Today

- Two-Way Between-Subjects Factorial Designs
  - 2 x 2 design
  - concept of interaction
  - model comparison approach
  - controlling type-I error
  - follow-up tests

The 2 x 2 Design

- hypothetical study:
- explore effects of biofeedback and drug therapy on blood pressure
- one approach could be:
  - 1 factor, four groups:
  - (1) biofeedback + drug
    (2) biofeedback, no drug
    (3) no biofeedback + drug
    (4) no biofeedback, no drug

TABLE 7.1 Blood Pressure Data for 2 × 2 Factorial Design

	Group			
	1: Biofeedback and Drug	*2: Biofeedback Alone*	*3: Drug Alone*	*4: Neither*
	158	188	186	185
	163	183	191	190
	173	198	196	195
	178	178	181	200
	168	193	176	180
Mean	**168**	**188**	**186**	**190**
s	7.9057	7.9057	7.9057	7.9057

TABLE 7.2 ANOVA for Data in Table 7.1

Source	*SS*	*df*	*MS*	*F*	*p*
Between	1540.00	3	513.33	8.21	.002
Within	1000.00	16	62.50		
Total	**2540.00**	**19**			

- [+1 +1 -1 -1]: effect of biofeedback: $F=8.00$, $p < .05$
- [+1 -1 +1 -1]: effect of drug: $F=11.52$, $p < .05$
- our conclusion would be that
  - both drug and biofeedback have an effect

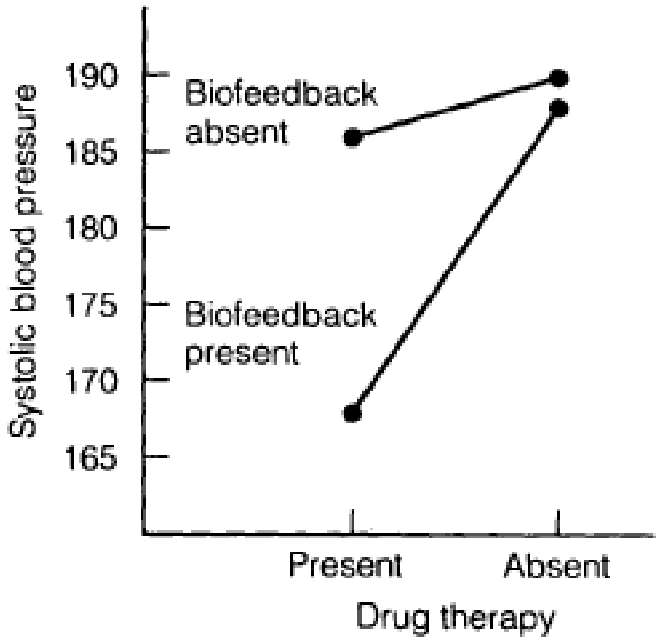
TABLE 7.3 Factorial Arrangement of Means from Table 7.1

		Biofeedback		
		Present	*Absent*	**Average**
Drug Therapy	*Present*	168	186	**177**
	Absent	188	190	**189**
	Average	**178**	**188**	**183**

- effect of drug therapy, averaged over levels of biofeedback
  - Present: 177
  - Absent: 189
  - F=11.52, p < .05; drug therapy has an effect on blood pressure
- effect of biofeedback, averaged over levels of drug therapy
  - Present: 178
  - Absent: 188
  - F=8.00, p < .05; biofeedback has an effect on blood pressure
- Is this an accurate representation of what's going on here?
- no! both main effects are driven by one cell
  - drug therapy + biofeedback

TABLE 7.3 Factorial Arrangement of Means from Table 7.1

		Biofeedback		
		Present	*Absent*	**Average**
Drug Therapy	*Present*	168	186	**177**
	Absent	188	190	**189**
	Average	**178**	**188**	**183**

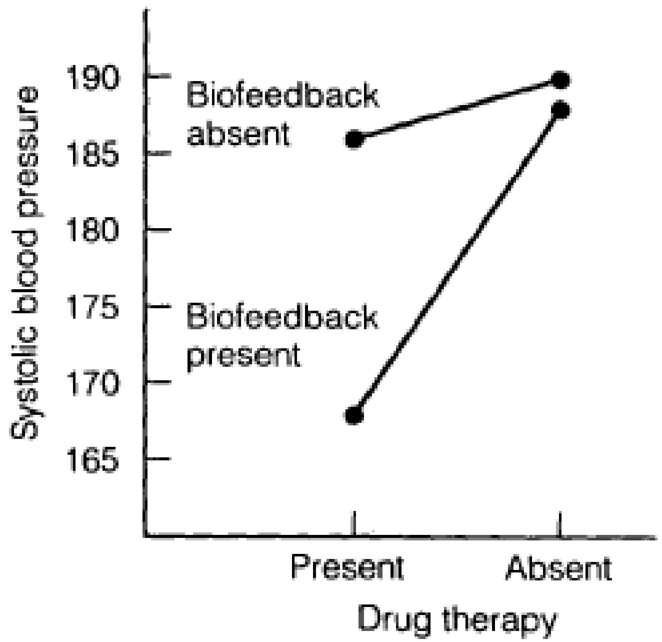


(a)

- there is an **interaction** between drug therapy and biofeedback
  - effect of drug therapy depends on the level of the biofeedback factor
  - effect of biofeedback depends on the level of the drug therapy factor
  - the level of biofeedback modulates the effect of drug therapy
  - the level of drug therapy modulates the effect of biofeedback

TABLE 7.3 Factorial Arrangement of Means from Table 7.1

		Biofeedback		
		Present	*Absent*	**Average**
Drug Therapy	*Present*	168	186	**177**
	Absent	188	190	**189**
	Average	**178**	**188**	**183**



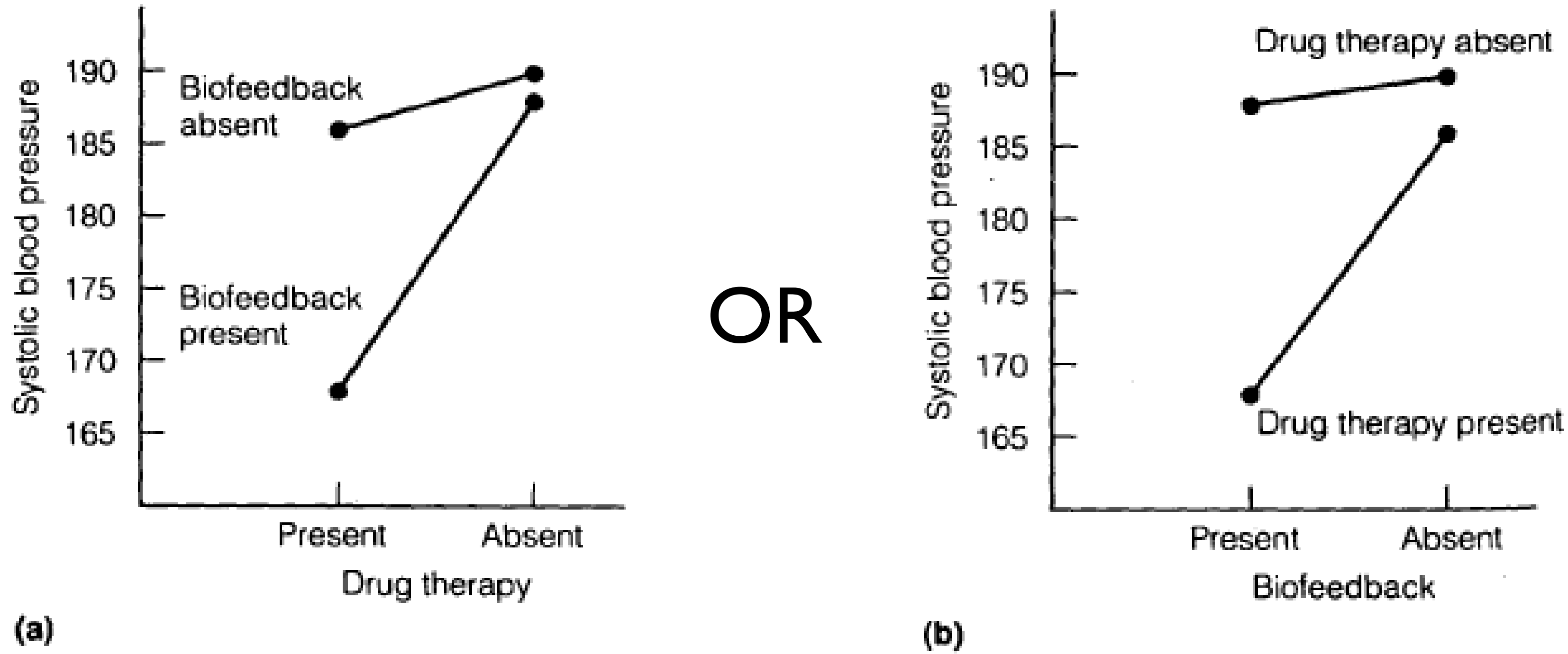
(a)

OR

- there is an **interaction** between drug therapy and biofeedback
  - effect of drug therapy depends on the level of the biofeedback factor
  - effect of biofeedback depends on the level of the drug therapy factor
  - the level of biofeedback modulates the effect of drug therapy
  - the level of drug therapy modulates the effect of biofeedback

TABLE 7.3 Factorial Arrangement of Means from Table 7.1

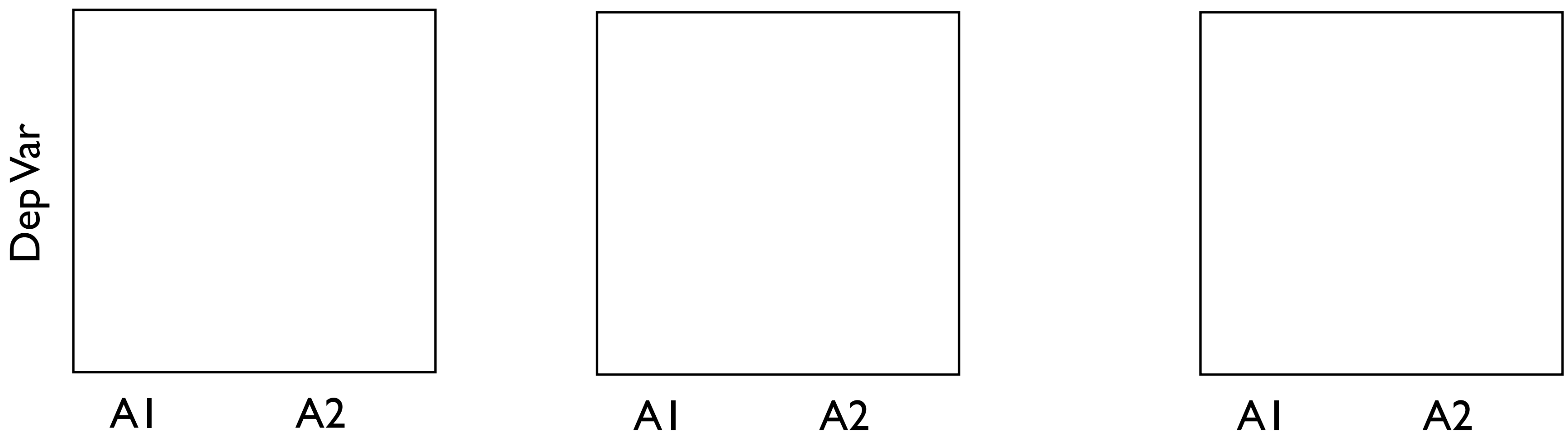
		Biofeedback		
		Present	*Absent*	**Average**
Drug Therapy	*Present*	168	186	**177**
	Absent	188	190	**189**
	Average	**178**	**188**	**183**



- there is an **interaction** between drug therapy and biofeedback
  - effect of drug therapy depends on the level of the biofeedback factor
  - effect of biofeedback depends on the level of the drug therapy factor
  - the level of biofeedback modulates the effect of drug therapy
  - the level of drug therapy modulates the effect of biofeedback

