PSYCH 710

Between-Subjects Factorial Designs
Main Effects & Interactions

Prof. Patrick Bennett

Crossed-factorial designs

- at least 2 independent variables/factors (A, B, C, etc.)
- each level of one factor combined with all levels of other factors
  - balanced: equal n per cell/condition
- is variation among cell means “statistically significant”?
  - is variation compatible with hypothesis that variation is due to chance?
- factorial ANOVA decomposes variation into main effects & interactions
  - main effect: effect associated with one factor
  - interaction: effect associated with <u>combination</u> of factors

---

Linear Model

subject i, row j, column k

main effects

interaction

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

intercept

residual

		Factor B		
		B₁	B₂	B₃
Factor A	A₁	A₁B₁	A₁B₂	A₁B₃
	A₂	A₂B₁	A₂B₂	A₂B₃
	A₃	A₃B₁	A₃B₂	A₃B₃
	A₄	A₄B₁	A₄B₂	A₄B₃

Main Effect of A

$$\hat{\mu}_{j.} = \bar{Y}_{j.}$$

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

$$\alpha_j \;=\; \mu_{j.} - \mu_{..}$$

$\mu_{j.}$ = marginal row mean [average of cell means in row j]

$\mu_{..}$ = grand mean (all scores)

$$\sum_{j=1}^{a} \alpha_j \;=\; 0$$

df$_A$ = a-1

		Factor B			Marginal Row Means
		B₁	B₂	B₃	
Factor A	A₁	A₁B₁	A₁B₂	A₁B₃	$\bar{Y}_{1.}$
	A₂	A₂B₁	A₂B₂	A₂B₃	$\bar{Y}_{2.}$
	A₃	A₃B₁	A₃B₂	A₃B₃	$\bar{Y}_{3.}$
	A₄	A₄B₁	A₄B₂	A₄B₃	$\bar{Y}_{4.}$

Main Effect of B

$$\hat{\mu}_{.k} = \bar{Y}_{.k}$$

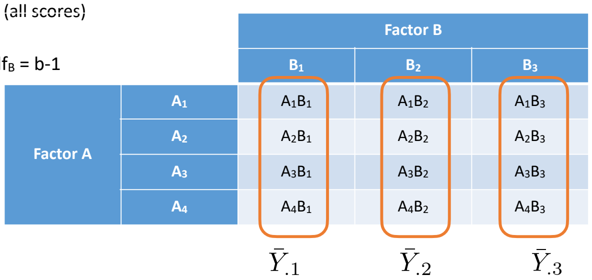
$$Y_{ijk} = \mu + \alpha_j + \boxed{\beta_k} + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

$$\beta_k = \mu_{.k} - \mu_{..}$$

$\mu_{.k}$ = marginal column mean [average of cell means in column k]

$\mu_{..}$ = grand mean (all scores)

$\sum_{k=1}^{b} \beta_k = 0$ df_B = b-1



		Factor B		
		B_1	B_2	B_3
Factor A	A_1	A_1B_1	A_1B_2	A_1B_3
	A_2	A_2B_1	A_2B_2	A_2B_3
	A_3	A_3B_1	A_3B_2	A_3B_3
	A_4	A_4B_1	A_4B_2	A_4B_3
		$\bar{Y}_{.1}$	$\bar{Y}_{.2}$	$\bar{Y}_{.3}$

Marginal Column Means

Interaction Effects (A x B)

$$\sum_{j=1}^{a} (\alpha\beta)_{jk} = 0 \text{ for each value of } k$$

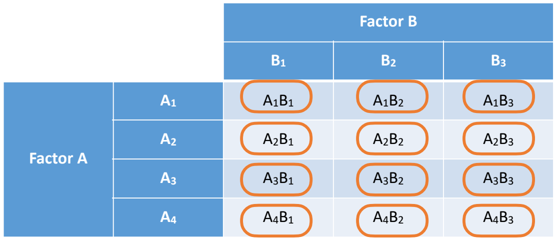
$$\sum_{k=1}^{b} (\alpha\beta)_{jk} = 0 \text{ for each value of } j$$

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

$$(\alpha\beta)_{jk} = \mu_{jk} - (\mu_{..} + \alpha_j + \beta_k)$$

i.e., interaction is what is left over after accounting for 2 main effects

df_{AxB} = (a-1)(b-1) = df_A x df_B



		Factor B		
		B_1	B_2	B_3
Factor A	A_1	A_1B_1	A_1B_2	A_1B_3
	A_2	A_2B_1	A_2B_2	A_2B_3
	A_3	A_3B_1	A_3B_2	A_3B_3
	A_4	A_4B_1	A_4B_2	A_4B_3

One interaction effect per cell

Least-squares Estimates of Parameters

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

$$\begin{aligned}
\alpha_j &= \bar{Y}_{j.} - \bar{Y}_{..} \\
\beta_k &= \bar{Y}_{.k} - \bar{Y}_{..} \\
(\alpha\beta)_{jk} &= \bar{Y}_{jk} - (\bar{Y}_{..} + (\bar{Y}_{j.} - \bar{Y}_{..}) + (\bar{Y}_{.k} - \bar{Y}_{..})) \\
&= \bar{Y}_{jk} - \bar{Y}_{j.} - \bar{Y}_{.k} + \bar{Y}_{..}
\end{aligned}$$

$$\bar{Y}_{jk} - \bar{Y}_{..} - \hat{\alpha}_{j.} - \hat{\beta}_{.k}$$

i.e., what is left over after accounting for 2 main effects

degrees-of-freedom:
df_A = (a-1)
df_B = (b-1)
df_{AxB} = (a-1)(b-1)

Interpretation of Model

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

$Y_{ijk} = \hat{Y}_{ijk} + \epsilon_{ijk}$	observed score = predicted score + residual
$\hat{Y}_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}$	predicted score = sum of intercept, main effects, and interaction
$\hat{Y}_{ijk} = \bar{Y}_{ij}$	predicted score = cell mean
$Y_{ijk} = \bar{Y}_{ij} + \epsilon_{ijk}$	observed score = cell mean + residual

anova breaks variation among cell means into 3 independent components: variation due to A, variation due to B, and variation due to AxB

model has 1 + (a-1) + (b-1) + (a-1)(b-1) free parameters, or one for each cell mean

Evaluating Main Effect of A

Null Hypothesis

$$\alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$$

$$\bar{Y}_{1.} = \bar{Y}_{2.} = \cdots = \bar{Y}_{j.}$$

Nested Models

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

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

determine if change in goodness-of-fit (sum-of-squared residuals) is unusually large assuming Null Hypothesis is true

Calculating Sum-of-Squares A (SS_A)

Nested Models

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

$$Y_{ijk} = \mu + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$$ reduced model [violates principle of marginality. Bad!]

$$Y_{ijk} = \mu + \epsilon_{ijk}$$ reduced model

$$Y_{ijk} = \mu + \alpha_j + \epsilon_{ijk}$$ full model (ignores B and AxB)

$$Y_{ijk} = \mu + \beta_k + \epsilon_{ijk}$$ reduced model

$$Y_{ijk} = \mu + \alpha_j + \beta_k + \epsilon_{ijk}$$ full model (ignores AxB)

SS_A is change in $SS_{residuals}$ that occurs when all alphas are set to zero.
For balanced designs, SS_A is the same in all cases.

