conditions.

Large number of possible interactions.

A 3-factor factorial has three 1st-**order interactions**, A X B, A X C, and B X C.; and a 2nd-**order interaction**, A X B X C.

The fixed model is assumed to be,

$$\mu_{ijk} = \mu + \mu_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk}$$

A 4-factor factorial has 6 first-order interactions, 4 2nd-order interactions, and one **third-order interaction** (A X B X C X D).

The testing of their significance, and more importantly, their interpretation becomes exceedingly complex.

Example of a three-way factorial ANOVA (Taken from: C.J. Monlezun.1979. Two-dimensional plots for interpreting interactions in the 3-factor analysis of variance model. The American Statistician 33:63-69.)

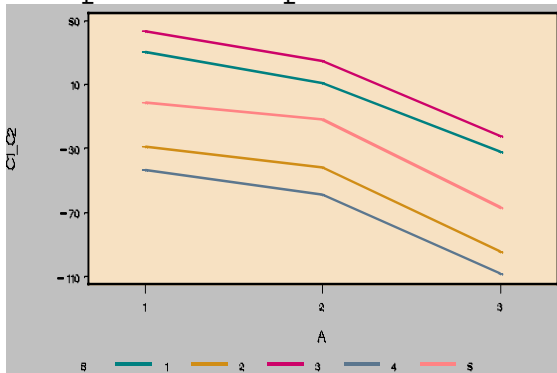
The following hypothetical population means for a 3x5x2 experiment are used to illustrate an example with **no three-way interactions**. A graphic technique to show the three way interactions is discussed.

	A1C1	A2C1	A3C1	A1C2	A2C2	A3C2
B1	61	38	81	31	27	113
B2	39	61	49	68	103	143
B3	121	82	41	78	57	63
B4	79	68	59	122	127	167
B5	91	31	61	92	43	128

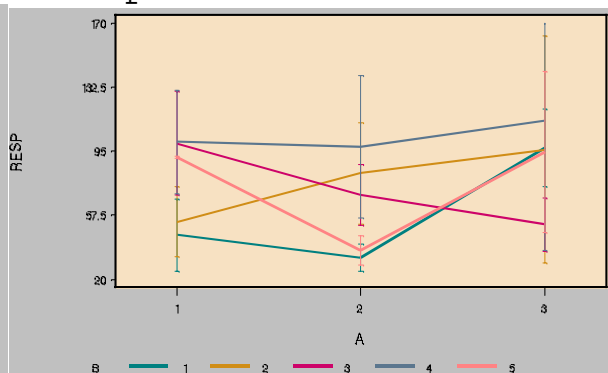
Three-way ANOVA. Analyst table for graphics of three way-interactions

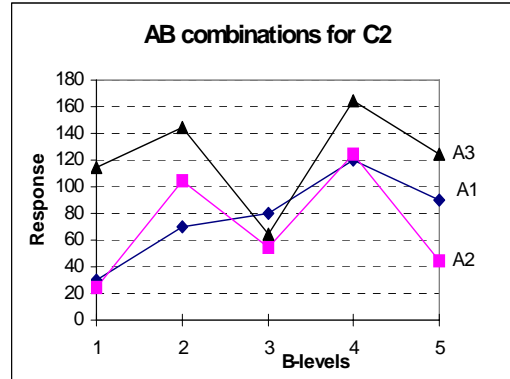
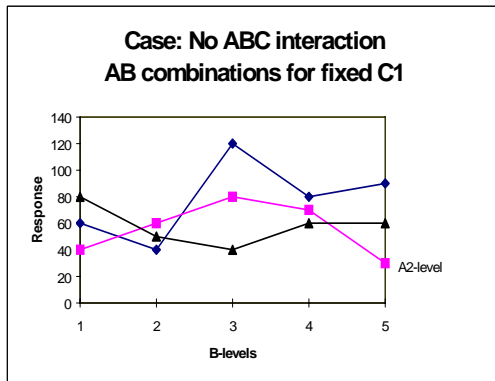
A	B	C	RESP	C1_C2
1	1	1	61	30
2	1	1	38	11
3	1	1	81	-32
1	1	2	31	30
2	1	2	27	11
3	1	2	113	-32
1	2	1	39	-29
2	2	1	61	-42
3	2	1	49	-94
1	2	2	68	-29
2	2	2	103	-42
3	2	2	143	-94
1	3	1	121	43
2	3	1	82	25
3	3	1	41	-22
1	3	2	78	43
2	3	2	57	25
3	3	2	63	-22
1	4	1	79	-43
2	4	1	68	-59
3	4	1	59	-108
1	4	2	122	-43
2	4	2	127	-59
3	4	2	167	-108
1	5	1	91	-1
2	5	1	31	-12
3	5	1	61	-67
1	5	2	92	-1
2	5	2	43	-12
3	5	2	128	-67

Resp C1- Resp C2 at AxB



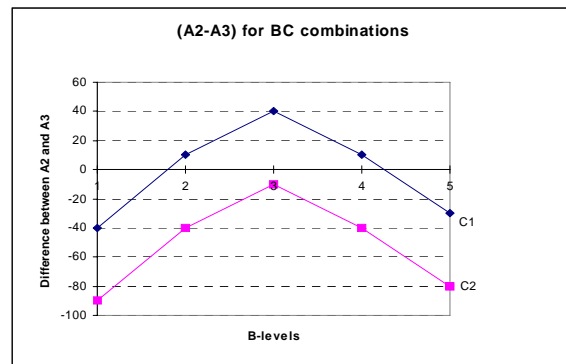
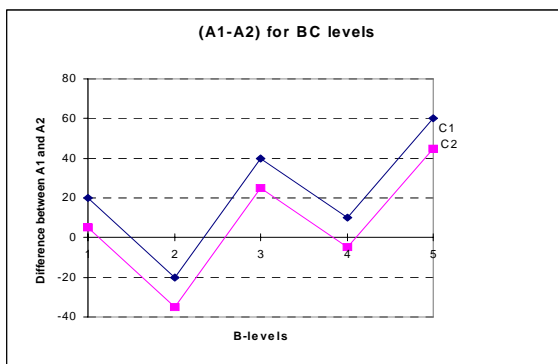
Resp AxB within 1C level





The lines of mean plots for fixed C1 (left) and C2 (right) levels are not parallel indicating a two-way interaction between A and B in both levels of C. The **1st order interaction** (AxB) now has **two values**: (AxB, c₁) and (AxB, c₂). The interaction term (AxB) is the average.

If, **the differences between different levels of A** are taken over levels of say, B, for the two different C levels, the plot of these differences reveals no interaction between BC. The lack of BC interaction with the differences between levels of A indicates that no ABC interaction is present in these means, i.e. $(\alpha\beta\gamma)_{ijk} = 0$.



The interpretation of a 3-factor interaction is that the effect of factor A depends on the precise combination of factors B and C.

For example if **A is nitrogen level** (0 or 3 cwt/a) and **B is plow depth** (7 or 11 in.). In a two-factor experiment, a significant AxB interaction indicates that the crop has a different response to N depending on plow depth. Now introduce the third factor **C, soil type** (loam or sand). **Then a nonzero (AxBxC) would mean that the difference in the response to N as a function of plow depth depends on the soil type.**