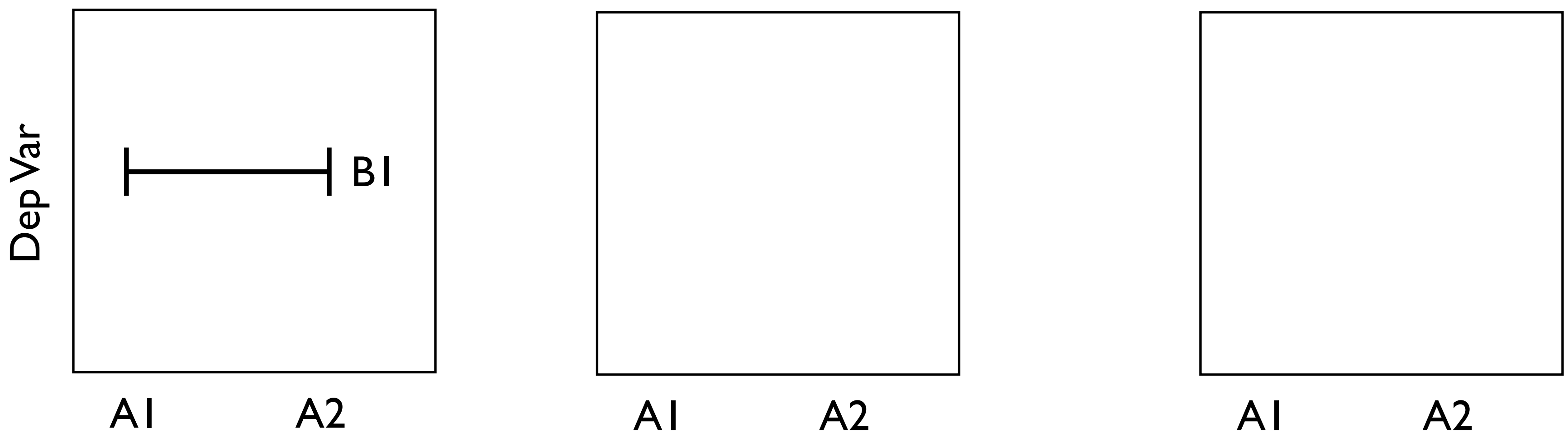
Main Effects

- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



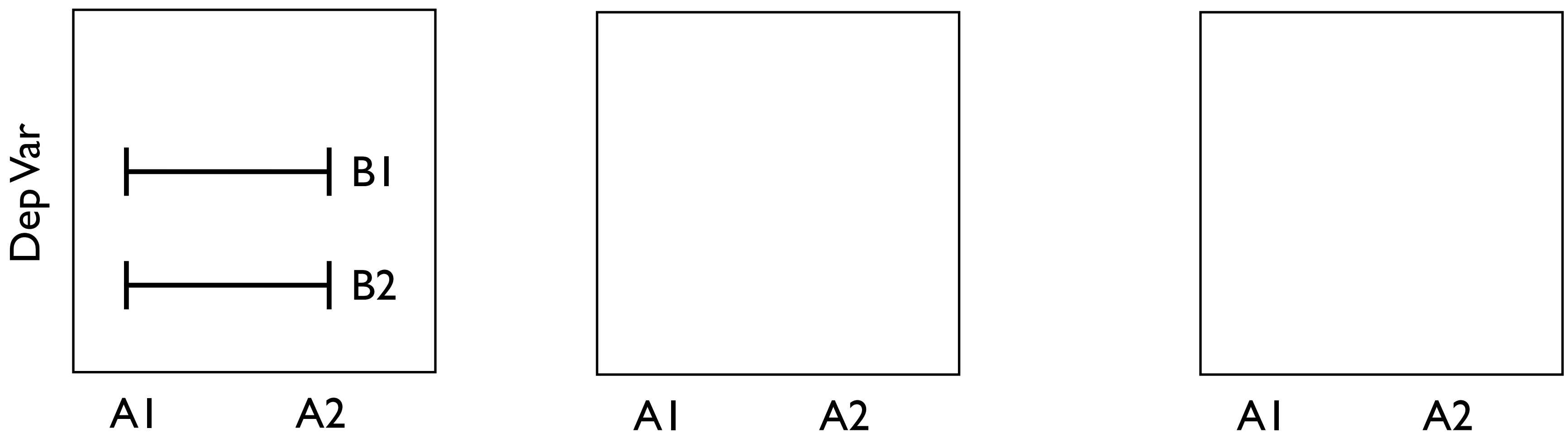
Main Effects

- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



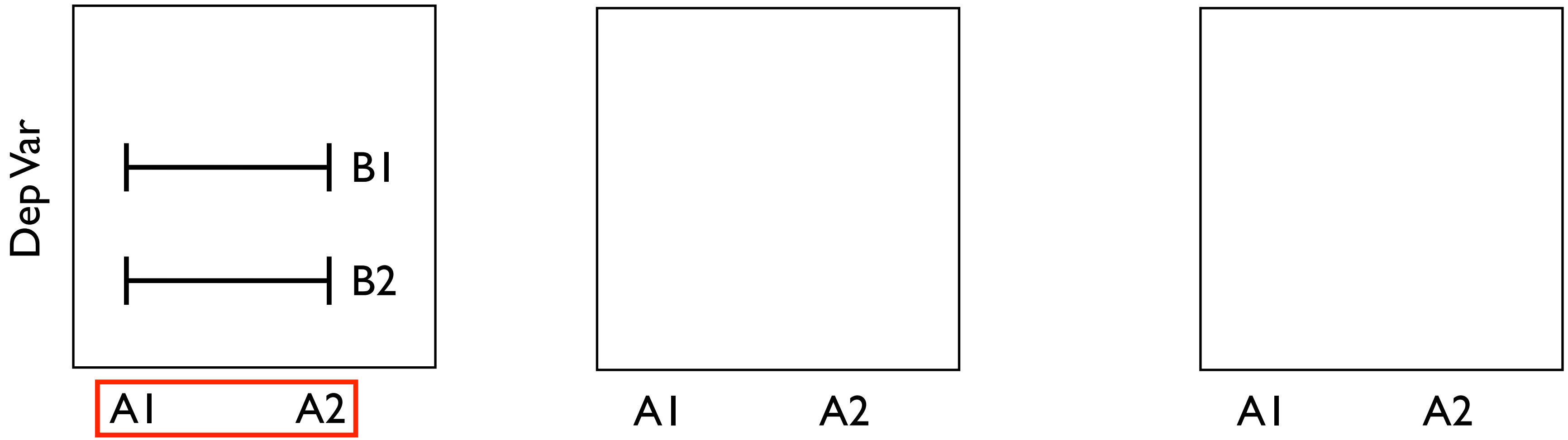
Main Effects

- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



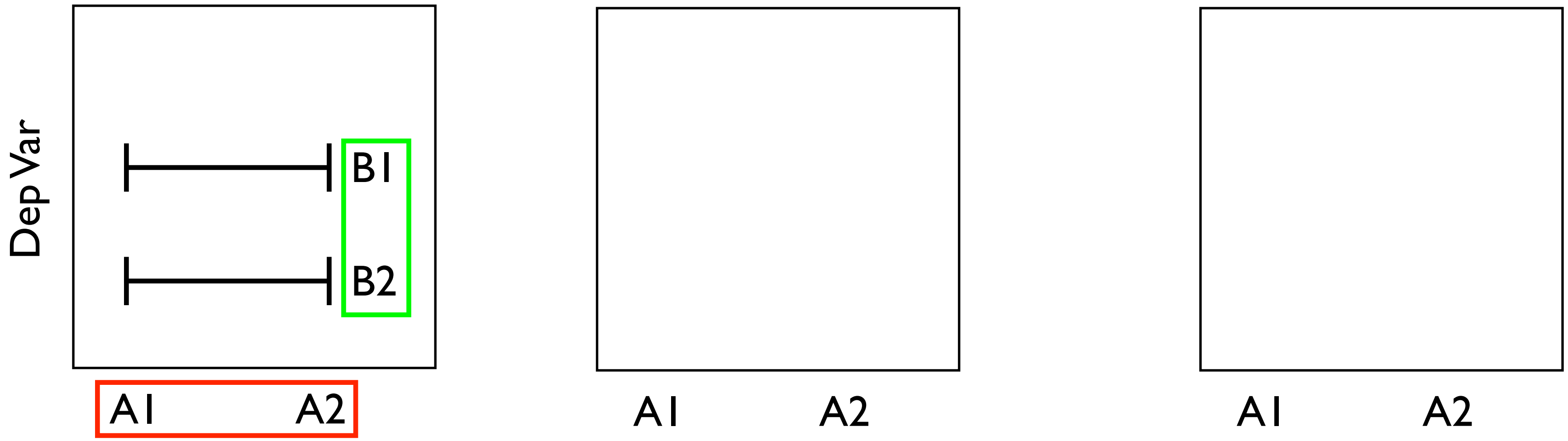
Main Effects

- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



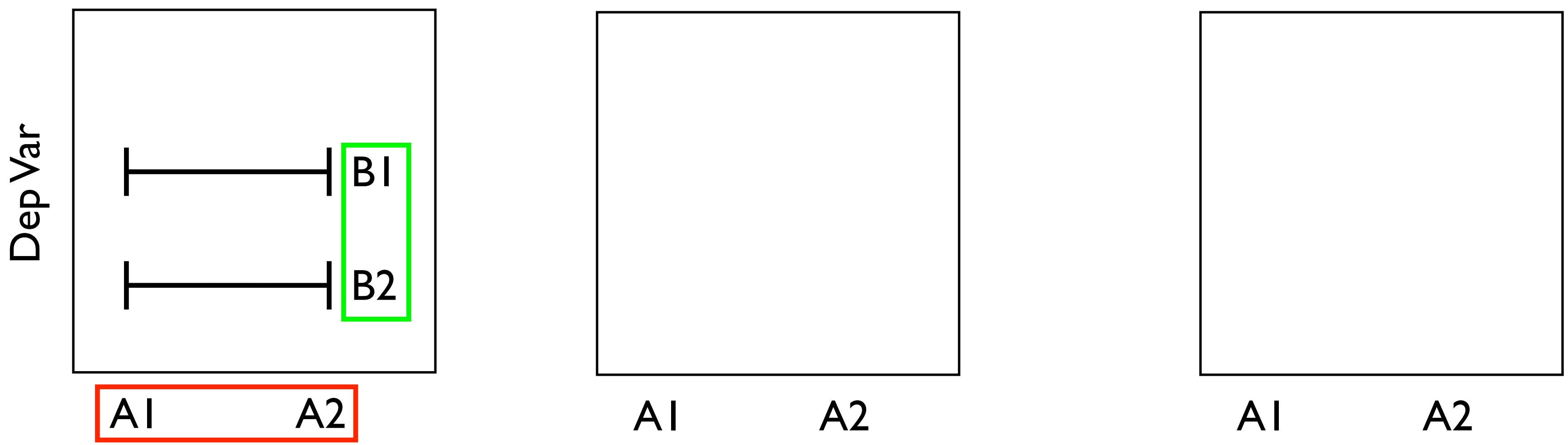
Main Effects

- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main Effects

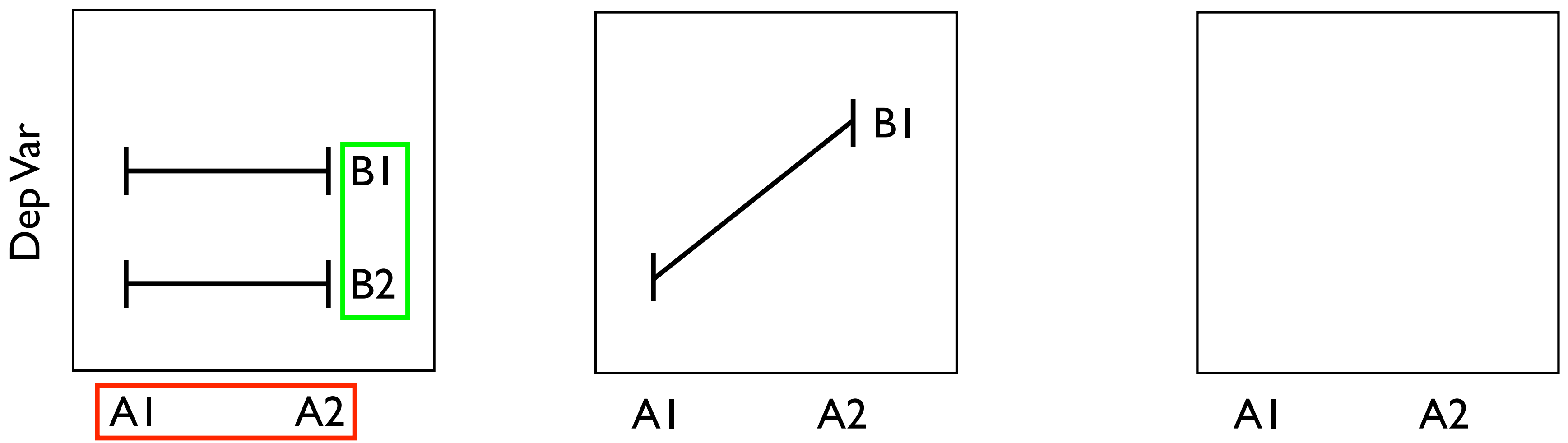
- typical main effects look like this
    - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B

Main Effects

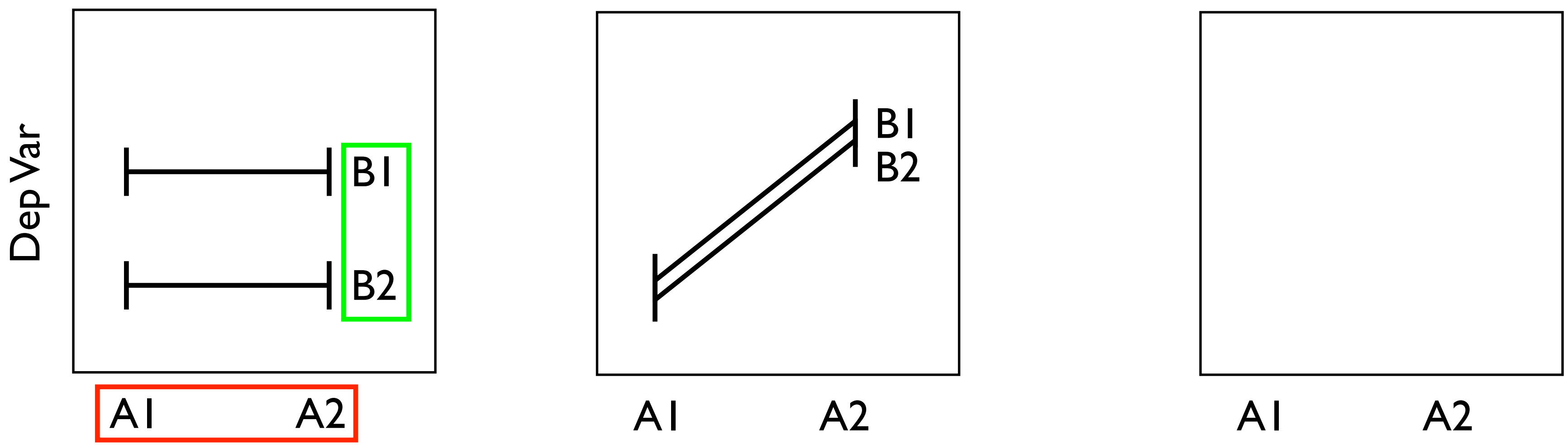
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B

Main Effects

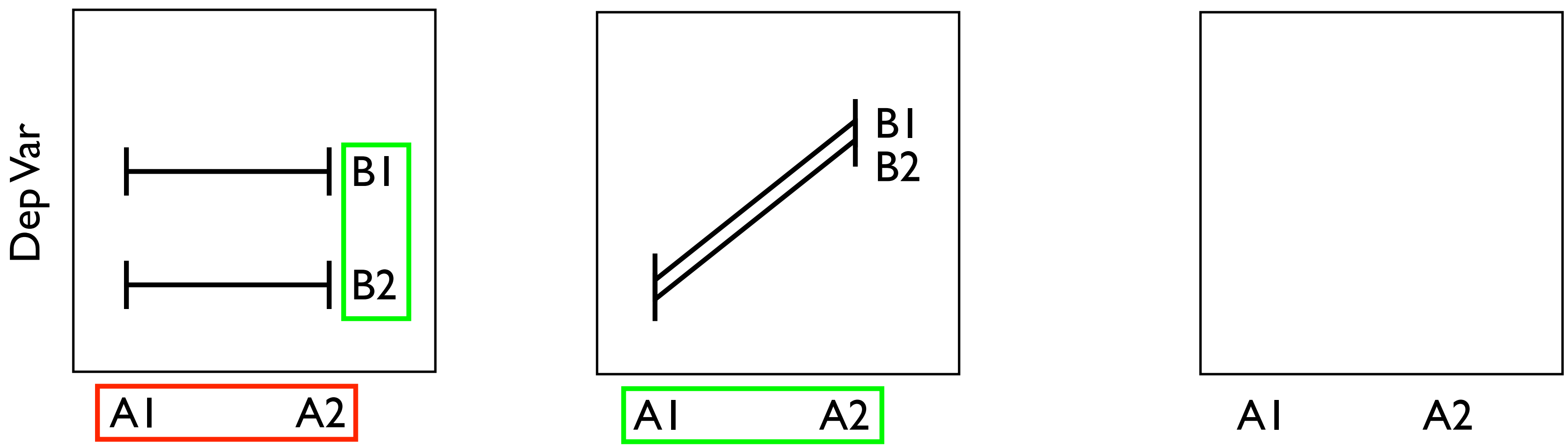
- typical main effects look like this
    - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B

Main Effects

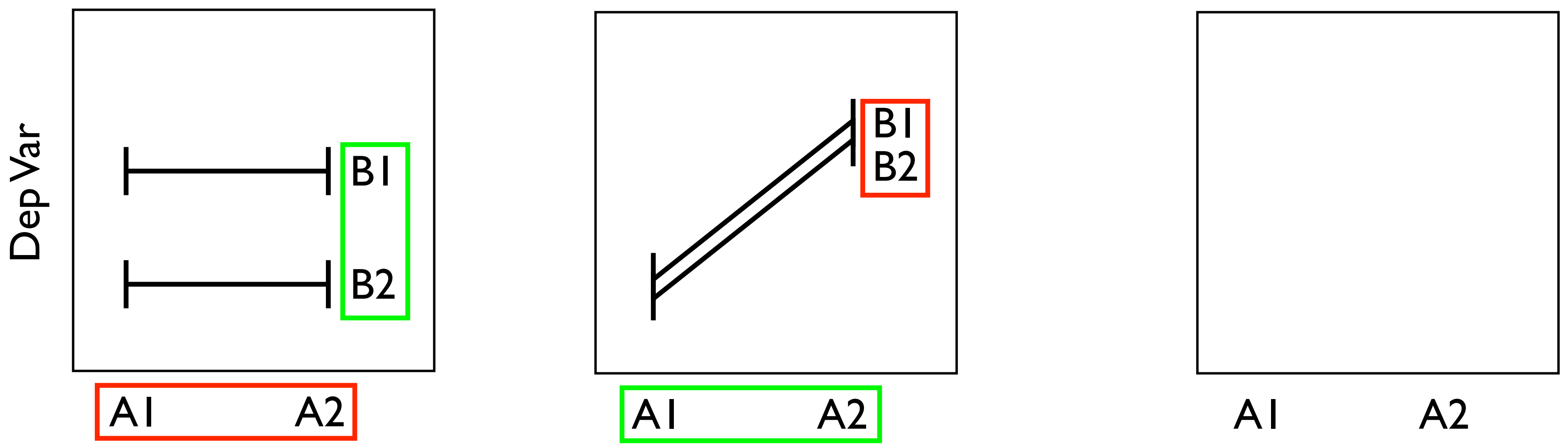
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B