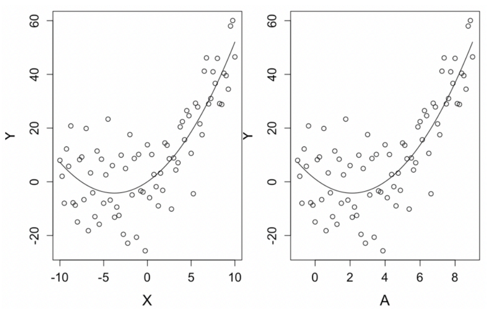
Marginality

- lower-order effects are marginal to higher-order effects
  - main effects are marginal to 2-way their interactions
  - 2-way interactions are marginal to their 3-way interactions
  - linear predictors (x) are marginal to quadratic & cubic predictors (e.g., x^2, x^3)
- principle of marginality: linear models that include higher-order effects/ predictors <u>should not exclude</u> the effects that are marginal to them
  - if a model includes A x B it should also include A and B
  - if a model includes x^2 it should also include x
- why? Because, in general, models that violate marginality are not invariant to <u>changes in measurement units</u>
  - http://pnb.mcmaster.ca/bennett/psy710/notes/mw7-marginality-example.html

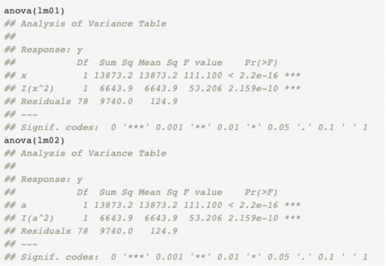
Marginality

http://pnb.mcmaster.ca/bennett/psy710/notes/mw7-marginality-example.html

a <- 4 + 0.5*x



ANOVAs are identical



```
anova(lm01)
## Analysis of Variance Table
##
## Response: y
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## x           1 13873.2 13873.2 111.100 < 2.2e-16 ***
## I(x^2)      1  6643.9  6643.9  53.206 2.159e-10 ***
## Residuals 78  9740.0   124.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
anova(lm02)
## Analysis of Variance Table
##
## Response: y
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## a           1 13873.2 13873.2 111.100 < 2.2e-16 ***
## I(a^2)      1  6643.9  6643.9  53.206 2.159e-10 ***
## Residuals 78  9740.0   124.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

both marginals obey marginality principle

Marginality

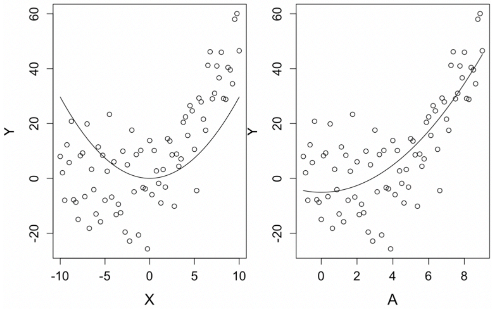
http://pnb.mcmaster.ca/bennett/psy710/notes/mw7-marginality-example.html

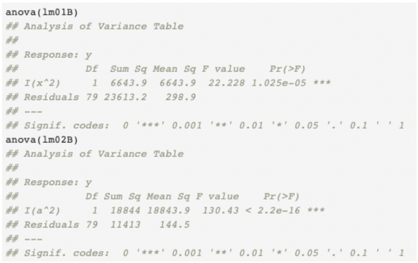
a <- 4 + 0.5*x

lm01 <- lm(y ~1+I(x^2))

lm02 <- lm(y ~1+I(a^2))

ANOVAs are <u>NOT</u> identical





```
anova(lm01B)
## Analysis of Variance Table
##
## Response: y
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## I(x^2)     1  6643.9  6643.9  22.228 1.025e-05 ***
## Residuals 79 23613.2   298.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
anova(lm02B)
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## I(a^2)     1  18844 18843.9  130.43 < 2.2e-16 ***
## Residuals 79  11413   144.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

both marginals violate marginality principle

Evaluating Main Effect of A

$$\alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$$

$$\bar{Y}_{1.} = \bar{Y}_{2.} = \cdots = \bar{Y}_{j.}$$

[null hypothesis]

$$SS_A = E_R - E_F = \sum_{j=1}^{a} \sum_{i=1}^{n} \alpha_j^2$$

$$SS_{Residuals} = E_F = \sum_{j=1}^{a} \sum_{k=1}^{b} \sum_{i=1}^{n} (Y_{ijk} - \bar{Y}_{jk})^2$$

[calculated using full model]

$$F_A = \frac{SS_A/(a-1)}{SS_{Residuals}/(ab(n-1))}$$

$F_A = MS_A \div MS_{residuals}$

df = {(a-1), ab(n-1)}

Calculating Sum-of-Squares B (SS$_B$)

Nested Models

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

full model

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

reduced model [violates principle of marginality. Bad!]

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

reduced model

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

full model (ignores A and AxB)

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

reduced model

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

full model (ignores AxB)

SS$_B$ is change in SS$_{residuals}$ that occurs when all betas are set to zero.
For balanced designs, SS$_B$ is the same in all cases.

Evaluating Main Effect of B

$$\beta_{.1} = \beta_{.2} = \cdots = \beta_{.b} = 0$$

$$\bar{Y}_{.1} = \bar{Y}_{.2} = \cdots = \bar{Y}_{.b} = 0$$

[null hypothesis]

$$SS_B = \sum_{k=1}^{b} \sum_{i=1}^{n} \beta_k^2$$

$$SS_{Residuals} = E_F = \sum_{j=1}^{a} \sum_{k=1}^{b} \sum_{i=1}^{n} (Y_{ijk} - \bar{Y}_{jk})^2$$

[calculated using full model]

$$F_B = \frac{SS_B/(b-1)}{SS_{Residual}/(ab(n-1))}$$

$F_B = MS_B \div MS_{residuals}$

df = {(b-1), ab(n-1)}

Calculating Sum-of-Squares AxB (SS_{AxB})

Nested Models

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

$$Y_{ijk} = \mu + \alpha_j + \beta_k + \epsilon_{ijk}$$ reduced model [does not violate principle of marginality!]

SS_{AxB} is change in $SS_{residuals}$ that occurs when all interaction effects are set to zero.

