Analysis of Variance (Anova)

Vera Demberg

Saarland University

WS 2021/2022

What we covered so far

We have now seen how to relate two variables when they are of different types:

- **response:** categorial
  predictor: categorial
  $\rightarrow$ binomial test / χ^2 / McNemar / Fisher's exact test
- **response:** ratio or interval
  predictor: categorial
  $\rightarrow$ t-test / Welch test / Wilcoxon test
- **response:** ratio or interval
  predictor: ratio or interval
  $\rightarrow$ correlation r / ρ / τ & regression

What we covered so far

We have now seen how to relate two variables when they are of different types:

- **response:** categorial
  predictor: categorial
  $\rightarrow$ binomial test / χ^2 / McNemar / Fisher's exact test
- **response:** ratio or interval
  predictor: categorial (BINARY!)
  $\rightarrow$ t-test / Welch test / Wilcoxon test
- **response:** ratio or interval
  predictor: ratio or interval
  $\rightarrow$ correlation r / ρ / τ & regression

What we covered so far

We have now seen how to relate two variables when they are of different types:

- **response:** categorial
  predictor: categorial
  $\rightarrow$ binomial test / χ^2 / McNemar / Fisher's exact test
- **response:** ratio or interval
  predictor: categorial (BINARY!)
  $\rightarrow$ t-test / Welch test / Wilcoxon test
- **response:** ratio or interval
  predictor: ratio or interval
  $\rightarrow$ correlation r / ρ / τ & regression

Today: what if we have more than two levels in a categorial variable, or more than one predictor variable?

More than two groups and variables

- Using the t-test, we compared the means of two groups
- for example, we compared the grades of students when those students were taught by tutor Anastasia vs. tutor Bernadette.
- i.e., there was a factor "Tutor" with levels "Anastasia" and "Bernadette"

More than two groups and variables

- Using the t-test, we compared the means of two groups
- for example, we compared the grades of students when those students were taught by tutor Anastasia vs. tutor Bernadette.
- i.e., there was a factor "Tutor" with levels "Anastasia" and "Bernadette"

Today, we look at methods for dealing with

- comparison of more than two means (i.e., what to do if we had a third tutor, Florentina)
- or comparisons with more than one predictor variable, for instance, comparing not only what effect a tutor had on performance, but also what effect the main study subject had on performance.

More than two - cntd

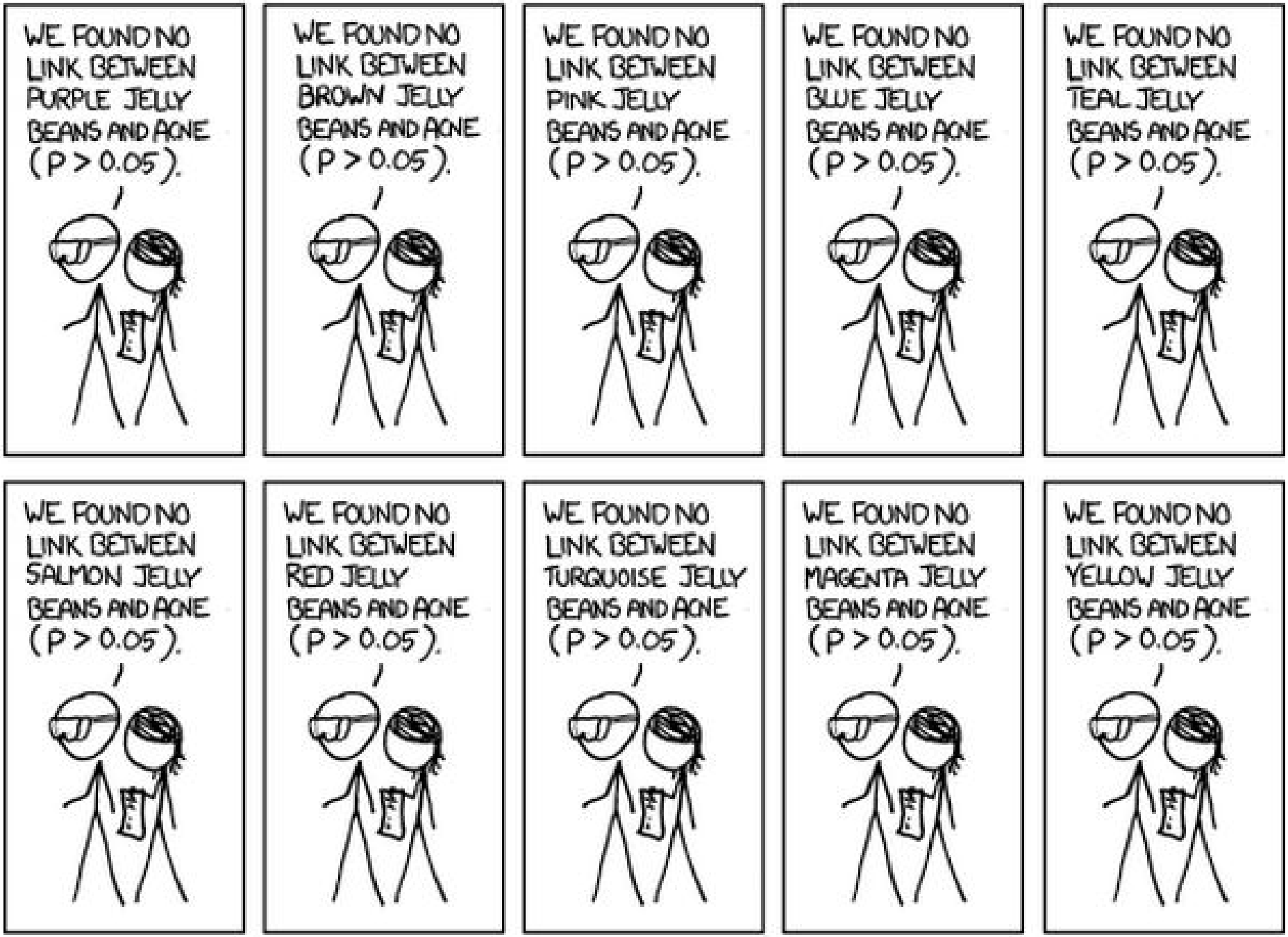
Let's first consider the case of three tutors.