Main Effects

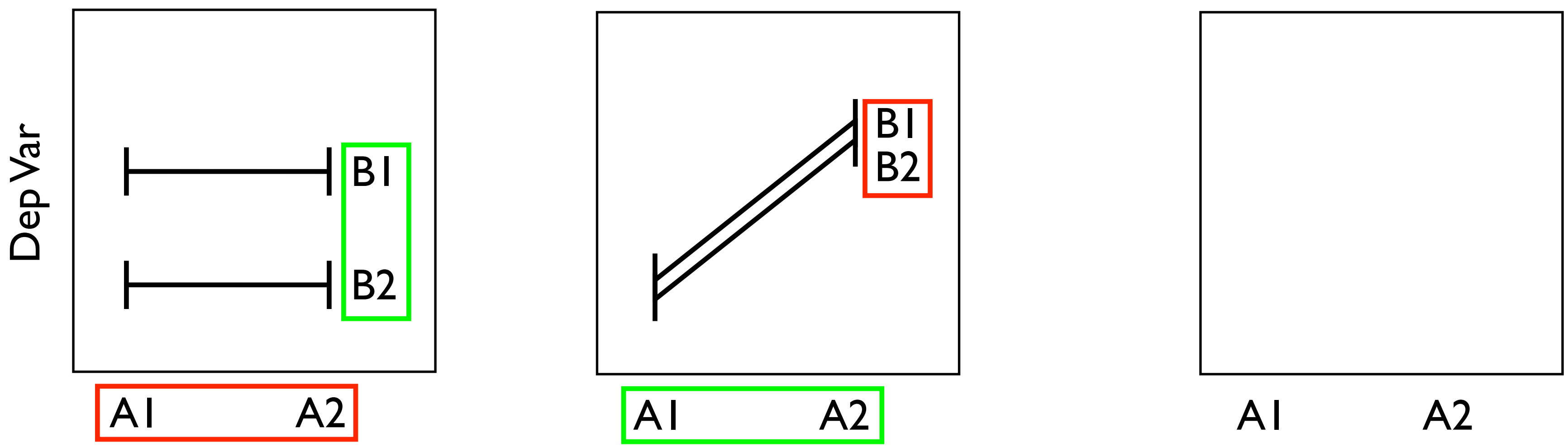
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B

Main Effects

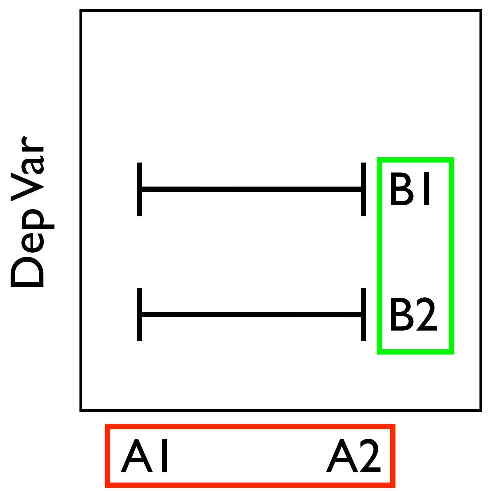
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



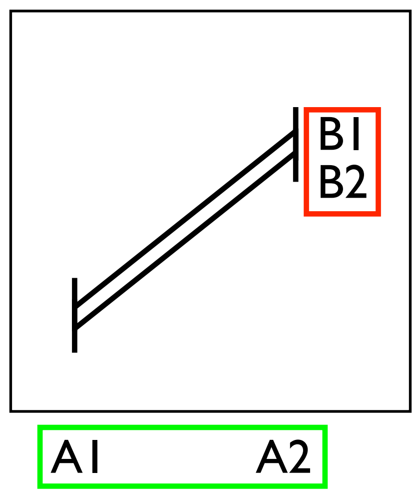
Main effect of B **Main effect of A**

Main Effects

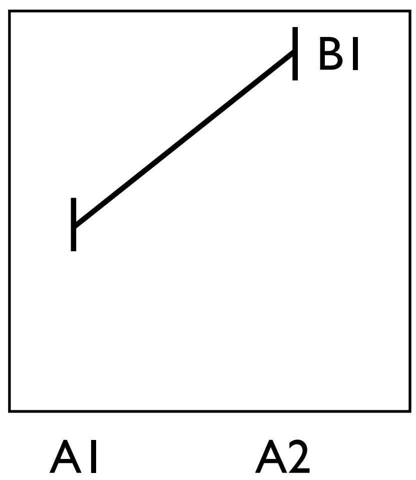
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B

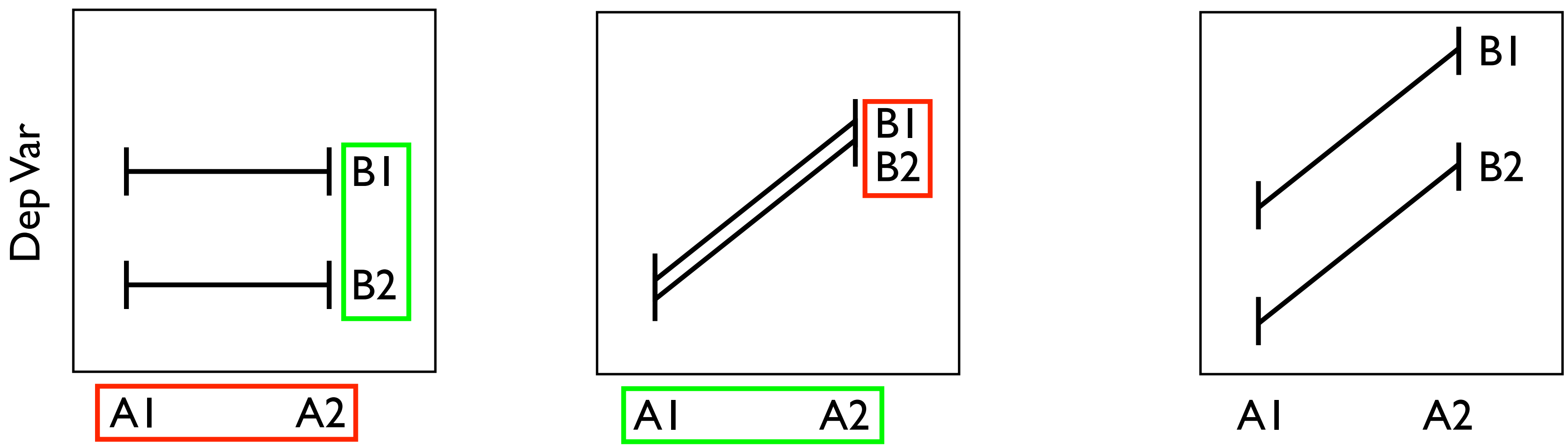


Main effect of A



Main Effects

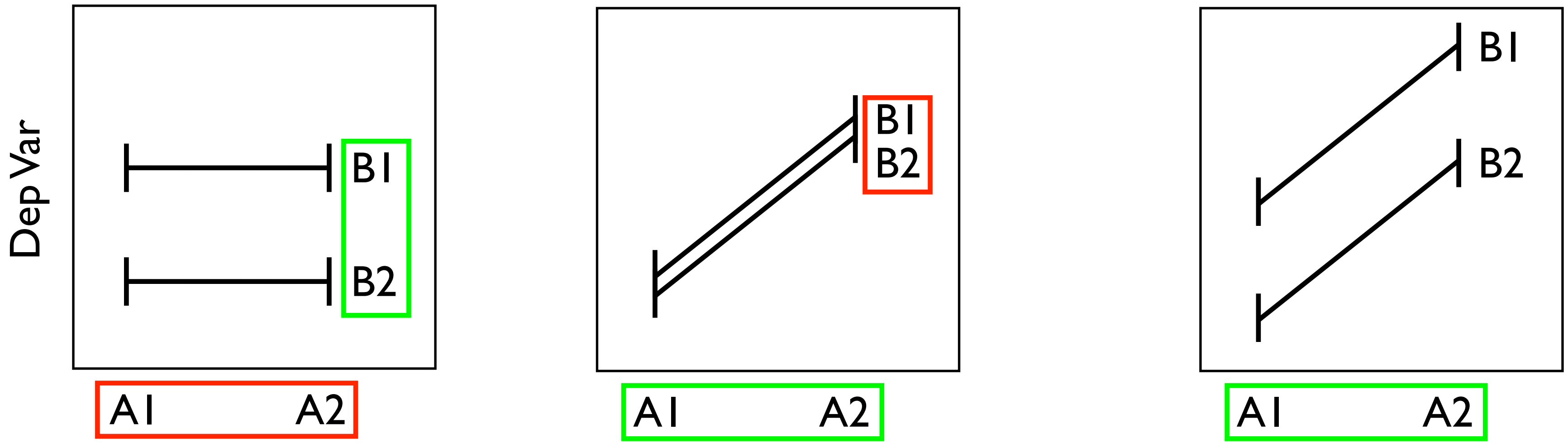
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B **Main effect of A**

Main Effects

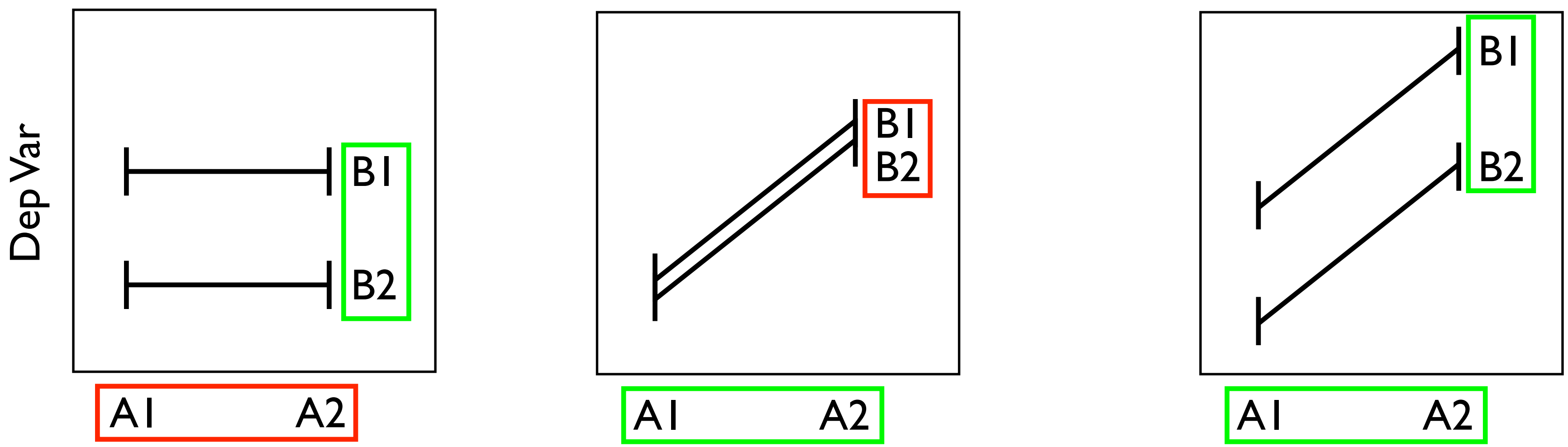
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B **Main effect of A**

Main Effects

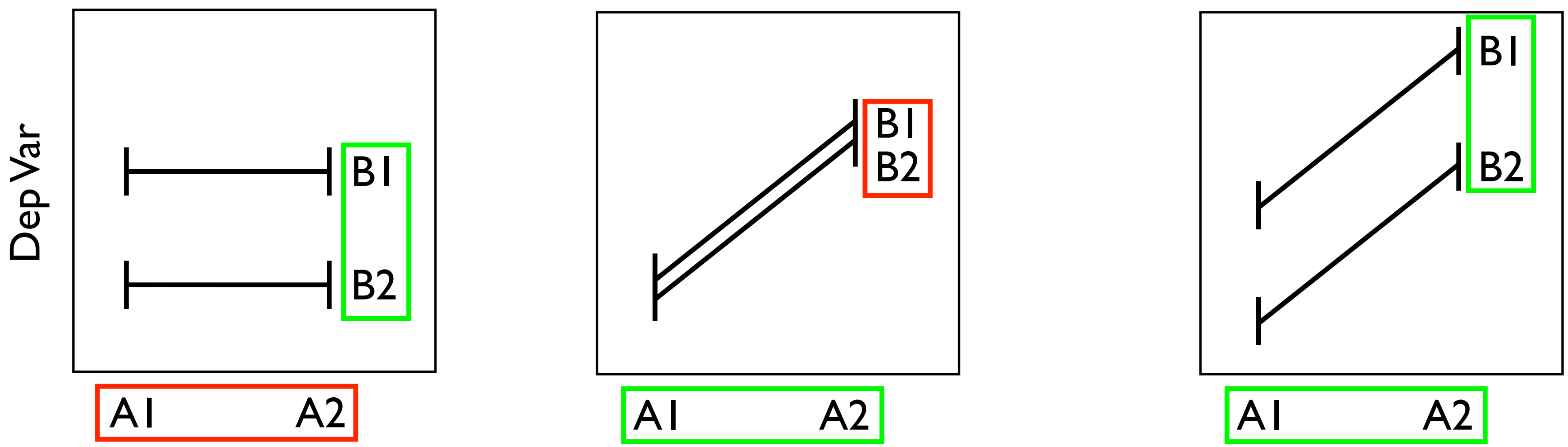
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B **Main effect of A**

Main Effects

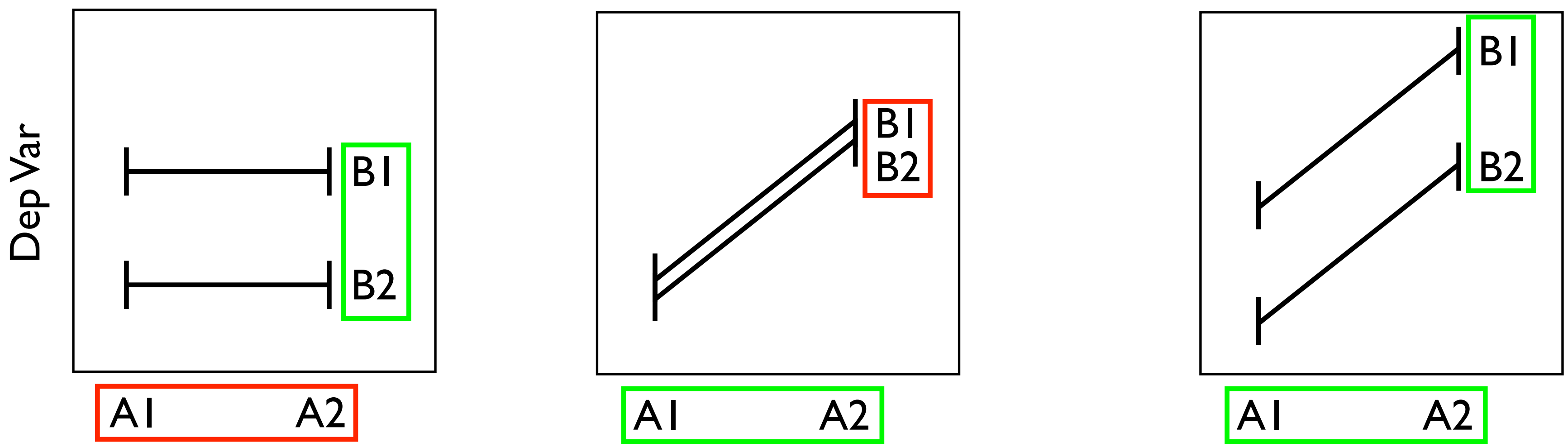
- typical main effects look like this
    - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design