Evaluating AxB Interaction

$$(\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{ab} = 0$$ [null hypothesis]

$$SS_{AB} = \sum_{j=1}^{a}\sum_{k=1}^{b}\sum_{i=1}^{n}(\alpha\beta)_{jk}^2$$

$$SS_{Residuals} = E_F = \sum_{j=1}^{a}\sum_{k=1}^{b}\sum_{i=1}^{n}(Y_{ijk} - \bar{Y}_{jk})^2$$ [calculated using full model]

$$F_{AB} = \frac{SS_{AB}/((a-1)(b-1))}{SS_{Residual}/(ab(n-1))}$$

$F_{AxB} = MS_{AxB} \div MS_{residuals}$

df = {(a-1)(b-1), ab(n-1)}

ANOVA table

	SS	df	MS	F	p
A	SS_A	df_A = a-1	$MS_A = SS_A/df_A$	MS_A/MS_W	
B	SS_B	df_B = b-1	$MS_B = SS_B/df_B$	MS_B/MS_W	
A x B	SS_{AxB}	df_{AxB} = (a-1)(b-1)	$MS_{AxB} = SS_{AxB}/df_{AxB}$	MS_{AxB}/MS_W	
Within Cell (residuals)	SS_W	df_W = ab(n-1)	$MS_W = SS_W/df_W$		

Interpreting Interactions

Interpreting AxB Interaction

- interaction effect is the difference between cell mean and the sum of the intercept & main effects:

$$(\alpha\beta)_{ij} = \bar{Y}_{ij} - \mu - \alpha_j - \beta_k$$

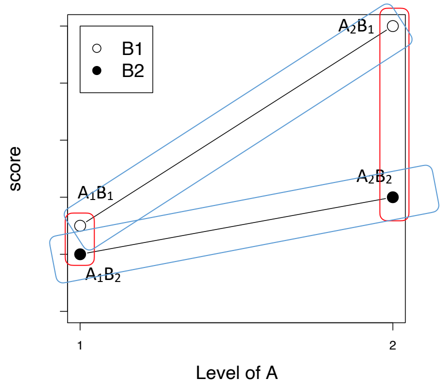
- if interaction effects are zero, then differences among cell means are due differences between alpha's and/or beta's (row & column effects)
- if interactions are not zero then effect of A depends on level of B (and effect of B depends on level of A)

Graphical Depiction of AxB Interaction

	Factor B	
Factor A	A_1B_1	A_1B_2
	A_2B_1	A_2B_2

AxB interaction implies that the effect of A <u>depends on</u> the level of B, and the effect of B <u>depends on</u> the level of A

$$(A_1B_1 - A_1B_2) \neq (A_2B_1 - A_2B_2)$$
$$(A_1B_1 - A_2B_1) \neq (A_1B_2 - A_2B_2)$$



Interactions

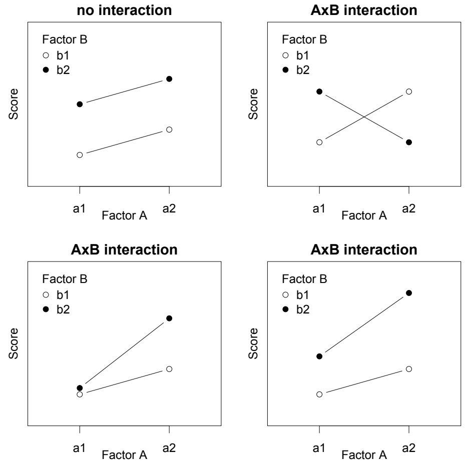
- 2-way (A x B) interactions assess whether the effect of one factor depends on the level of the other factor
  - Does the effect of A depend on the level of B?
  - Does the effect of B depend on the level of A?
- 3-way (A x B x C) interactions assess whether the <u>interaction</u> between 2 factors depends on the level of the 3rd factor
  - Does the AxB interaction depend on the level of C?
  - Does the AxC interaction depend on the level of B?
  - Does the BxC interaction depend on the level of A?

Graphical representation of 2-way interactions

Significant AxB interaction implies:

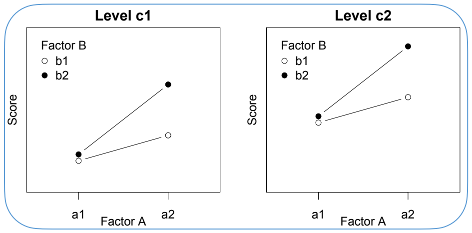
- Difference between a1 & a2 depends on level of B.
- Difference between b1 & b2 depends on level of A.

Significant 2-way interactions imply a significant deviation from parallelism.



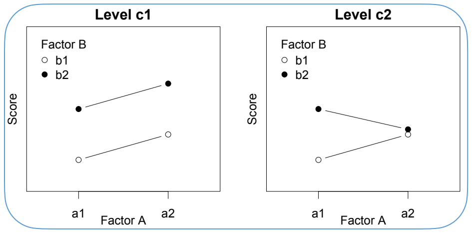
Graphical representation of 3-way interactions

- Significant AxB interaction
- Non-significant AxBxC interaction



Significant 3-way interactions imply that the deviation from parallelism in a 2-way interaction depends on the level of the 3rd factor

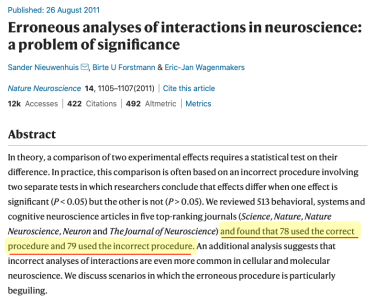
- Significant AxBxC interaction
- AxB interaction depends on level of C
- Significant AxB interaction in c2 but not c1



Incorrect Interpretations of Interactions

common problem in published scientific papers

- It often is interesting to find interactions
- But often they are “discovered” with faulty statistical methods



Published: 26 August 2011

Erroneous analyses of interactions in neuroscience: a problem of significance

Sander Nieuwenhuis, Birte U Forstmann & Eric-Jan Wagenmakers

Nature Neuroscience **14**, 1105–1107(2011) | Cite this article

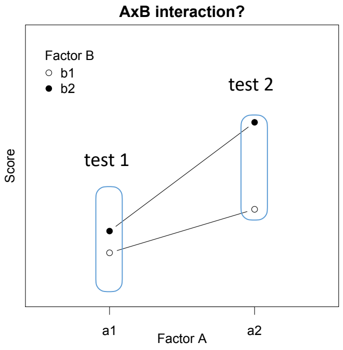
12k Accesses | **422** Citations | **492** Altmetric | Metrics

Abstract

In theory, a comparison of two experimental effects requires a statistical test on their difference. In practice, this comparison is often based on an incorrect procedure involving two separate tests in which researchers conclude that effects differ when one effect is significant ($P < 0.05$) but the other is not ($P > 0.05$). We reviewed 513 behavioral, systems and cognitive neuroscience articles in five top-ranking journals (*Science*, *Nature*, *Nature Neuroscience*, *Neuron* and *The Journal of Neuroscience*) and found that 78 used the correct procedure and 79 used the incorrect procedure. An additional analysis suggests that incorrect analyses of interactions are even more common in cellular and molecular neuroscience. We discuss scenarios in which the erroneous procedure is particularly beguiling.