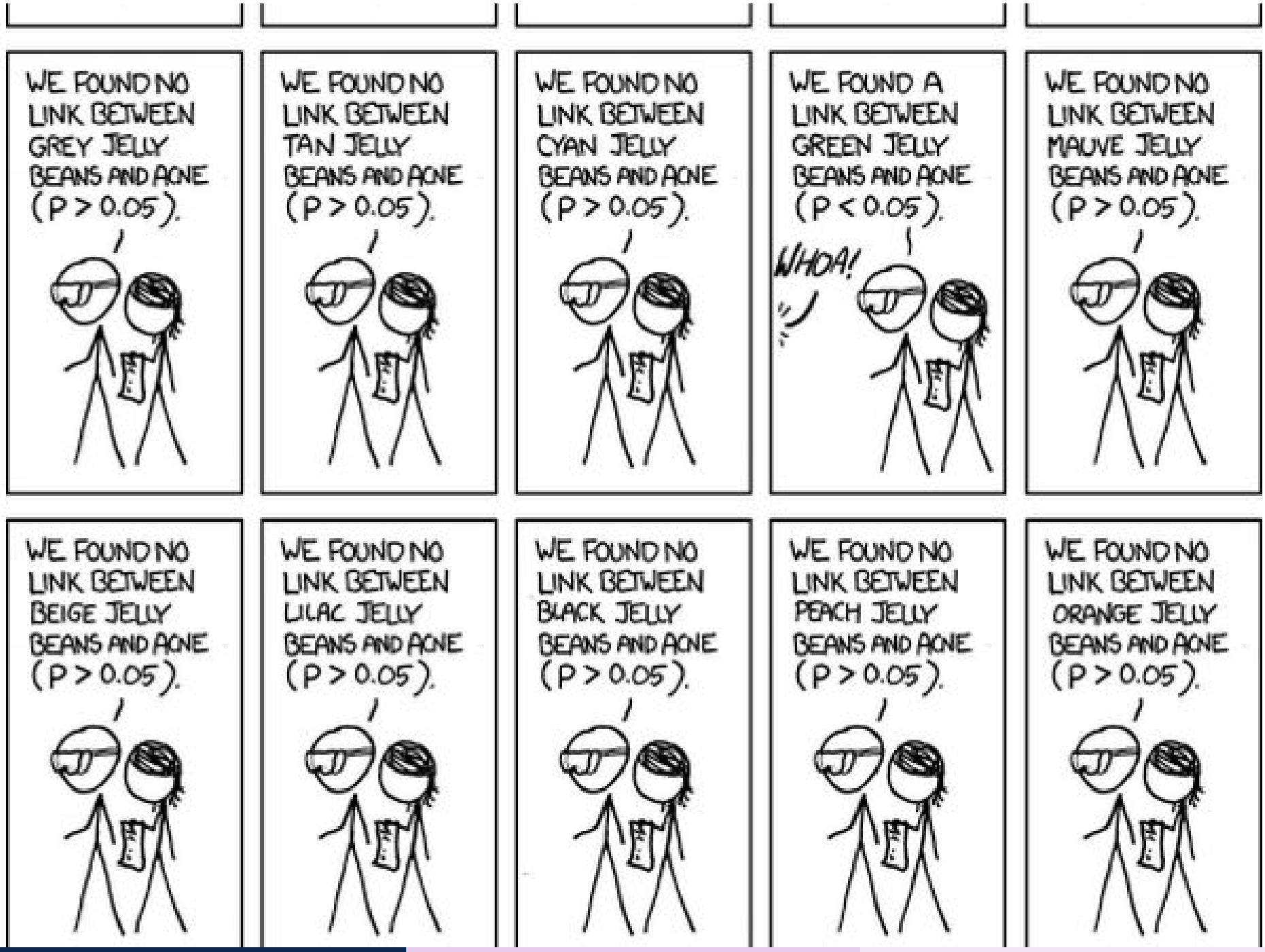
- simplest idea: run multiple t tests (Bernadette vs. Anastasia, Anastasia vs. Florentina and Bernadette vs. Florentina)
  $N * (N - 1)/2$ t tests where N is the number of levels of a factor