Main effect of B **Main effect of A** **Main effects of A and B**

Main Effects

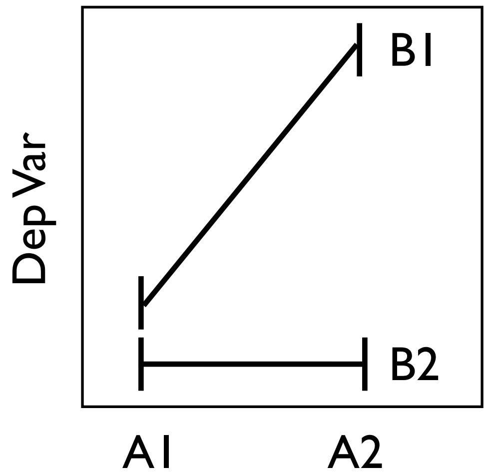
- typical main effects look like this
  - Factor A (A1, A2) and Factor B (B1, B2) fully crossed design

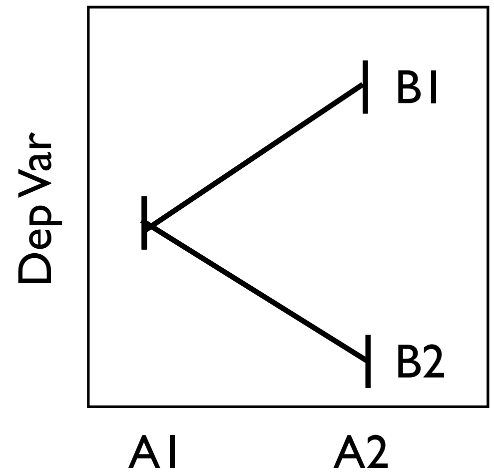


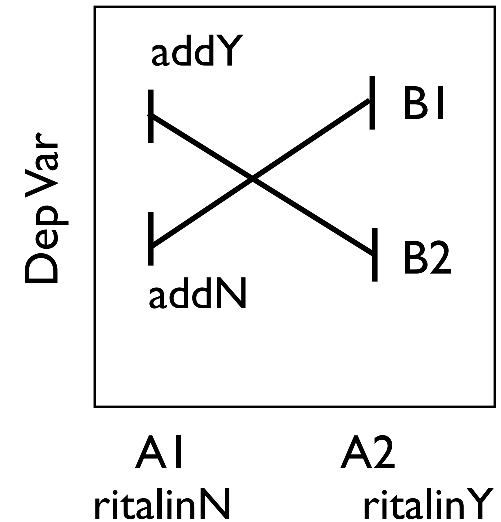
Main effect of B **Main effect of A** **Main effects of A and B**

in all 3 cases: **no A x B interaction effect**

What about these datasets?







A:

B:

AxB:

A:

B:

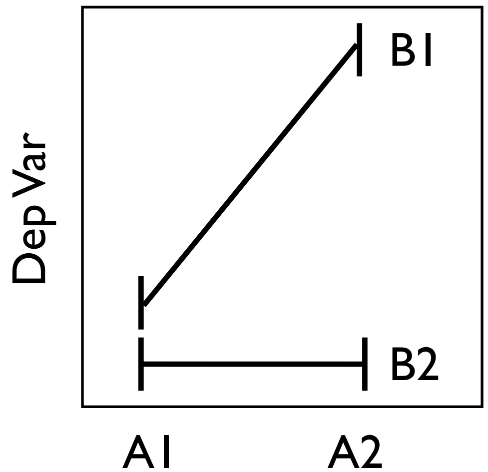
AxB:

A:

B:

AxB:

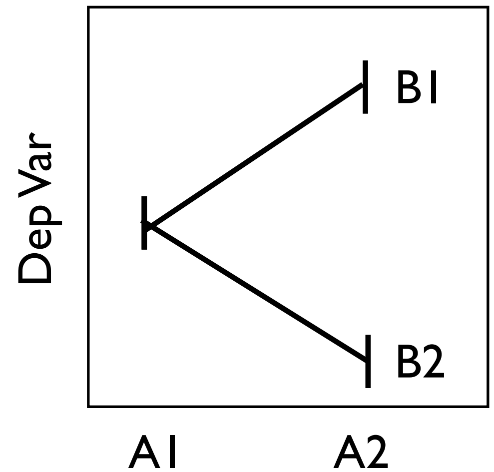
What about these datasets?



A:

B:

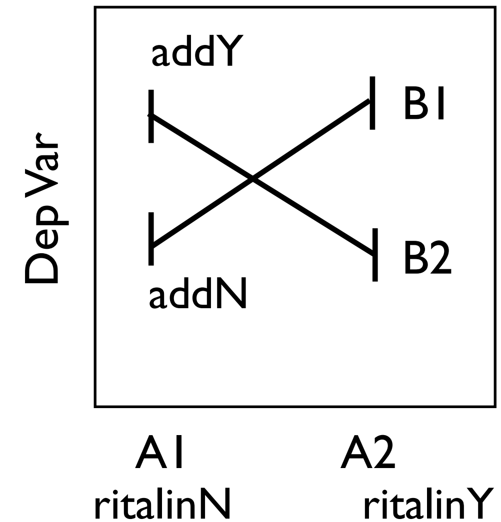
AxB:



A:

B:

AxB:

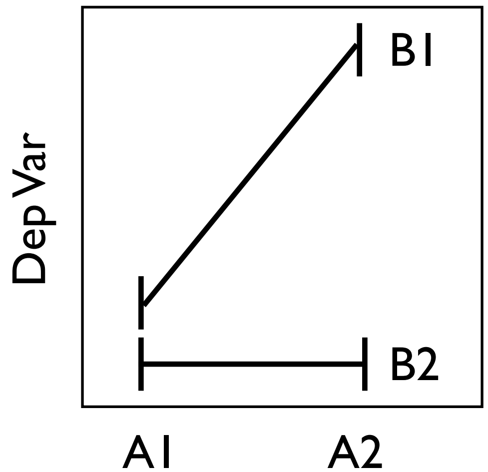


A:

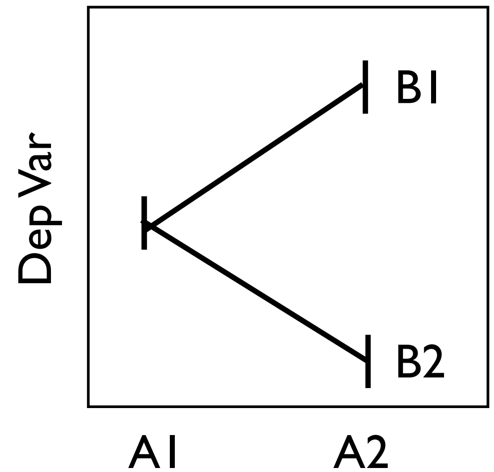
B:

AxB:

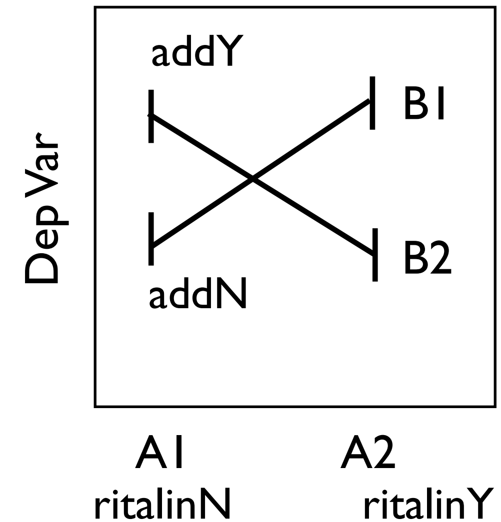
What about these datasets?



A:
B:
AxB:

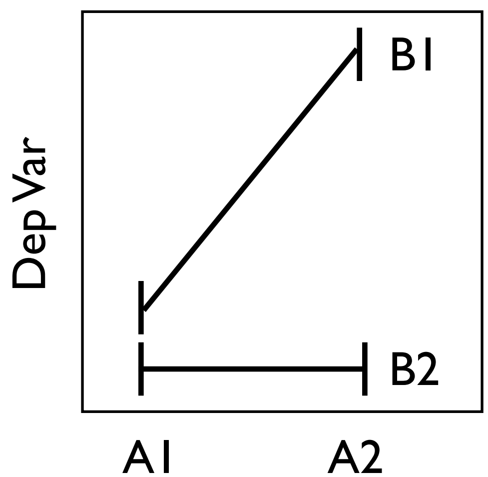


A:
B:
AxB:

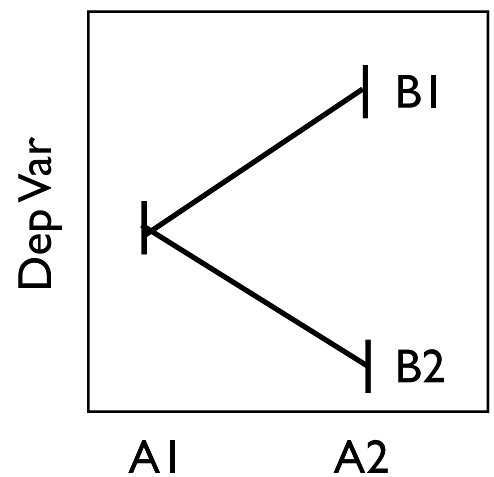


A:
B:
AxB:

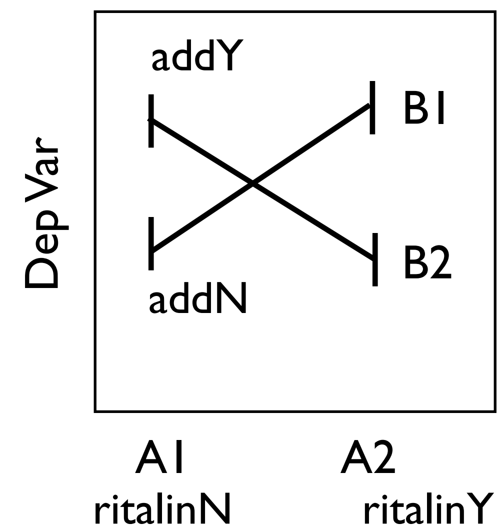
What about these datasets?



A:
B:
AxB:

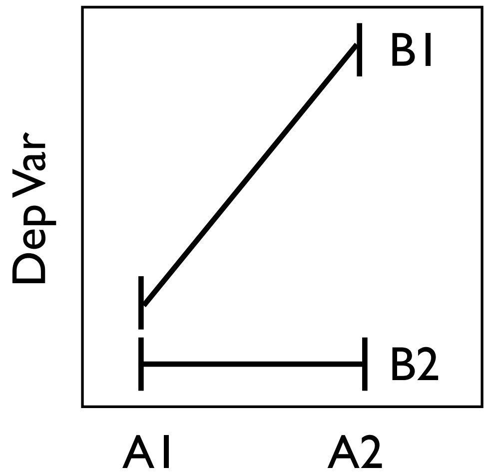


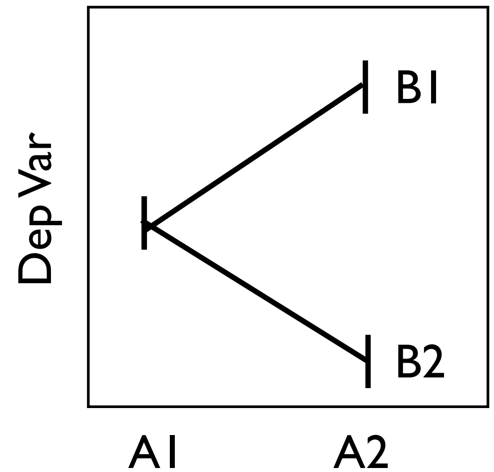
A:
B:
AxB:

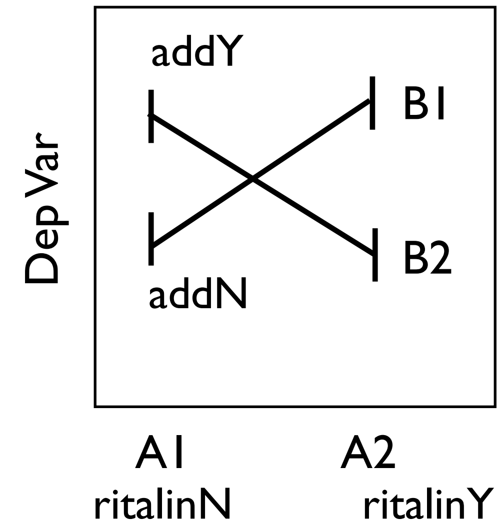


A:
B:
AxB:

What about these datasets?





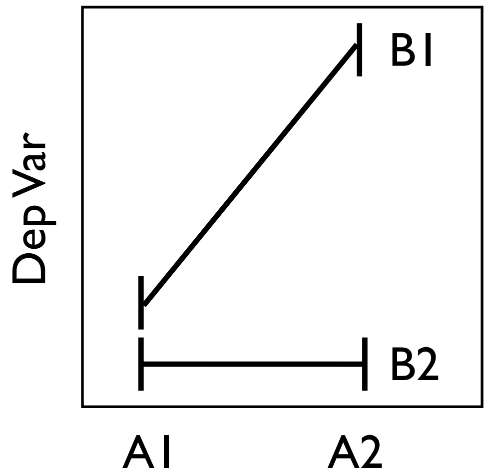


A:
B:
AxB:

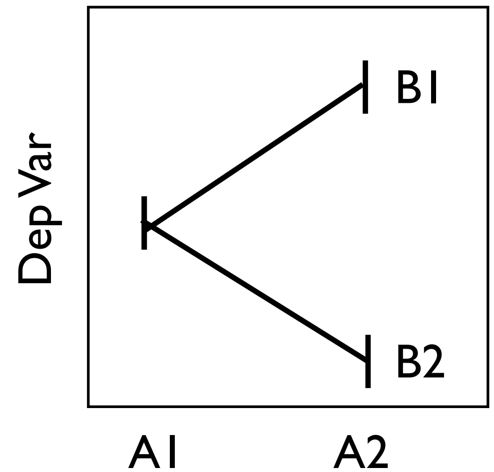
A:
B:
AxB:

A:
B:
AxB:

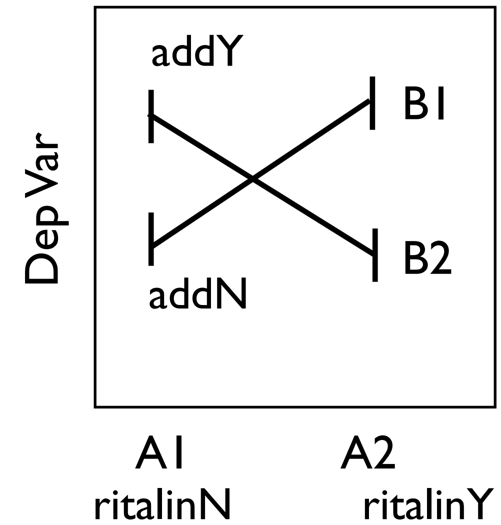
What about these datasets?



A:
B:
AxB:

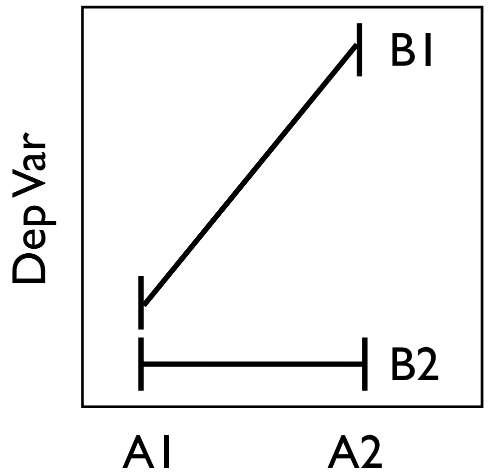


A:
B:
AxB:



A:
B:
AxB:

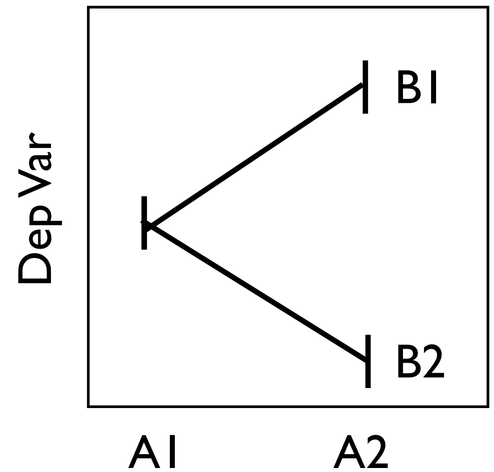
What about these datasets?