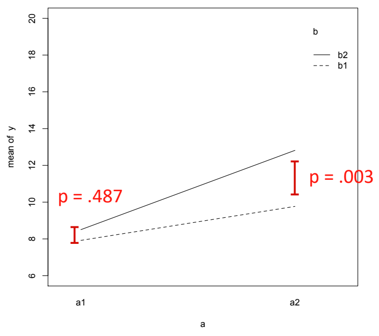
Incorrect Interpretations of Interactions

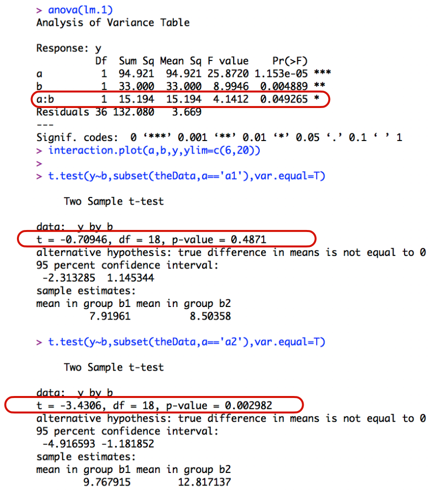
- Interaction determines if <u>difference between 2 differences</u> is significant
  - Is {(b_2-b_1) at a_1} <u>minus</u> {(b_2-b_1) at a_2} significantly different from zero?
- Not the same as doing <u>separate</u> tests of (b_2-b_1) at a_1 and (b_2-b_1) at a_2
- Example: if t-test 1 is not significant but t-test 2 is significant, will AxB interaction be significant?
  - not necessarily…



---

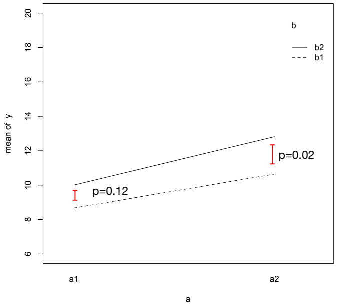
only 1 t-test is significant and AxB interaction <u>is</u> significant





```
> anova(lm.1)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)    
a          1  94.921  94.921 25.8720 1.153e-05 ***
b          1  33.000  33.000  8.9946  0.004889 ** 
a:b        1  15.194  15.194  4.1412  0.049265 *  
Residuals 36 132.080   3.669                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> interaction.plot(a,b,y,ylim=c(6,20))
> 
> t.test(y~b,subset(theData,a=='a1'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -0.70946, df = 18, p-value = 0.4871
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.313285  1.145344
sample estimates:
mean in group b1 mean in group b2 
         7.91961          8.50358 

> t.test(y~b,subset(theData,a=='a2'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -3.4306, df = 18, p-value = 0.002982
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.916593 -1.181852
sample estimates:
mean in group b1 mean in group b2 
        9.767915        12.817137 
```

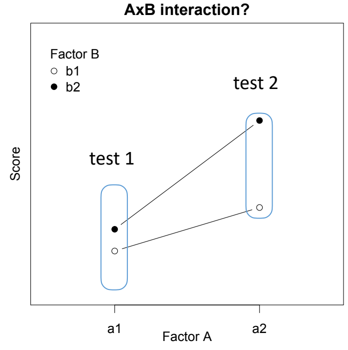
only 1 t-test is significant and AxB interaction is not significant



```
> anova(lm.1)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)    
a          1  57.285  57.285 15.6138 0.0003473 ***
b          1  30.769  30.769  8.3863 0.0063906 ** 
a:b        1   1.765   1.765  0.4811 0.4923813    
Residuals 36 132.080   3.669                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> interaction.plot(a,b,y,ylim=c(6,20))
> 
> t.test(y~b,subset(theData,a=='a1'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -1.6206, df = 18, p-value = 0.1225
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.0632850  0.3953442
sample estimates:
mean in group b1 mean in group b2 
         8.66961         10.00358 

> t.test(y~b,subset(theData,a=='a2'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -2.4462, df = 18, p-value = 0.02494
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.0415928 -0.3068517
sample estimates:
mean in group b1 mean in group b2 
        10.64291         12.81714 
```

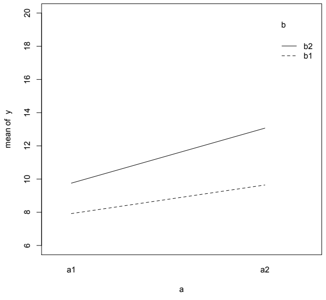
Incorrect Interpretations of Interactions

- Suppose both tests are significant… does that mean that the effect of B does not depend on A?
- i.e., that the AxB interaction is not significant?
  - not necessarily



---

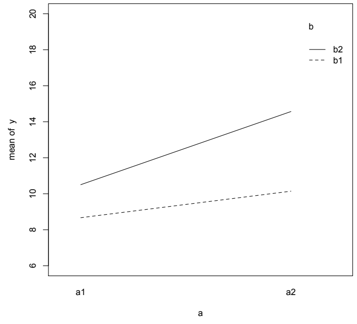
both t-tests are significant and AxB interaction is not significant



```
> anova(lm.1)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)    
a          1  63.425  63.425 17.2872 0.0001897 ***
b          1  69.121  69.121 18.8399 0.0001104 ***
a:b        1   6.322   6.322  1.7232 0.1975881    
Residuals 36 132.080   3.669                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> interaction.plot(a,b,y,ylim=c(6,20))
> 
> t.test(y~b,subset(theData,a=='a1'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -2.2281, df = 18, p-value = 0.03887
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.5632850 -0.1046558
sample estimates:
mean in group b1 mean in group b2 
         7.91961          9.75358 

> t.test(y~b,subset(theData,a=='a2'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -3.8525, df = 18, p-value = 0.001167
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -5.291593 -1.556852
sample estimates:
mean in group b1 mean in group b2 
        9.642915        13.067137 
```

---

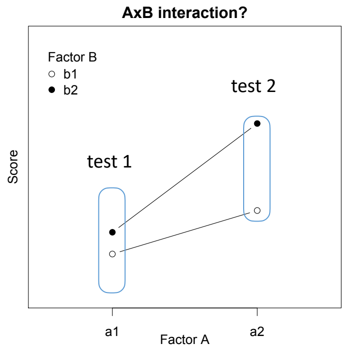
both t-tests are significant and AxB interaction is significant



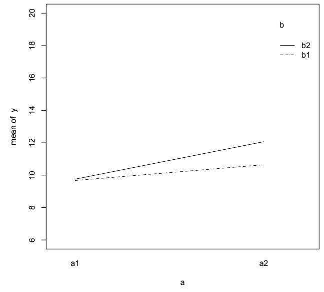
```
> anova(lm.1)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)    
a          1  76.642  76.642 20.8897 5.536e-05 ***
b          1  97.912  97.912 26.6872 9.039e-06 ***
a:b        1  16.774  16.774  4.5718   0.03936 *  
Residuals 36 132.080   3.669                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> interaction.plot(a,b,y,ylim=c(6,20))
> 
> t.test(y~b,subset(theData,a=='a1'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -2.2281, df = 18, p-value = 0.03887
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.5632850 -0.1046558
sample estimates:
mean in group b1 mean in group b2 
         8.66961         10.50358 

> t.test(y~b,subset(theData,a=='a2'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -4.9776, df = 18, p-value = 9.746e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -6.291593 -2.556852
sample estimates:
mean in group b1 mean in group b2 
        10.14291         14.56714 
```

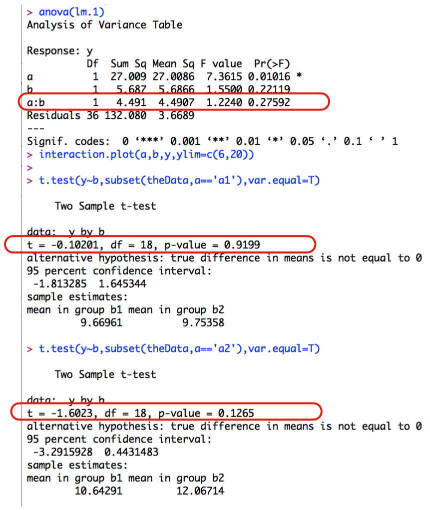
Incorrect Interpretations of Interactions

- Suppose that both tests are not significant… does that mean that the effect of B does not depend on A?
- i.e., that the AxB interaction is not significant?
  - not necessarily



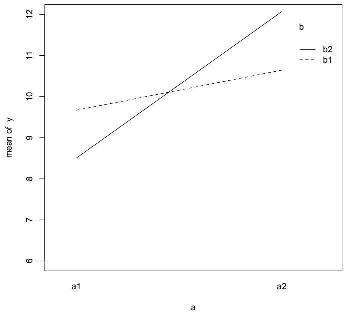
both t-tests are not significant and AxB interaction is not significant

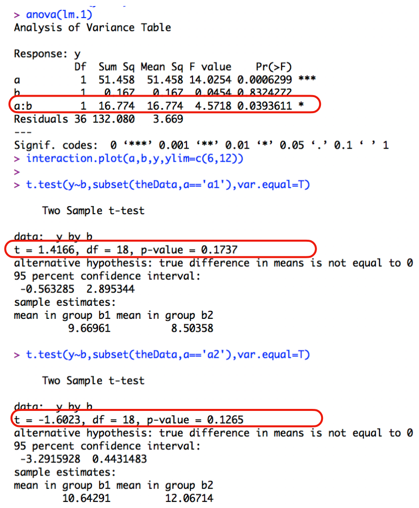




```
> anova(lm.1)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value  Pr(>F)
a          1  27.009 27.0086  7.3615 0.01016 *
b          1   5.687  5.6866  1.5500 0.22119
a:b        1   4.491  4.4907  1.2240 0.27592
Residuals 36 132.080  3.6689
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> interaction.plot(a,b,y,ylim=c(6,20))
>
> t.test(y~b,subset(theData,a=='a1'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -0.10201, df = 18, p-value = 0.9199
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.813285  1.645344
sample estimates:
mean in group b1 mean in group b2
         9.66961          9.75358

> t.test(y~b,subset(theData,a=='a2'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -1.6023, df = 18, p-value = 0.1265
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.2915928  0.4431483
sample estimates:
mean in group b1 mean in group b2
        10.64291         12.06714
```

both t-tests are not significant and AxB interaction is significant





```
> anova(lm.1)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
a          1  51.458  51.458 14.0254 0.0006299 ***
b          1   0.167   0.167  0.0454 0.8324272
a:b        1  16.774  16.774  4.5718 0.0393611 *
Residuals 36 132.080   3.669
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> interaction.plot(a,b,y,ylim=c(6,12))
>
> t.test(y~b,subset(theData,a=='a1'),var.equal=T)

	Two Sample t-test

data:  y by b
t = 1.4166, df = 18, p-value = 0.1737
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.563285  2.895344
sample estimates:
mean in group b1 mean in group b2
         9.66961          8.50358

> t.test(y~b,subset(theData,a=='a2'),var.equal=T)

	Two Sample t-test

data:  y by b
t = -1.6023, df = 18, p-value = 0.1265
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.2915928  0.4431483
sample estimates:
mean in group b1 mean in group b2
        10.64291         12.06714
```

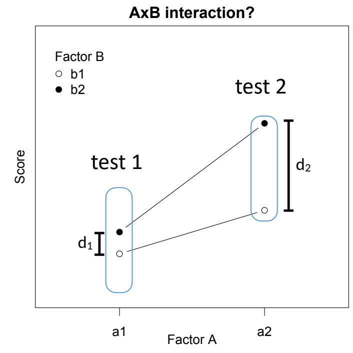
A x B ≠ Pairwise tests of differences

Interaction

	test 1 & test 2 are not significant	test 1 & test 2 are significant	test 1 OR test 2 is significant
$d_1 = d_2$	✓	✓	✓
$d_1 \neq d_2$	✓	✓	✓

Knowing if AxB interaction is significant does not inform us about pairwise tests.

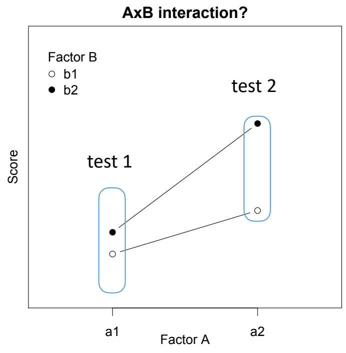
Knowing if pairwise tests are significant does not inform us about AxB interaction.



Incorrect Interpretations of Interactions

Nieuwehnuis, Forstmann, & Wagenmakers, Nature Neurosci, 2011, 14, 1105-1107

- To determine if the effect of one variable depends on another…
  - e.g., if the effect of one variable differs between groups or ages or genders
- …you need to assess the interaction between the 2 variables
- do not rely on significance tests performed separately on the different groups



R example

7.6 an example

In this section I will illustrate how to analyze data collected in an experiment that used a balanced factorial design. There are two *factors*, *A* and *B*, and each factor has two levels. The dependent variable is denoted by y. Finally, there are 6 scores per cell. Here is how I initialized R and created the fake data:

```
> options(contrasts=c("contr.sum","contr.poly"))
> a<-rep(c(-1,-1,1,1),each=6)
> b<-rep(c(-1,1,-1,1),each=6)
> ab<-rep(c(-1,1,1,-1),each=6)
> y<-10+2*a+1*b+0.5*ab
> set.seed(123456);
> nz<-rnorm(y)
> y<-y+nz;
> af<-factor(a,labels=c("a1","a2"),ordered=F)
> bf<-factor(b,labels=c("b1","b2"),ordered=F)
> myData <- data.frame(y,af,bf)
> names(myData) <- c("y","A","B")
```

Sequential Sums of Squares

```
> lm.full.model <- lm(y ~ 1 + A + B + A:B, data=myData)
> anova(lm.full.model)
```

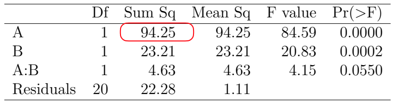
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	94.25	94.25	84.59	0.0000
B	1	23.21	23.21	20.83	0.0002
A:B	1	4.63	4.63	4.15	0.0550
Residuals	20	22.28	1.11		

Table 1: ANOVA table for full model.

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

SS_A obtained with a different comparison

```
> lm.01 <- lm(y ~ 1, data=myData)
> lm.02.a <- lm(y ~ 1 + A, data=myData)
> anova(lm.01, lm.02.a)
```



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	94.25	94.25	84.59	0.0000
B	1	23.21	23.21	20.83	0.0002
A:B	1	4.63	4.63	4.15	0.0550
Residuals	20	22.28	1.11		

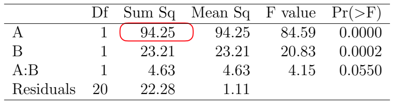
Table 1: ANOVA table for full model.

```
Analysis of Variance Table

Model 1: y ~ 1
Model 2: y ~ 1 + A
  Res.Df     RSS Df Sum of Sq      F    Pr(>F)
1     23 144.370
2     22  50.124  1    94.246 41.366 1.801e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note that SS_A is the same value as the one listed in full ANOVA table.
F differs because the estimate of population error variance (the F's denominator) differs.

SS_A obtained with a third comparison

```
> lm.02.b <- lm(y ~ 1 + B, data=myData)
> lm.03 <- lm(y ~ 1 + A + B, data=myData)
> anova(lm.02.b, lm.03)
```



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	94.25	94.25	84.59	0.0000
B	1	23.21	23.21	20.83	0.0002
A:B	1	4.63	4.63	4.15	0.0550
Residuals	20	22.28	1.11		

Table 1: ANOVA table for full model.