A:

B:

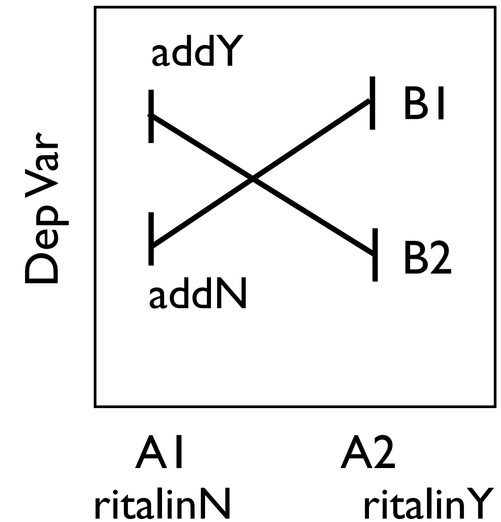
AxB:



A:

B:

AxB:

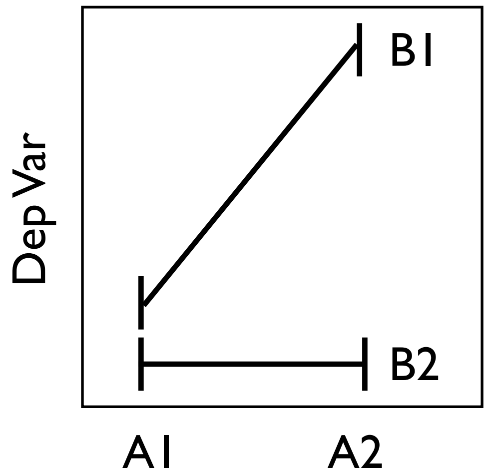


A:

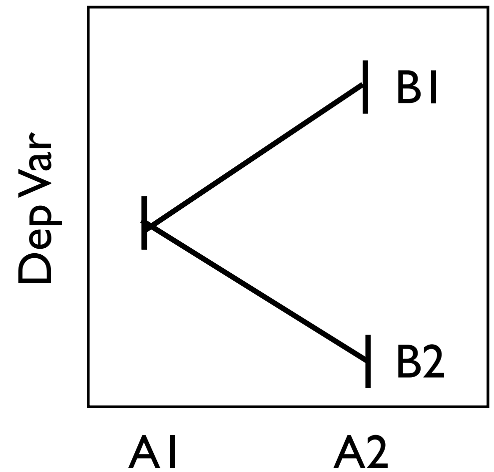
B:

AxB:

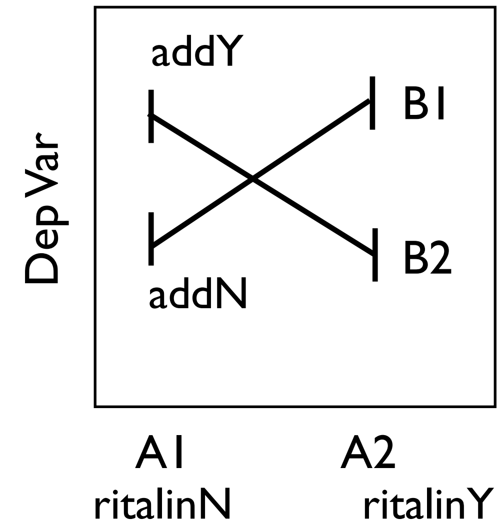
What about these datasets?



A:
B:
AxB:

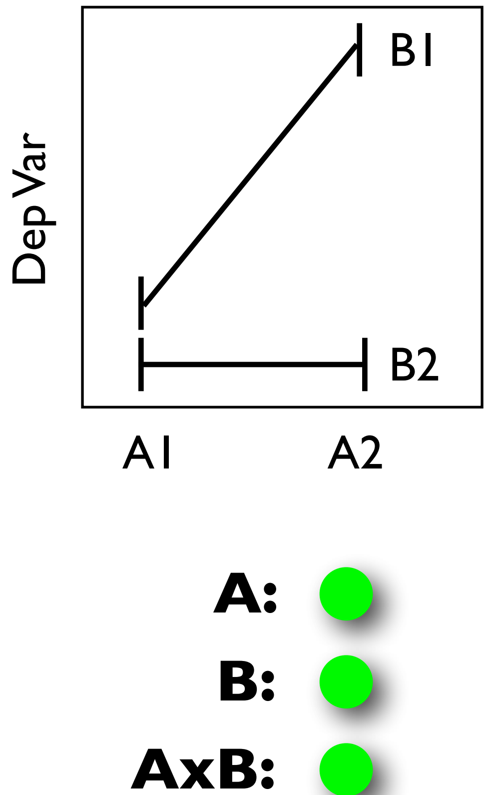


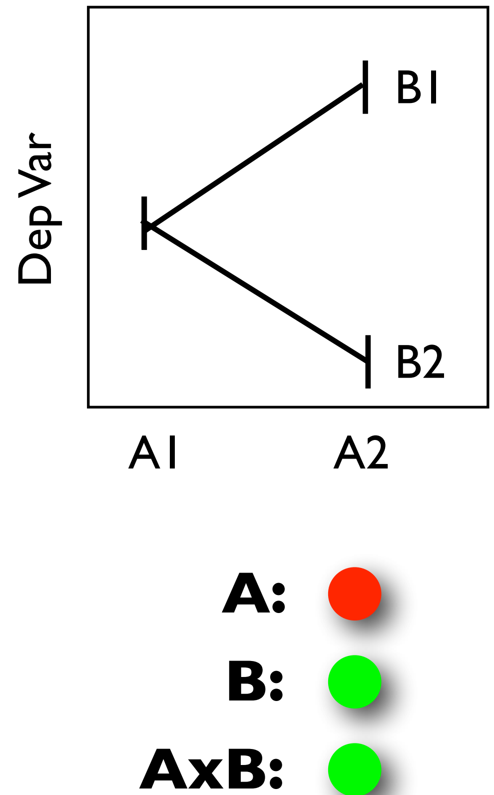
A:
B:
AxB:

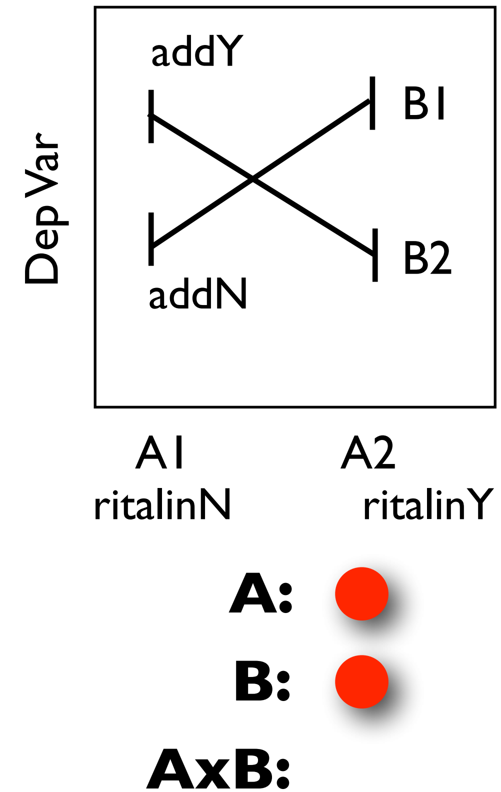


A:
B:
AxB:

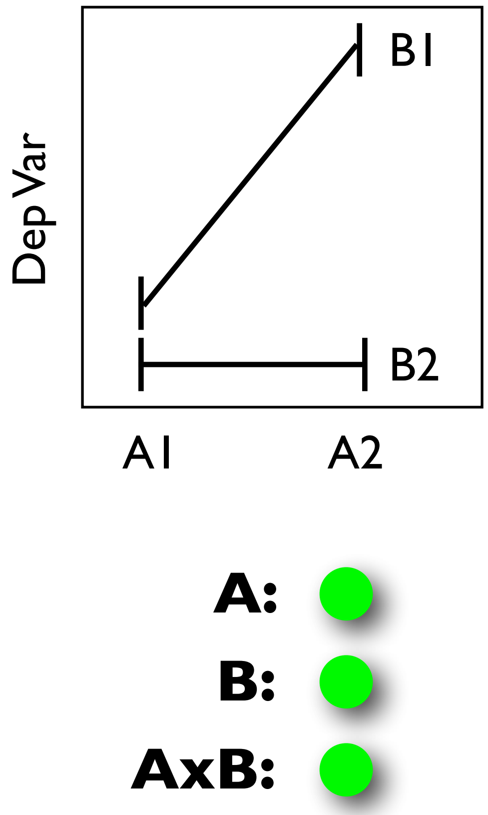
What about these datasets?

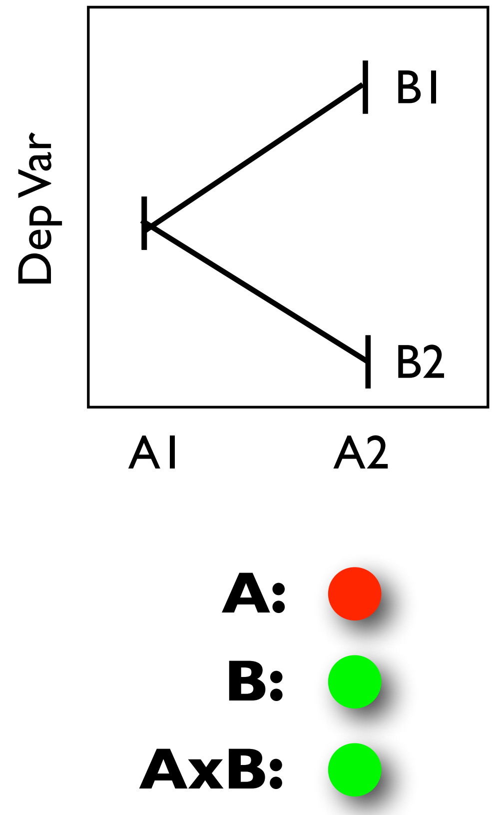


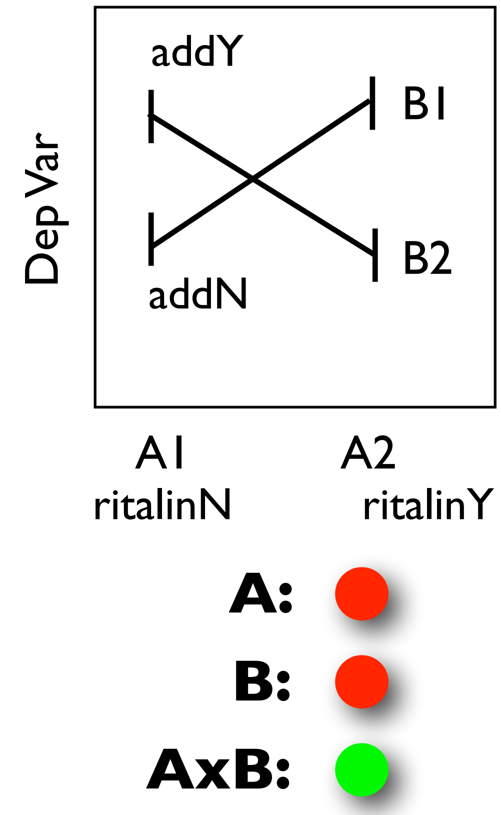




What about these datasets?







rule of thumb

- parallel lines: main effect
- non-parallel lines: interaction effect

Two Factor Design: Model Comparison Approach

- Let's assume two factors
  - Factor A with a levels
  - Factor B with b levels
- Fully crossed design
  - every level of factor A is tested with every level of factor B
  - total # groups (cells) is a x b
- we will see how to formulate in terms of model comparisons:
  - main effect of A
  - main effect of B
  - interaction effect A x B

Our approach will be as before

1. write the equation for the full and restricted models
2. derive the equations for model **error** Er and Ef
3. derive the expressions for **degrees of freedom** dfR and dfF
4. end up with an equation for the **F ratio**

The Full Model

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

- Yijk is an individual score in the jth level of factor A and the kth level of factor B (i indexes subjects within each (j,k) cell)
- μ is the overall mean of all cells
- α_j is the effect of the jth level of factor A
- β_k is the effect of the kth level of factor B
- $(\alpha\beta)_{jk}$ is the interaction effect of level j of A and level k of B

Hypothesis testing using Restricted Models

- Two-Factor (A x B) design: 3 null hypotheses to be tested:
  - main effect of A
  - main effect of B
  - interaction effect A x B
- We will formulate a separate restricted model for each hypothesis test
- **each test will involve the <u>same full model</u>**
- we will use the usual **F test**:

$$F = \frac{(E_R - E_F)/(df_R - df_F)}{E_F/df_F}$$

Main effect of A

- **full model:** $Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$

- null hypothesis is that A main effect is zero

$$H_0 : \alpha_1 = \alpha_2 = ... = \alpha_a = 0$$

- restricted model:

$$Y_{ijk} = \mu + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

From Chapter 7:

$$E_F = \sum_{allobs} (Y_{ijk} - \bar{Y}_{jk})^2$$

$$E_R - E_F = n \sum_{j=1}^{a} (\bar{Y}_j - \bar{Y})^2$$

$$df_F = ab(n-1)$$

$$df_R - df_F = a - 1$$

- so now we can do our F-test!

$$F = \frac{(E_R - E_F)/(df_R - df_F)}{E_F/df_F}$$

denominator is always the same as MS_W from ANOVA table

Main Effect of B

- full model again is:

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

- restricted model is:

$$Y_{ijk} = \mu + \alpha_j + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

- See Chapter 7 for equations for EF and ER-EF

Interaction effect AB

- full model again is:

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

- restricted model:

$$Y_{ijk} = \mu + \alpha_j + \beta_k + \epsilon_{ijk}$$

Controlling Alpha level

- huh? we are doing three tests here and we are doing nothing about controlling the Type-I error rate. Why not?
- each test is conceptualized as a separate "family" of tests
- each test is addressing an **independent** question
- the approach is to control the **family-wise alpha level** at 0.05
- each **major effect** (A, B, AB) is considered a **family**
- within each family of tests we control alpha at 0.05 level

Controlling Alpha level

- so we are allowing experiment-wise alpha level to exceed 0.05
- we are controlling the family-wise alpha level at 0.05
- does this seem rather arbitrary to you?
- it's not entirely arbitrary .... BUT
- it's not entirely non-arbitrary either
- statistics is a framework for formulating rational approaches to inferences based on data
- you are responsible for your own convincing arguments

Follow-up Tests

- ok - so we've done F-tests for the main effect A, main effect B, and interaction effect AB; now what?
- investigate the nature of each significant effect
- there is a good rule of thumb for how to proceed:

Follow-up Tests

- **first look at the interaction effect**
- **IF** interaction effect is significant,
  - perform analyses of "simple effects"
  - (i.e. investigate the nature of the interaction)
  - and DON'T bother looking into the main effects (they are not informative anyway)
- **ELSE** if interaction effect is not significant,
  - perform contrasts within each significant main effect to understand the nature of the differences
- so if interaction is significant don't bother looking at the main effects