```
Analysis of Variance Table

Model 1: y ~ 1 + B
Model 2: y ~ 1 + A + B
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1     22 121.16
2     21  26.91  1    94.246 73.549 2.654e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note that SS_A is the same value as the one listed previously.
Again, F differs because the estimate of population error variance differs across comparisons.

SS_A computed with 4th comparison of nested models

- in theory, we could drop A from full model
  - full: y ~ 1 + A + B + A:B
  - reduced: y ~ 1 + B + A:B
  - $SS_A = \triangle SS_{resid}$
- However, this comparison does not work (easily)

```
> lm.04 <- lm(y ~ 1 + B + A:B,data=myData)
> anova(lm.04, lm.full.model)
```

R and Marginality

```
> lm.04 <- lm(y ~ 1 + B + A:B,data=myData)
> anova(lm.04, lm.full.model)

Analysis of Variance Table

Model 1: y ~ 1 + B + A:B
Model 2: y ~ 1 + A + B + A:B
  Res.Df    RSS Df  Sum of Sq F Pr(>F)
1     20 22.284
2     20 22.284  0 3.5527e-15
```

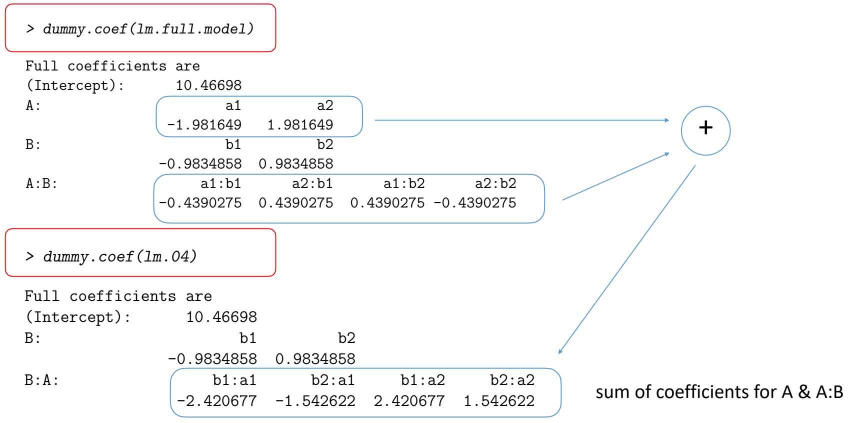
Notice that df hasn't changed although we dropped A from model.

R and Marginality

df for A has been incorporated into AxB interaction

```
> anova(lm.04)

Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
B          1 23.214  23.214  20.835 0.0001883 ***
B:A        2 98.872  49.436  44.370  4.43e-08 ***
Residuals 20 22.284   1.114
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R and Marginality



```
> dummy.coef(lm.full.model)
```

```
Full coefficients are
(Intercept):       10.46698
A:                       a1         a2
                  -1.981649   1.981649
B:                       b1         b2
                 -0.9834858  0.9834858
A:B:                  a1:b1      a2:b1      a1:b2      a2:b2
                 -0.4390275  0.4390275  0.4390275 -0.4390275
```

```
> dummy.coef(lm.04)
```

```
Full coefficients are
(Intercept):        10.46698
B:                        b1          b2
                  -0.9834858   0.9834858
B:A:                   b1:a1       b2:a1       b1:a2       b2:a2
                   -2.420677   -1.542622    2.420677    1.542622
```

sum of coefficients for A & A:B

R and Marginality

- R intentionally makes it difficult (though not impossible) to violate marginality
- later we will use drop1 command to evaluate models that violate marginality

---

alternatives to lm & aov

```
library(afex)
N <- dim(myData)[1] # number of rows/subjects
myData$subj <- factor(x=1:N,labels="s")
car.model.01 <- aov_car(y~1+A+B+A:B+Error(subj),data=myData)
ez.model.01 <- aov_ez(id="subj",dv="y",between=c("A","B"),data=myData)
# nice(car.model.01) # this command gives same anova table as next line...
nice(ez.model.01,es="pes") # pes is partial eta-squared


## Anova Table (Type 3 tests)
##
## Response: y
##   Effect    df  MSE         F  pes p.value
## 1      A 1, 20 1.11 84.59 *** .809   <.001
## 2      B 1, 20 1.11 20.83 *** .510   <.001
## 3    A:B 1, 20 1.11    4.15 + .172    .055
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

decomposing a main effect

analyses of “sub-effects”

Analyzing a Significant Main Effect

```
> mw.aov.01<-aov(score˜drug*biofeedback,data=mw75)
> summary(mw.aov.01)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
drug	2.00	1882.07	941.03	5.21	0.01
biofeedback	1.00	1904.03	1904.03	10.54	0.00
drug:biofeedback	2.00	1248.07	624.03	3.46	0.05
Residuals	24.00	4334.80	180.62		

Table 3: ANOVA for biofeedback experiment.

Tukey HSD [all pairs of marginal means]

```
TukeyHSD(mw.aov.02,which="drug")


##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = score ˜ drug * biofeedback, data = mw75)
##
## $drug
##              diff       lwr   upr  p adj
## drugY-drugX 18.1   3.09063 33.11 0.0160
## drugZ-drugX 15.1   0.09063 30.11 0.0484
## drugZ-drugY -3.0 -18.00937 12.01 0.8724
```

Analyzing a Significant Main Effect

testing sub-effects with linear contrasts

```
# levels(mw75£drug) # "drugX""drugY""drugZ"
c1 <- c(-1,1,0)
c2 <- c(-1,-1,2)
myC <- cbind(c1,c2)
contrasts(mw75$drug) <- myC
mw.aov.02<-aov(score˜drug*biofeedback,data=mw75)
summary(mw.aov.02,split=list(drug=list(1,2)))


##                           Df Sum Sq Mean Sq F value Pr(>F)
## drug                       2   1882     941    5.21 0.0132 *
##   drug: C1                 1   1638    1638    9.07 0.0060 **
##   drug: C2                 1    244     244    1.35 0.2565
## biofeedback                1   1904    1904   10.54 0.0034 **
## drug:biofeedback           2   1248     624    3.46 0.0480 *
##   drug:biofeedback: C1     1   1110    1110    6.15 0.0206 *
##   drug:biofeedback: C2     1    138     138    0.76 0.3907
## Residuals                 24   4335     181
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

compares drugs x & y ignoring z

N.B. contrast for biofeedback is c(-1,1) because it has only 2 levels

contrast is significant
MARGINAL x, y means differ

contrast x biofeedback
interaction is significant

Analyzing a Significant Main Effect

testing sub-effects with linear contrasts

$R^2_{\text{alerting}} = SS(\psi) / SS(\text{Drug}) = 1638/1882= 0.87$

$R^2_{\text{effect size}} = SS(\psi) / SS(\text{Total}) = 1638/9369 = 0.17$

```
summary(mw.aov.02,split=list(drug=list(1,2)))

##                              Df Sum Sq Mean Sq F value Pr(>F)
## drug                          2   1882     941    5.21 0.0132 *
##   drug: C1                    1   1638    1638    9.07 0.0060 **
##   drug: C2                    1    244     244    1.35 0.2565
## biofeedback                   1   1904    1904   10.54 0.0034 **
## drug:biofeedback              2   1248     624    3.46 0.0480 *
##   drug:biofeedback: C1        1   1110    1110    6.15 0.0206 *
##   drug:biofeedback: C2        1    138     138    0.76 0.3907
## Residuals                    24   4335     181
```