Follow-up Tests

- **Further Investigations of Main Effects**
- upon finding a significant main effect, the precise effect is not known
- we do not know in what way the different levels of the factor differ
- contrasts are formed and tested in the same way as in a one-way design
- e.g. to test a contrast ψ in the main effect of A (averaged over levels of B):

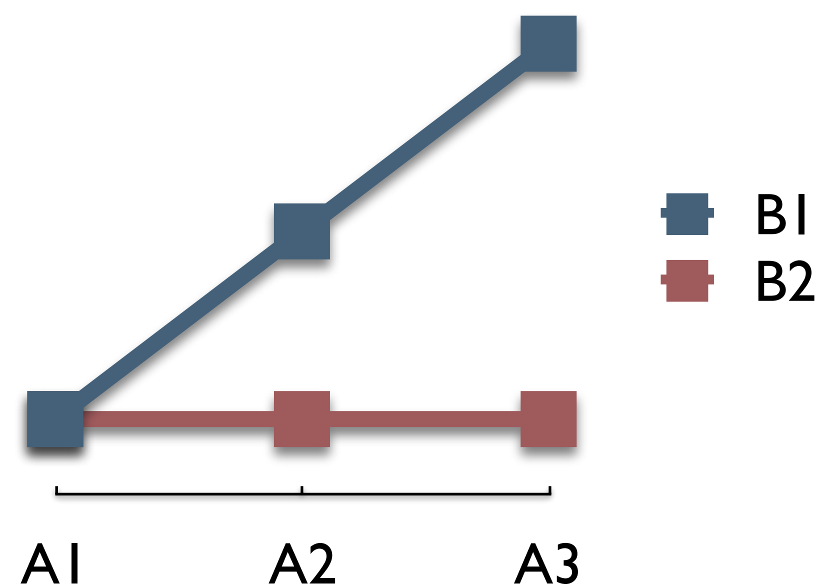
$$SS_\psi = nb(\psi)^2 / \sum_{j=1}^{a} c_j^2 \qquad F = SS_\psi / MS_W$$

Follow-up Tests

- critical value of F (Fcrit) will depend on the same kinds of decisions we discussed in Chapter 5 on multiple-comparison procedures
- lots of possibilities including:
  - no correction
  - Bonferroni / Bonferroni-Holm
  - Tukey
  - Scheffé
- I can tell you about different approaches but ultimately it's up to you to decide how to control family-wise alpha level

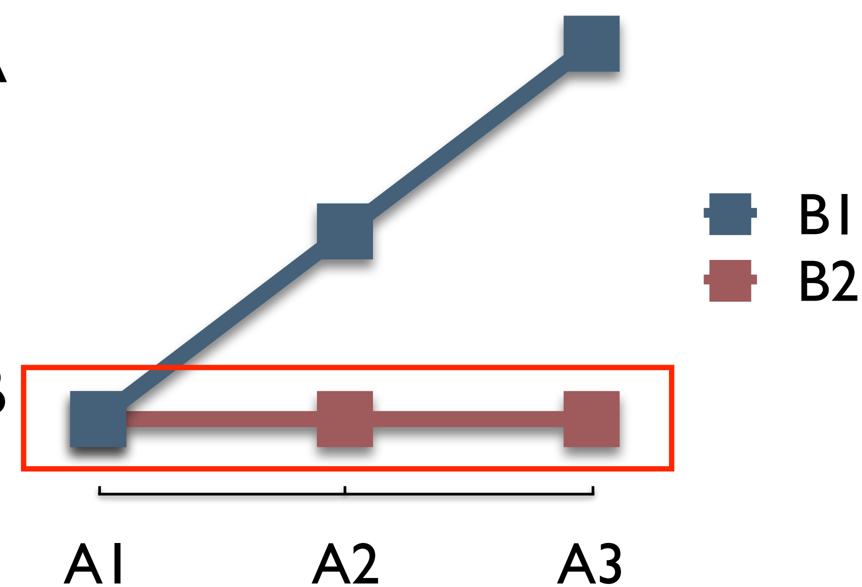
Investigating Interactions: Simple Effects

- like testing contrasts of a main effect, **except** we perform contrasts <u>separately in each level of the other factor</u>
- *like* a mini one-way anova (but **NOT** a one-way anova)
- e.g. two-factors A (3 levels) and B (2 levels)
- let's say we have a significant AB interaction
- test contrasts across levels of A
  - but within each level of B separately
- OR alternatively,
  test contrasts across levels of B
  - but within each level of A separately



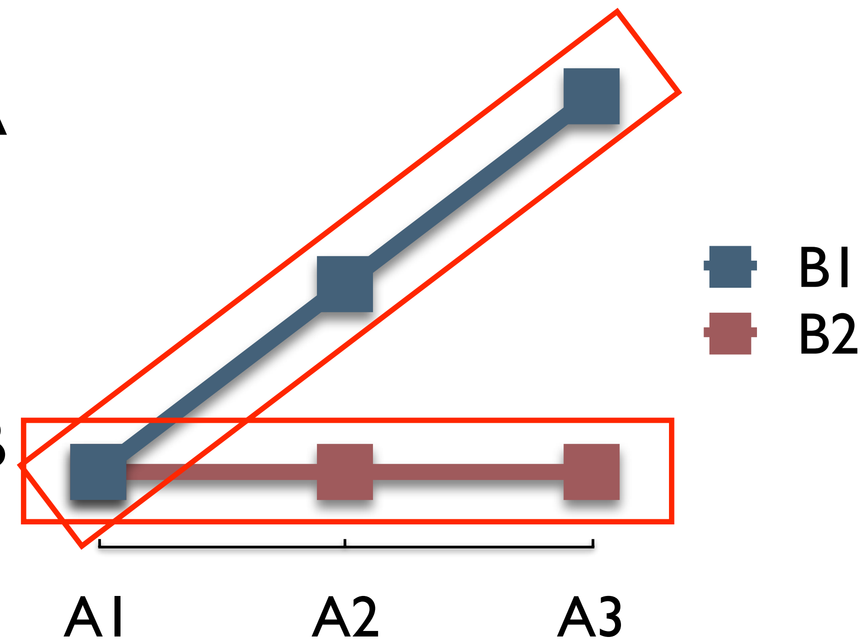
Investigating Interactions: Simple Effects

- like testing contrasts of a main effect, **except** we perform contrasts separately in each level of the other factor
- *like* a mini one-way anova (but **NOT** a one-way anova)
- e.g. two-factors A (3 levels) and B (2 levels)
- let's say we have a significant AB interaction
- test contrasts across levels of A
  - but within each level of B separately
- OR alternatively,
  test contrasts across levels of B
  - but within each level of A separately



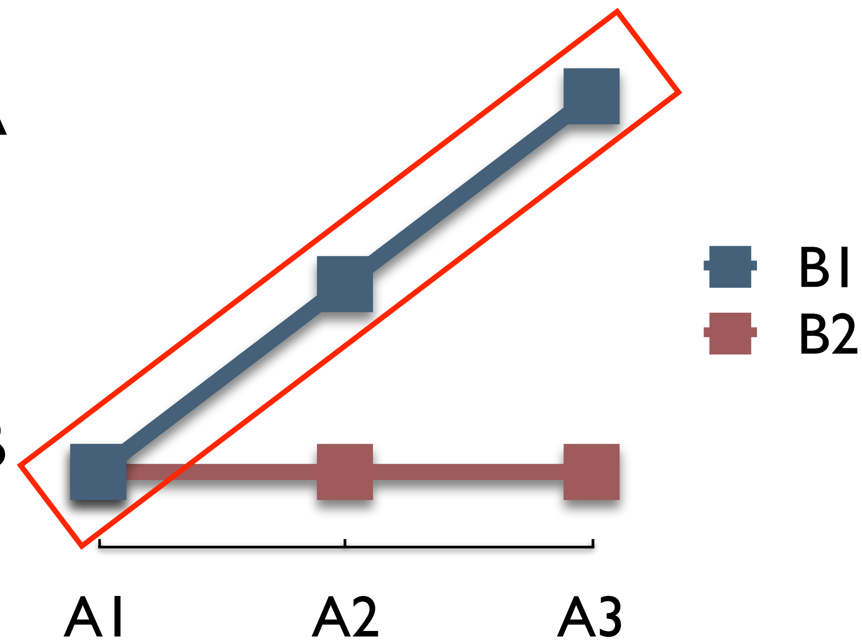
Investigating Interactions: Simple Effects

- like testing contrasts of a main effect, **except** we perform contrasts <u>separately in each level of the other factor</u>
- *like* a mini one-way anova (but <u>**NOT** a one-way anova</u>)
- e.g. two-factors A (3 levels) and B (2 levels)
- let's say we have a significant AB interaction
- test contrasts across levels of A
  - but within each level of B separately
- OR alternatively, test contrasts across levels of B
  - but within each level of A separately



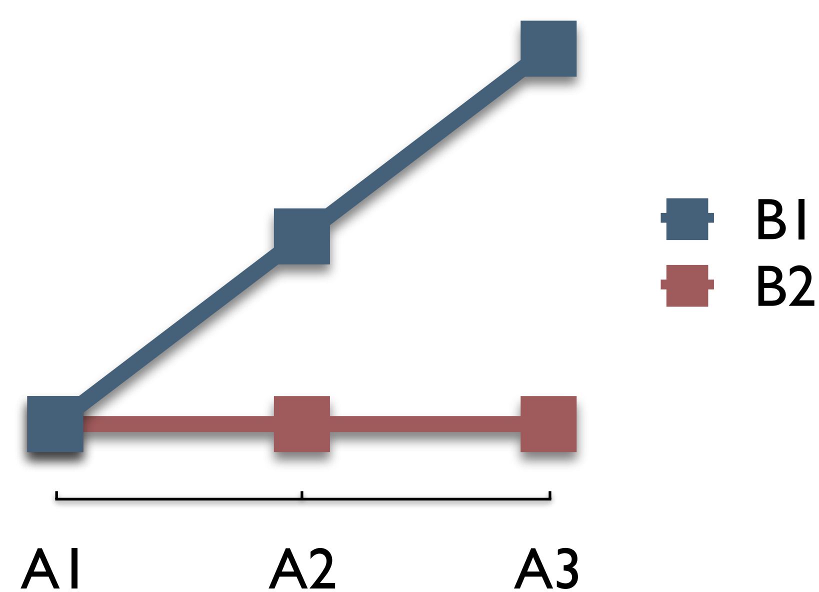
Investigating Interactions: Simple Effects

- like testing contrasts of a main effect, **except** we perform contrasts <u>separately in each level of the other factor</u>
- *like* a mini one-way anova (but <u>**NOT** a one-way anova</u>)
- e.g. two-factors A (3 levels) and B (2 levels)
- let's say we have a significant AB interaction
- test contrasts across levels of A
  - but within each level of B separately
- OR alternatively,
  test contrasts across levels of B
  - but within each level of A separately



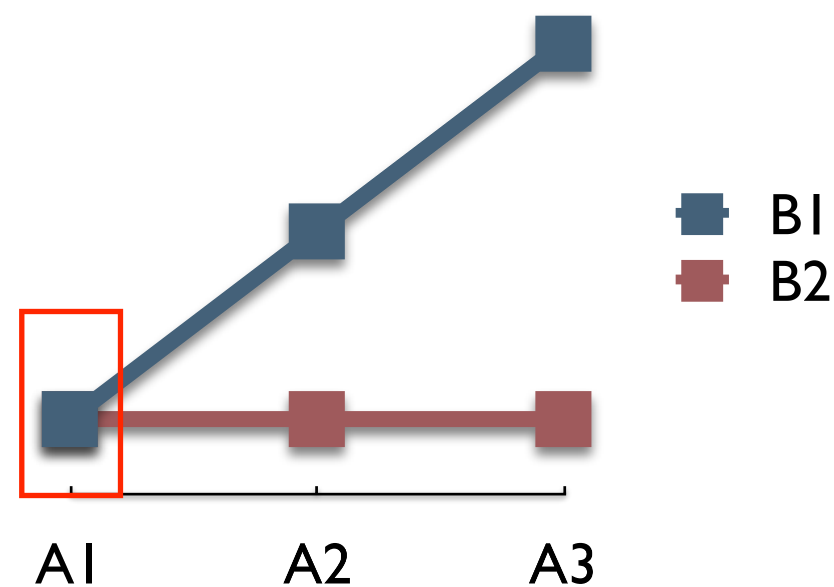
Investigating Interactions: Simple Effects

- like testing contrasts of a main effect, **except** we perform contrasts <u>separately in each level of the other factor</u>
- *like* a mini one-way anova (but **NOT** a one-way anova)
- e.g. two-factors A (3 levels) and B (2 levels)
- let's say we have a significant AB interaction
- test contrasts across levels of A
  - but within each level of B separately
- OR alternatively,
  test contrasts across levels of B
  - but within each level of A separately



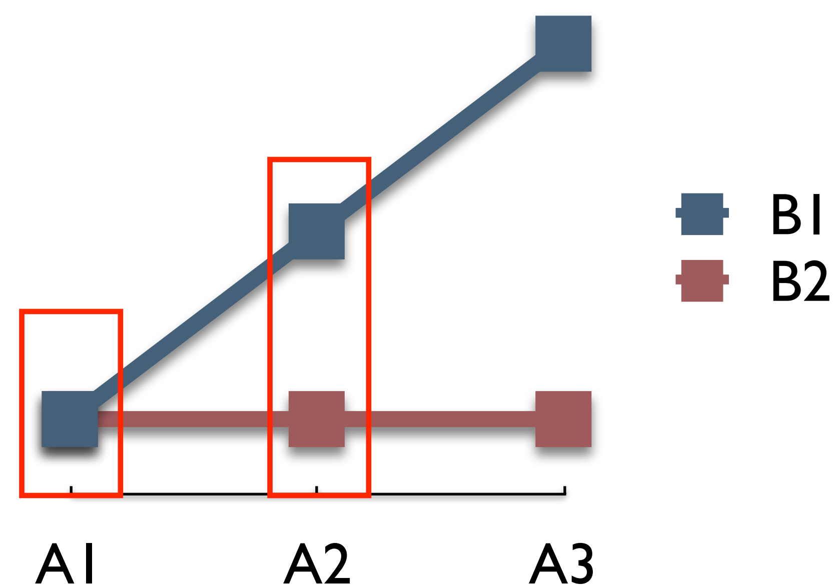
Investigating Interactions: Simple Effects

- like testing contrasts of a main effect, **except** we perform contrasts <u>separately in each level of the other factor</u>
- *like* a mini one-way anova (but <u>**NOT** a one-way anova</u>)
- e.g. two-factors A (3 levels) and B (2 levels)
- let's say we have a significant AB interaction
- test contrasts across levels of A
  - but within each level of B separately
- OR alternatively, test contrasts across levels of B
  - but within each level of A separately



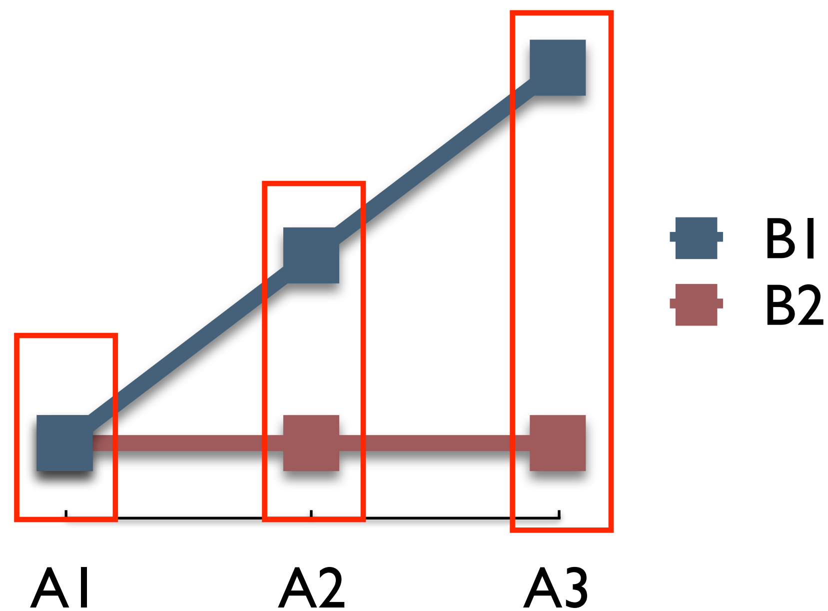
Investigating Interactions: Simple Effects