$R^2_{\text{alerting}} = SS(\psi) / SS(\text{Drug x Bio}) = 1110/1248 = 0.89$

$R^2_{\text{effect size}} = SS(\psi) / SS(\text{Total}) = 1110/9369 = 0.12$

decomposing an interaction

simple main effects

Example (section 7.11 in course notes)

tf: task factor (3 levels)
gf: group factor (3 levels)

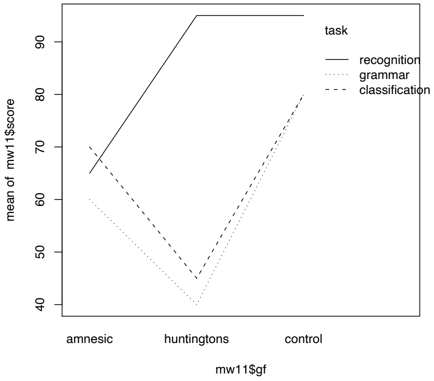
```
> mw11.lm.01 <- lm(score ~ 1+ gf + tf + gf:tf, data = mw11)
> anova(mw11.lm.01)

 Analysis of Variance Table

 Response: score
           Df Sum Sq Mean Sq F value    Pr(>F)
 gf         2   5250 2625.00 16.6373  7.64e-06 ***
 tf         2   5250 2625.00 16.6373  7.64e-06 ***
 gf:tf      4   5000 1250.00  7.9225 0.0001092 ***
 Residuals 36   5680  157.78
 ---
 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

 > MS.w <- 157.8
 > df.w <- 36
```

Example Table 7.11 in Textbook

```
interaction.plot(x.factor=mw11$gf,
                 trace.factor=mw11$tf,
                 response=mw11$score,
                 trace.label="task")
```

illustration of group x task interaction

analyze simple main effect of task in each group



Simple Main Effect of Task

1)

Analyze task <u>within each group</u>:

```
> lm.task.hunt <- lm(score ~ 1 + tf, data = subset(mw11, gf == "huntingtons"))
> lm.task.amnesic <- lm(score ~ 1 + tf, data = subset(mw11, gf == "amnesic"))
> lm.task.control <- lm(score ~ 1 + tf, data = subset(mw11, gf == "control"))
```

2)

<u>Print ANOVA table:</u>

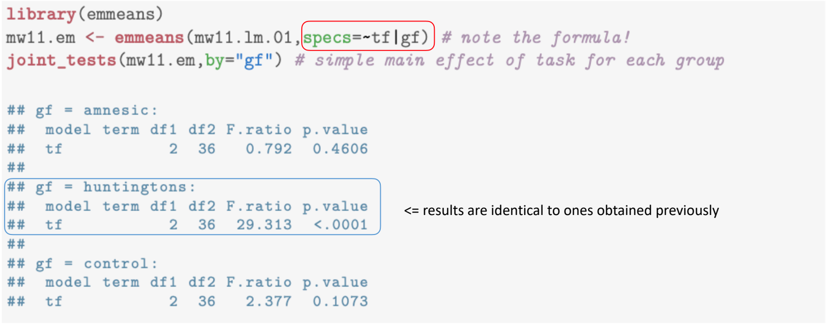
```
> anova(lm.task.hunt)

Analysis of Variance Table

Response: score
          Df Sum Sq Mean Sq F value    Pr(>F)
tf         2   9250  4625.0  29.365 2.385e-05 ***
Residuals 12   1890   157.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3)

<u>Recalculate F & P values:</u>

```
> (F.task.hunt <- 4625/MS.w)

[1] 29.30925

> (p.task.hunt <- 1 - pf(F.task.hunt, df1 = 2, df2 = df.w))

[1] 2.791744e-08
```

Simple Main Effect of Task (emmeans)

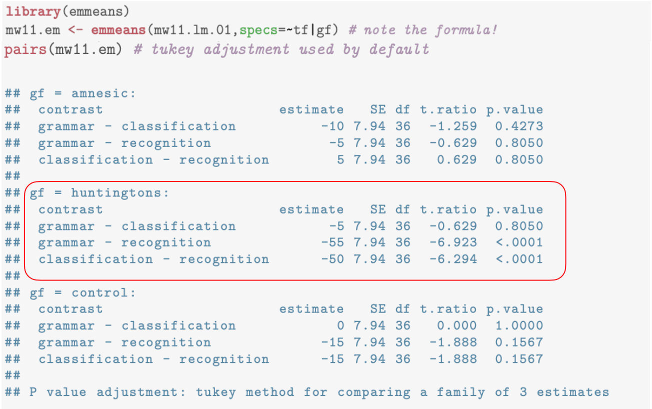


```
library(emmeans)
mw11.em <- emmeans(mw11.lm.01,specs=~tf|gf) # note the formula!
joint_tests(mw11.em,by="gf") # simple main effect of task for each group

## gf = amnesic:
##  model term df1 df2 F.ratio p.value
##  tf           2  36   0.792  0.4606
##
## gf = huntingtons:
##  model term df1 df2 F.ratio p.value
##  tf           2  36  29.313  <.0001
##
## gf = control:
##  model term df1 df2 F.ratio p.value
##  tf           2  36   2.377  0.1073
```

<= results are identical to ones obtained previously

mw11.em <- emmeans(mw11.lm.01, **specs="tf", by="gf"**) # different way of specifying emmeans object

Tukey HSD applied to Simple Main Effect (emmeans)



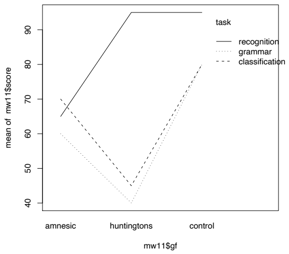
```
library(emmeans)
mw11.em <- emmeans(mw11.lm.01,specs=~tf|gf) # note the formula!
pairs(mw11.em) # tukey adjustment used by default