- like testing contrasts of a main effect, **except** we perform contrasts <u>separately in each level of the other factor</u>
- *like* a mini one-way anova (but <u>**NOT** a one-way anova</u>)
- e.g. two-factors A (3 levels) and B (2 levels)
- let's say we have a significant AB interaction
- test contrasts across levels of A
  - but within each level of B separately
- OR alternatively,
  test contrasts across levels of B
  - but within each level of A separately



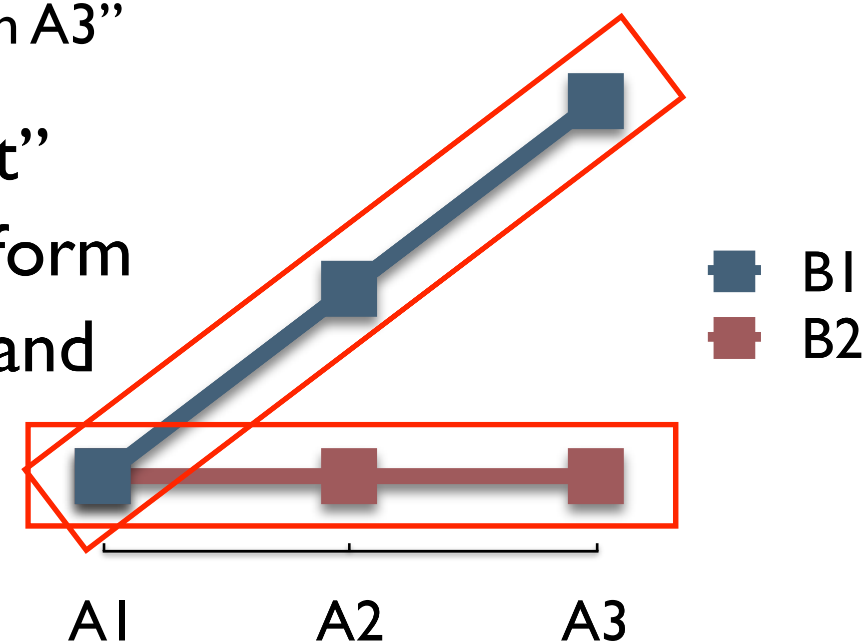
Investigating Interactions: Simple Effects

- like testing contrasts of a main effect, **except** we perform contrasts <u>separately in each level of the other factor</u>
- *like* a mini one-way anova (but <u>**NOT** a one-way anova</u>)
- e.g. two-factors A (3 levels) and B (2 levels)
- let's say we have a significant AB interaction
- test contrasts across levels of A
  - but within each level of B separately
- OR alternatively,
  test contrasts across levels of B
  - but within each level of A separately



Investigating Interactions: Simple Effects

- test contrasts across levels of A
  - but within each level of B separately
  - called A "within B1" and A "within B2" simple effects
- OR alternatively,
  test contrasts across levels of B
  - but within each level of A separately
  - B "within A1", B "within A2", B "within A3"
- Upon a significant "simple effect"
  we would then proceed to perform
  additional contrasts to understand
  the nature of the differences



Investigating Interactions

$$F = \frac{SS_{contrast}/df_{contrast}}{MS_W}$$

- we can perform an F test on **any contrast** we want as long as we can compute SS_contrast and df_contrast
- MS_W always comes directly from ANOVA table

$$SS_\psi = n(\psi)^2 / \sum_{j=1}^{a} c_j^2$$

this equation is your friend

- see Chapter 7 for some numerical examples

Type-I Error Rate

- when you test a bunch of contrasts in order to follow up a significant interaction effect, what should you do to control Type-I error rate?
- one school of thought: nothing! you are only performing the tests if the interaction is significant at 0.05 - so probability that any of the followup tests will be a Type-I error is also 0.05
- M & D don't like this - they say this logic can be flawed if the interaction null hypothesis is "partially" true
- what to do depends on what you constitute as a "family"

Type-I Error Rate

- M & D: suggest we consider all tests regarding differences among levels of a given factor as a separate "family" of tests
- Goal should be to maintain alpha = 0.05 <u>within each family</u>
- they suggest a Bonferroni-like approach
- take # of tests done in each family and divide the alpha level (0.05) by that number
- I suggest: if # comparisons is small (2 or 3) this is ok. If # comparisons is much greater than 2 or 3, use Tukey instead

Statistical Power

- Chapter 7 gives some computational formulas for computing statistical power of
    - main effect of A
    - main effect of B
    - interaction effect AB
- We won't go into it here
- Read it on your own time

Non-orthogonal designs

- orthogonal design = a design with equal number of subjects within each cell
- non-orthogonal design = a design with different numbers of subjects within each cell
- There is controversy about best approach for analysing non-orthogonal designs
- one approach is to compute a new version of n called a “harmonic mean”, sort of like an average # of subjects
- read about it in the Chapter
- my advice: avoid non-orthogonal designs

Advantages of Factorial Designs

Advantages of Factorial Designs

- suppose we are interested in effects of various treatments for hypertension: biofeedback vs drugs X, Y, Z (vs nothing)

Advantages of Factorial Designs

- suppose we are interested in effects of various treatments for hypertension: biofeedback vs drugs X, Y, Z (vs nothing)
- is it better to conduct a 2 x 3 factorial study OR two separate single-factor studies?

Advantages of Factorial Designs

- suppose we are interested in effects of various treatments for hypertension: biofeedback vs drugs X, Y, Z (vs nothing)
- is it better to conduct a 2 x 3 factorial study OR two separate single-factor studies?
- factorial design enables us to test for an interaction

Advantages of Factorial Designs

- suppose we are interested in effects of various treatments for hypertension: biofeedback vs drugs X, Y, Z (vs nothing)
- is it better to conduct a 2 x 3 factorial study OR two separate single-factor studies?
- factorial design enables us to test for an interaction
- factorial design allows for greater generalizability

Advantages of Factorial Designs

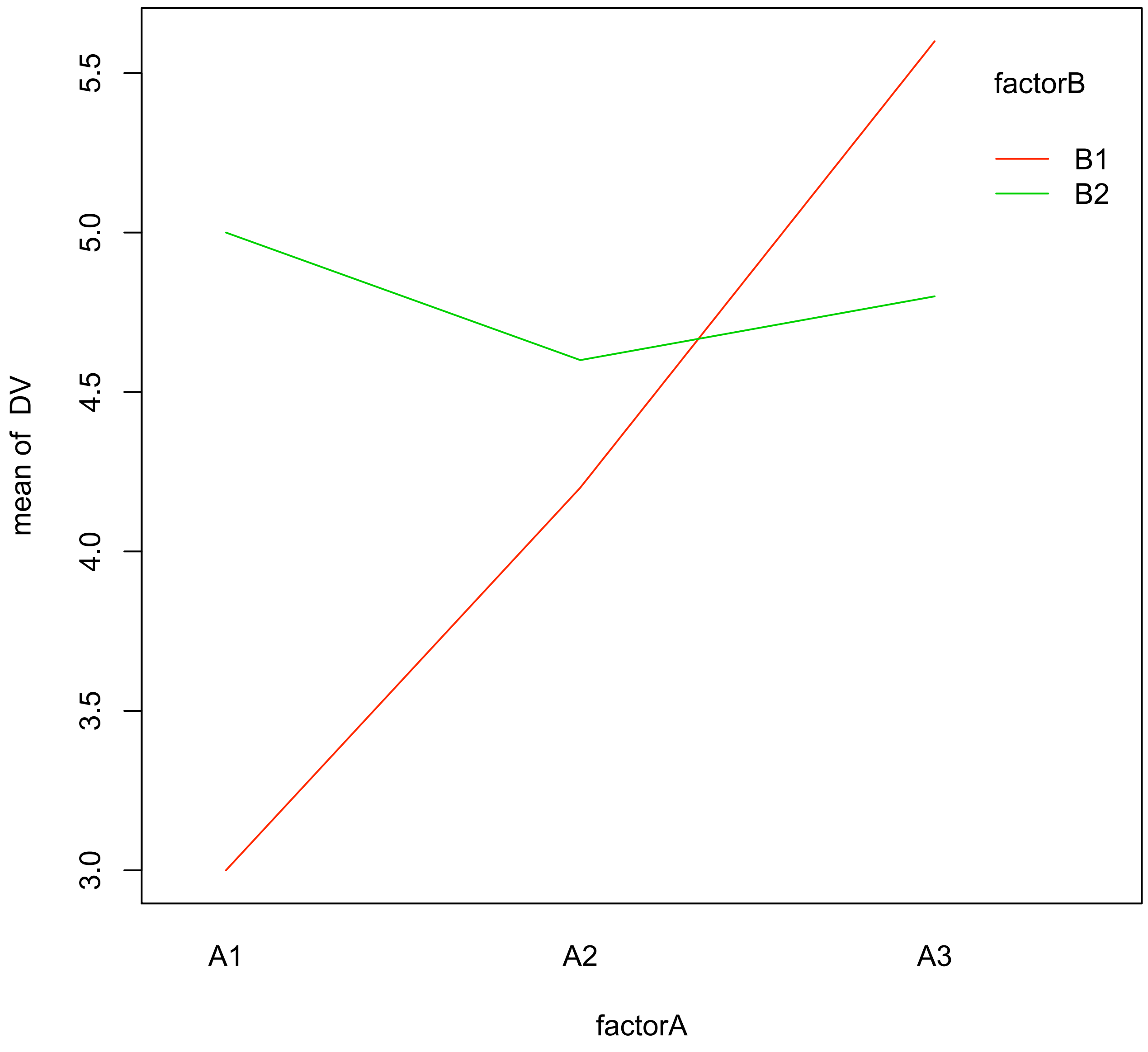
- suppose we are interested in effects of various treatments for hypertension: biofeedback vs drugs X, Y, Z (vs nothing)
- is it better to conduct a 2 x 3 factorial study OR two separate single-factor studies?
- factorial design enables us to test for an interaction
- factorial design allows for greater generalizability

★ factorial design can produce the **same statistical power** as 2 single-factor designs using **half as many subjects**!

An example using R

Group	B1	B2
A1	2,3,4,3,3 (**3.00**)	4,5,6,5,5 (**5.00**)
A2	3,4,5,4,5 (**4.20**)	6,5,4,4,4 (**4.60**)
A3	4,6,5,6,7 (**5.6**)	5,4,6,5,4 (**4.8**)

http://www.gribblelab.org/stats2019/code/twoWay.R
http://www.gribblelab.org/stats2019/data/2waydata.csv



factorB
B1
B2
mean of DV
3.0
3.5
4.0
4.5
5.0
5.5
A1
A2
A3
factorA

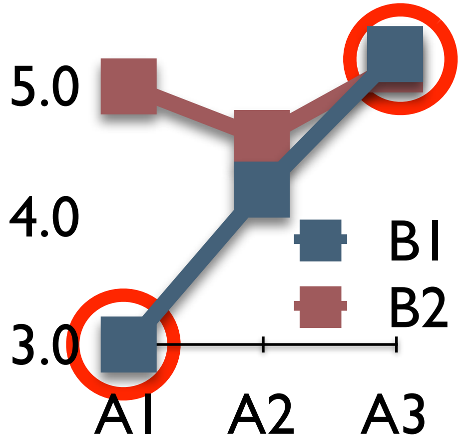
```
summary(aov(DV~factorA*factorB))
```

```
                Df  Sum Sq Mean Sq F value   Pr(>F)
factorA          2  7.4667  3.7333  4.9778 0.015546 *
factorB          1  2.1333  2.1333  2.8444 0.104648
factorA:factorB  2  9.8667  4.9333  6.5778 0.005275 **
Residuals       24 18.0000  0.7500
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- what now? possibilities:
- "simple effects" (mini-anova) of A within B1 & within B2
- simple effects of B within A1, within A2 and within A3
- or just go directly to pairwise contrasts

```
                  Df  Sum Sq Mean Sq F value   Pr(>F)
factorA            2  7.4667  3.7333  4.9778 0.015546 *
factorB            1  2.1333  2.1333  2.8444 0.104648
factorA:factorB    2  9.8667  4.9333  6.5778 0.005275 **
Residuals         24 18.0000  0.7500
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- as a demonstration, let’s do the following contrast within B1
  - A1 vs A3
- and the same contrast within B2
  - A1 vs A3
- strategy for controlling Type-I error?
  - how about since we are doing 2 tests we divide each alpha by 2



$$\psi = \sum_{j=1}^{a} c_j \mu_j$$

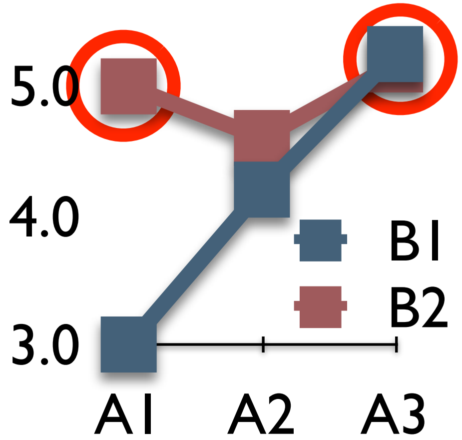
$$SS_\psi = n(\psi)^2 / \sum_{j=1}^{a} c_j^2$$

$$F = \frac{SS_{contrast}/df_{contrast}}{MS_W}$$

```
                 Df  Sum Sq Mean Sq F value   Pr(>F)
factorA           2  7.4667  3.7333  4.9778 0.015546 *
factorB           1  2.1333  2.1333  2.8444 0.104648
factorA:factorB   2  9.8667  4.9333  6.5778 0.005275 **
Residuals        24 18.0000  0.7500
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- for A1 vs A3 **within B1**
- psi = (+1)(3.00) + (-1)(5.6) = -2.6
- SS = 5((-2.6)^2) / ((+1)^2 + (-1)^2) = 33.8 / 2 = 16.9
- df_contrast = 1
- MS_W = 0.75; df_denom = 24 (from ANOVA table)
- Fobs = 16.9 / 0.75 = **22.53**
- `pf(22.5333,1,24,lower.tail=F) -> ` **`p=0.000079`**

uncorrected for Type-I error



$$\psi = \sum_{j=1}^{a} c_j \mu_j$$

$$SS_\psi = n(\psi)^2 / \sum_{j=1}^{a} c_j^2$$

$$F = \frac{SS_{contrast}/df_{contrast}}{MS_W}$$

```
                 Df  Sum Sq Mean Sq F value   Pr(>F)
factorA           2  7.4667  3.7333  4.9778 0.015546 *
factorB           1  2.1333  2.1333  2.8444 0.104648
factorA:factorB   2  9.8667  4.9333  6.5778 0.005275 **
Residuals        24 18.0000  0.7500
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- for A1 vs A3 **within B2**
- psi = (+1)(5.00) + (-1)(4.8) = 0.2
- SS = 5((0.2)^2) / ((+1)^2 + (-1)^2) = 0.2 / 2 = 0.1
- df_contrast = 1
- MS_W = 0.75; df_denom = 24 (from ANOVA table)
- Fobs = 0.1 / 0.75 = **0.133**
- `pf(0.133,1,24,lower.tail=F) -> ` **`p=0.719`**

uncorrected for Type-I error

Controlling Alpha Level

- As we saw there are other approaches
- If you are following up tests based on how the data look post-hoc, perhaps you would feel more comfortable using Tukey tests instead
- If you are performing a whole bunch of planned tests then perhaps Bonferroni will actually be too conservative and you might feel better using Scheffé
- Here is the rule to follow:
- you must have some well defined and well understood rationale for how (or if) you control for Type-I error

tukeyHSD(myanova)

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = DV ~ factorA * factorB, data = mydata)

$factorA
      diff        lwr      upr     p adj
A2-A1  0.4 -0.5671951 1.367195 0.5639204
A3-A1  1.2  0.2328049 2.167195 0.0131180
A3-A2  0.8 -0.1671951 1.767195 0.1185021

$factorB
           diff        lwr      upr     p adj
B2-B1 0.5333333 -0.1193286 1.185995 0.1046482

$`factorA:factorB`
            diff         lwr       upr     p adj
A2:B1-A1:B1  1.2 -0.49352039 2.8935204 0.2782133
A3:B1-A1:B1  2.6  0.90647961 4.2935204 0.0009965
A1:B2-A1:B1  2.0  0.30647961 3.6935204 0.0141717
A2:B2-A1:B1  1.6 -0.09352039 3.2935204 0.0717436
A3:B2-A1:B1  1.8  0.10647961 3.4935204 0.0326534
A3:B1-A2:B1  1.4 -0.29352039 3.0935204 0.1475933
A1:B2-A2:B1  0.8 -0.89352039 2.4935204 0.6911401
A2:B2-A2:B1  0.4 -1.29352039 2.0935204 0.9761219
A3:B2-A2:B1  0.6 -1.09352039 2.2935204 0.8783892
A1:B2-A3:B1 -0.6 -2.29352039 1.0935204 0.8783892
A2:B2-A3:B1 -1.0 -2.69352039 0.6935204 0.4690617
A3:B2-A3:B1 -0.8 -2.49352039 0.8935204 0.6911401
A2:B2-A1:B2 -0.4 -2.09352039 1.2935204 0.9761219
A3:B2-A1:B2 -0.2 -1.89352039 1.4935204 0.9990353
A3:B2-A2:B2  0.2 -1.49352039 1.8935204 0.9990353
```

3-Factor ANOVA

The 2 x 2 x 2 Design

- same example as last time
- test effects of different therapies for hypertension
- last time: 2 x 2
  - biofeedback (yes/no) x drug therapy (yes/no)
- now add a 3rd factor: diet therapy (yes/no)
- 3 factor design: 2 x 2 x 2
- subjects randomly assigned to one of 8 possible groups

	Diet Absent		**Diet Present**	
	Biofeedback Present	*Biofeedback Absent*	*Biofeedback Present*	*Biofeedback Absent*
Drug Present	180	205	170	190
Drug Absent	200	210	185	190

The 2 x 2 x 2 Design

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A
  - main effect of B

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A
  - main effect of B
  - main effect of C

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A
  - main effect of B
  - main effect of C
- Three 2-Way Interaction Effects

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A
  - main effect of B
  - main effect of C
- Three 2-Way Interaction Effects
  - AB interaction

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A
  - main effect of B
  - main effect of C
- Three 2-Way Interaction Effects
  - AB interaction
  - AC interaction

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A
  - main effect of B
  - main effect of C
- Three 2-Way Interaction Effects
  - AB interaction
  - AC interaction
  - BC interaction

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A
  - main effect of B
  - main effect of C
- Three 2-Way Interaction Effects
  - AB interaction
  - AC interaction
  - BC interaction
- One 3-Way Interaction Effect

The 2 x 2 x 2 Design

- There are **7 effects** in a 3 Factor design:
- Three Main Effects
  - main effect of A
  - main effect of B
  - main effect of C
- Three 2-Way Interaction Effects
  - AB interaction
  - AC interaction
  - BC interaction
- One 3-Way Interaction Effect
  - ABC interaction

Main Effects

- main effect for a factor involves comparing the levels of that factor after averaging over all other factors
- e.g. main effect of Factor A (biofeedback):
  - average over levels of B and C
  - marginal means for Factor A are:
  - **Biofeedback Present**: (180 + 200 + 170 + 185)/4 = 183.75
  - **Biofeedback Absent**: (205 + 210 + 190 + 190)/4 = 198.75
- Main effect of B and of C in a similar fashion

C

	Diet Absent		Diet Present	
B	A *Biofeedback Present*	*Biofeedback Absent*	*Biofeedback Present*	*Biofeedback Absent*
Drug Present	180	205	170	190
Drug Absent	200	210	185	190

Two-Way Interactions

Two-Way Interactions

- AB Interaction
  - average over Factor C
  - when averaged over Factor C, the effect of Factor A is different depending on the level of Factor B

Two-Way Interactions

- AB Interaction
  - average over Factor C
  - when averaged over Factor C, the effect of Factor A is different depending on the level of Factor B
- AC Interaction
  - average over Factor B
  - when averaged over Factor B, the effect of Factor A is different depending on the level of Factor C

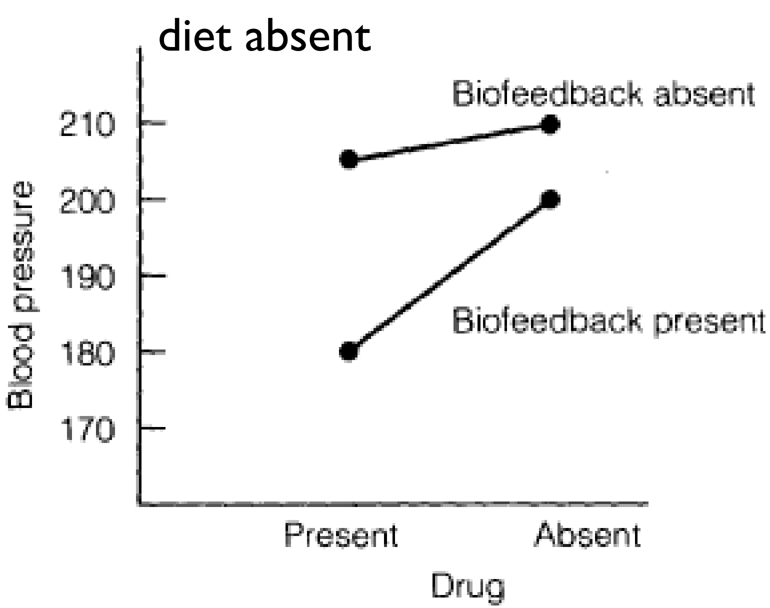
Two-Way Interactions

- AB Interaction
  - average over Factor C
  - when averaged over Factor C, the effect of Factor A is different depending on the level of Factor B
- AC Interaction
  - average over Factor B
  - when averaged over Factor B, the effect of Factor A is different depending on the level of Factor C
- BC Interaction
  - average over Factor A
  - when averaged over Factor A, the effect of Factor B is different depending on the level of Factor C

Three-Way Interaction

- review: meaning of a two-way interaction (e.g. AB)
  - the Main Effect of A is different depending on the level of B
- meaning of a three-way interaction (e.g. ABC)
  - the AB interaction is different depending on the level of C
  - or
  - the AC interaction is different depending on the level of B
  - or
  - the BC interaction is different depending on the level of A
  - (all are equivalent statements)
  - some may have greater meaning than others in the context of your particular experiment

- I find it easiest to understand three-way interactions by referring to a graphical display of the data
- strategy: plot the two-way interaction multiple times, at each level of the third factor
- e.g. plot the drug x biofeedback interaction (1) for the diet absent level and (2) for the diet present level
- <u>the 2-way drug x biofeedback interaction is different when diet is is absent vs when diet is present</u>



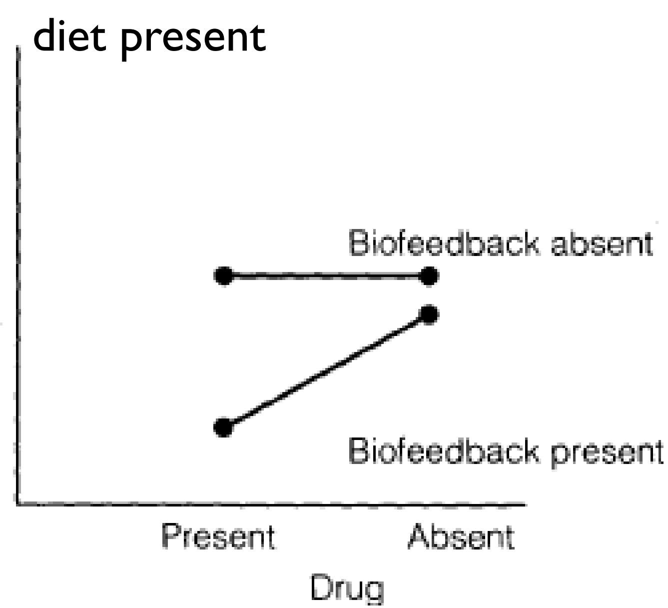


TABLE 8.8 Meaning of Effects in a Three-Way A × B × C Design

	Meaning
Main Effects	
A	Comparison of marginal means of A factor, averaging over levels of B and C
B	Comparison of marginal means of B factor, averaging over levels of A and C
C	Comparison of marginal means of C factor, averaging over levels of A and B
Two-Way Interactions	
A × B	Examines whether the A effect is the same at every level of B, averaging over levels of C (equivalently, examines whether the B effect is the same at every level of A, averaging over levels of C)
A × C	Examines whether the A effect is the same at every level of C, averaging over levels of B (equivalently, examines whether the C effect is the same at every level of A, averaging over levels of B)
B × C	Examines whether the B effect is the same at every level of C, averaging over levels of A (equivalently, examines whether the C effect is the same at every level of B, averaging over levels of A)
Three-Way Interaction	
A × B × C	Examines whether the two-way A × B interaction is the same at every level of C (equivalently, examines whether the two-way A × C interaction is the same at every level of B; equivalently, examines whether the two-way B × C interaction is the same at every level of A)

Model Comparison Approach

- just as before we can write a full model that contains all seven effects
- for each significance test (7 of them) we can write a restricted model in which the effect being tested is absent
- just as before we end up with an F-ratio
- just as before the denominator is equal to the MS_W from the ANOVA table
- **See Chapter 8 M&D for all the details**

Implications of a Three-Way Interaction

- Two-way interactions cannot be interpreted unambiguously
- e.g. there may be a significant two-way interaction between A and B within C1 but not within C2
- **so: do not interpret two-way interactions if the three-way interaction is significant!**
- (just like our previous rule about not interpreting main effects if a two-way interaction is significant)

Implications of a Three-Way Interaction

- Also do not interpret main effects
- effect of one factor depends on the level of BOTH of the other 2 factors
- doesn't make sense to average over levels of the other 2 factors
- in general: rule is: if three-way interaction is significant, do not interpret 2-way interactions OR main effects
- go directly to follow-up tests to understand the nature of the three-way interaction

General Guidelines for Analyzing Effects

- a flowchart is shown in chapter
- looks more complicated than it should
- basic idea: start by looking at highest-order effect (3-way interaction)
- if significant, do follow-up tests within each level of a factor
- if not significant, move down to lower-order effects (2-way interactions)
- repeat

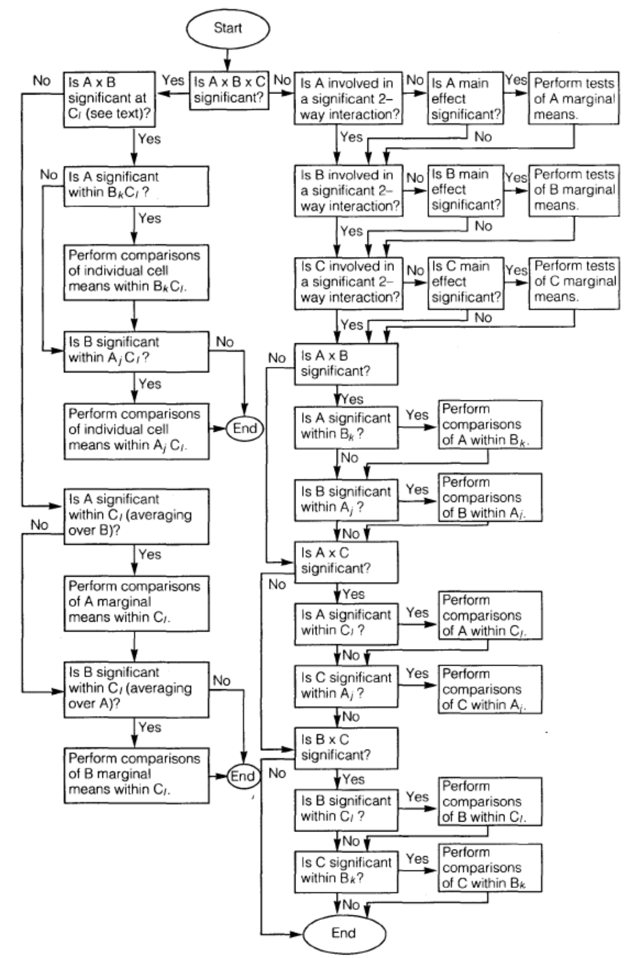


FIGURE 8.2 General guideline for analyzing effects in a three-factor design.

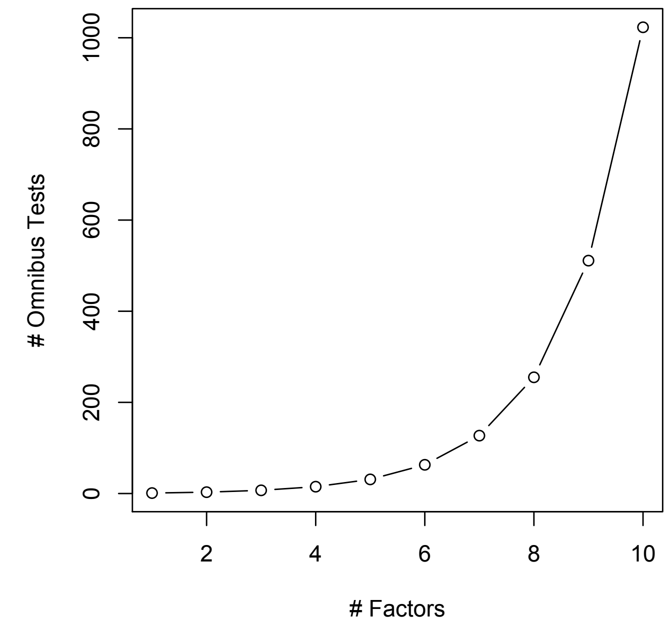
General Guidelines for Analyzing Effects

- issues to consider (just as before)
- for follow-up tests, are they planned or post-hoc?
- how are you going to correct (if you do at all) for Type-I error?
- what's the best way of displaying your data graphically?

Higher-Order Designs

- 4-Factors (**15** omnibus tests)
  - 4 x main effects: A, B, C, D
  - 6 x 2-way interactions: AB, AC, AD, BC, BD, CD
  - 4 x 3-way interactions: ABC, ABD, ACD, BCD
  - 1 x 4-way interaction: ABCD
- # of groups:
  - e.g. A(2) B(2) C(2) D(2) : 2 x 2 x 2 x 2 = 16 groups!
  - e.g. A(3) B(3) C(3) D(3) : 3 x 3 x 3 x 3 = 81 groups!
  - this is ridiculous
- in any case - can you really interpret a 4-way interaction?
- difficult to {visualize, articulate, explain, understand}

Higher-Order Designs



- 4-Factors (**15** omnibus tests)
  - 4 x main effects: A, B, C, D
  - 6 x 2-way interactions: AB, AC, AD, BC, BD, CD
  - 4 x 3-way interactions: ABC, ABD, ACD, BCD
  - 1 x 4-way interaction: ABCD
- # of groups:
  - e.g. A(2) B(2) C(2) D(2) : 2 x 2 x 2 x 2 = 16 groups!
  - e.g. A(3) B(3) C(3) D(3) : 3 x 3 x 3 x 3 = 81 groups!
  - this is ridiculous
- in any case - can you really interpret a 4-way interaction?
- difficult to {visualize, articulate, explain, understand}