## gf = amnesic:
##  contrast                      estimate   SE df t.ratio p.value
##  grammar - classification           -10 7.94 36  -1.259  0.4273
##  grammar - recognition               -5 7.94 36  -0.629  0.8050
##  classification - recognition         5 7.94 36   0.629  0.8050
##
## gf = huntingtons:
##  contrast                      estimate   SE df t.ratio p.value
##  grammar - classification            -5 7.94 36  -0.629  0.8050
##  grammar - recognition              -55 7.94 36  -6.923  <.0001
##  classification - recognition       -50 7.94 36  -6.294  <.0001
##
## gf = control:
##  contrast                      estimate   SE df t.ratio p.value
##  grammar - classification             0 7.94 36   0.000  1.0000
##  grammar - recognition              -15 7.94 36  -1.888  0.1567
##  classification - recognition       -15 7.94 36  -1.888  0.1567
##
## P value adjustment: tukey method for comparing a family of 3 estimates
```

decomposing an interaction

interaction contrasts

evaluating *specific* interactions

- often an experimenter expects an interaction
  - a particular sub-effect on one factor to varies across levels of other factor
- Example: predictions for task x group interaction
  - Task sub-effect: difference between implicit & explicit memory tasks
    - task contrast = (0.5*grammar + 0.5*classification) - (1*recognition) ≠ 0
  - prediction: task contrast differs between Amnesic & Huntington's patients
    - Amnesic patients have deficits on explicit, not implicit, memory tasks
    - Huntington's patients have deficits on implicit, not explicit, memory tasks
    - Task $\text{Contrast}_{\text{Amnesic}}$ - Task $\text{Contrast}_{\text{Hunt}}$ ≠ 0



---

creating interaction contrast weights

derive weights to determine if implicit/explicit contrast differs between Amnesic & Huntington's groups

weights for task contrast

		Grammar	Classification	Recognition
		0.5	0.5	-1
Amnesic	1			
Huntington's	-1			
Control	0			

weights for group contrast

---

creating interaction contrast weights

derive weights to determine if implicit/explicit contrast differs between Amnesic & Huntington's groups

		Grammar	Classification	Recognition
		0.5	0.5	-1
Amnesic	1	0.5	0.5	-1
Huntington's	-1	-0.5	-0.5	1
Control	0	0	0	0

cell weight = row x column

---

interaction contrast

```
contrasts(mw11$tf) <- cbind(c(0.5,0.5,-1))
contrasts(mw11$gf) <- cbind(c(1,-1,0))
mw11.aov.02 <- aov(score ~ 1 + tf*gf,data=mw11)
summary(mw11.aov.02,split=list(tf=list(task=1),gf=list(group=1)))

##                     Df Sum Sq Mean Sq F value  Pr(>F)
## tf                   2   5250    2625   16.64 7.6e-06 ***
##   tf: task           1   5062    5062   32.09 2.0e-06 ***
## gf                   2   5250    2625   16.64 7.6e-06 ***
##   gf: group          1    187     187    1.19 0.28290
## tf:gf                4   5000    1250    7.92 0.00011 ***
##   tf:gf: task.group  1   4594    4594   29.12 4.5e-06 ***
## Residuals           36   5680     158
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(R2.alerting <- 4594/5000) # strength of association for our contrast

## [1] 0.9188
```

$R^2_{\text{effect}} = SS_{\text{contrast}}/SS_{\text{total}} = 4594/(5250+5250+5000+5680) = 0.217$

$R^2_{\text{alerting}} = SS_{\text{contrast}}/SS_{\text{interaction}} = 4594/(5000) = \underline{\mathbf{0.919}}$ ← almost all of the interaction is accounted for by our contrast

alternative method

apply contrast to 1-way layout of our 2x3 design

---

creating interaction contrast weights

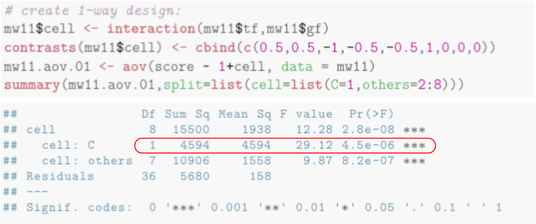
derive weights to determine if implicit/explicit contrast differs between Amnesic & Huntington’s groups

Cell	task	group	task x group
Amnesic-Grammar	0.5	1	0.5
Amnesic-Classification	0.5	1	0.5
Amnesic-Recog	-1	1	-1
Huntington’s - Grammar	0.5	-1	-0.5
Huntington’s - Classification	0.5	-1	-0.5
Huntington’s - Recog	-1	-1	1
Control - Grammar	0.5	0	0
Control - Classification	0.5	0	0
Control - Recog	-1	0	0

---

interaction contrast

(convert 2-way factorial to 1-way design)



```r
# create 1-way design:
mw11$cell <- interaction(mw11$tf,mw11$gf)
contrasts(mw11$cell) <- cbind(c(0.5,0.5,-1,-0.5,-0.5,1,0,0,0))
mw11.aov.01 <- aov(score ~ 1+cell, data = mw11)
summary(mw11.aov.01,split=list(cell=list(C=1,others=2:8)))
```

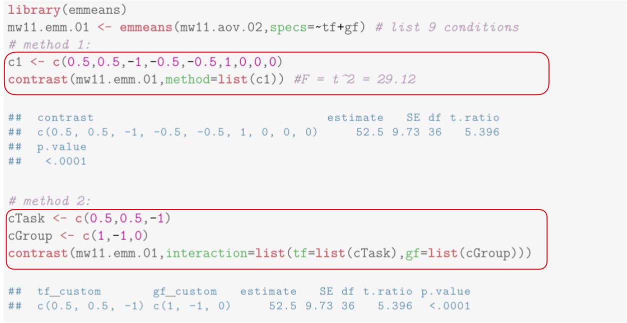
```
##                 Df Sum Sq Mean Sq F value  Pr(>F)
## cell             8  15500    1938   12.28 2.8e-08 ***
##   cell: C        1   4594    4594   29.12 4.5e-06 ***
##   cell: others   7  10906    1558    9.87 8.2e-07 ***
## Residuals       36   5680     158
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
          levels.mw11.cell.    w
1         grammar.amnesic  0.5
2  classification.amnesic  0.5
3     recognition.amnesic -1.0
4      grammar.huntingtons -0.5
5 classification.huntingtons -0.5
6  recognition.huntingtons  1.0
7         grammar.control  0.0
8  classification.control  0.0
9     recognition.control  0.0
```

$R^2_{effect} = SS_{contrast}/SS_{total} = 4594/(15500+5680) = 0.217$

---

interaction contrast with emmeans



```r
library(emmeans)
mw11.emm.01 <- emmeans(mw11.aov.02,specs=~tf+gf) # list 9 conditions
# method 1:
c1 <- c(0.5,0.5,-1,-0.5,-0.5,1,0,0,0)
contrast(mw11.emm.01,method=list(c1)) #F = t^2 = 29.12
```

```
##  contrast                                  estimate   SE df t.ratio
##  c(0.5, 0.5, -1, -0.5, -0.5, 1, 0, 0, 0)       52.5 9.73 36   5.396
##  p.value
##   <.0001
```

$F = t^2 = 5.396^2 = 29.12$

```r
# method 2:
cTask <- c(0.5,0.5,-1)
cGroup <- c(1,-1,0)
contrast(mw11.emm.01,interaction=list(tf=list(cTask),gf=list(cGroup)))
```

```
##  tf_custom        gf_custom    estimate   SE df t.ratio p.value
##  c(0.5, 0.5, -1) c(1, -1, 0)      52.5 9.73 36   5.396  <.0001
```

$F = t^2 = 5.396^2 = 29.12$