More than two - cntd

Let's first consider the case of three tutors.

- simplest idea: run multiple t tests (Bernadette vs. Anastasia, Anastasia vs. Florentina and Bernadette vs. Florentina)
  $N * (N - 1)/2$ t tests where N is the number of levels of a factor
- **crucially**, running multiple t tests is not a good idea because the Type I error rate (α) increases very quickly when we carry out multiple tests on the same sample of data


WE FOUND NO LINK BETWEEN PURPLE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BROWN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PINK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLUE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TEAL JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN SALMON JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN RED JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TURQUOISE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN MAGENTA JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN YELLOW JELLY BEANS AND ACNE (P > 0.05).


WE FOUND NO LINK BETWEEN GREY JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TAN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN CYAN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND A LINK BETWEEN GREEN JELLY BEANS AND ACNE (P < 0.05).
WHOA!
WE FOUND NO LINK BETWEEN MAUVE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BEIGE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN LILAC JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLACK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PEACH JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN ORANGE JELLY BEANS AND ACNE (P > 0.05).

ANOVA = Analysis of Variance

- One-way ANOVA: one factor (I.V.), with more than two levels
- Two-way ANOVA: two factors (I.V.'s)
  - when each of the two factors has only two levels, then we have a 2X2 design
  - very common

ANOVA = Analysis of Variance

- One-way ANOVA: one factor (I.V.), with more than two levels
- Two-way ANOVA: two factors (I.V.'s)
  - when each of the two factors has only two levels, then we have a 2X2 design
  - very common

Table of contents:

Table of Contents

Effect of cognate status on lexical decision times

Example

Participants

German native speakers with Dutch as L2 and English as L3

Cognate status	non-cognate	L2 cognate	L3 (and L2) cognate
	Regenschirm	Stuhl	Finger

Question

Does the time it takes to decide whether a stimulus is a word differ depending on whether it is phonologically related to a word with the same meaning (a cognate) in the participant's second or third language?

One-way unrelated ANOVA - Example

Some fake data. Say we tested 10 words in each group and obtained the following numbers (let's assume for now we don't have any repeated measures, i.e. each data points is an average across participants).

item	non-cognate	L2 cognate	L3 cognate
1	580	540	560
2	600	560	580
3	560	520	540
4	540	500	520
5	585	545	565
6	575	535	555
7	510	470	490
8	610	570	590
9	605	575	595
10	545	505	525
mean	**571**	**532**	**552**

One-way unrelated ANOVA - Partitioning the variance

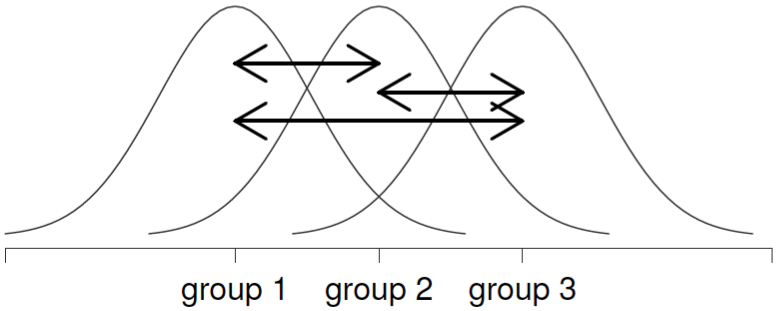
We need to compute two variances (i.e., two sources of variation):

1. the variation explained by the factor of interest (SS_B)[1]
2. the variation explained by sampling error (SS_E) aka SS_W
3. which of course together make up the total variation in the D.V. (SS_{TOT})

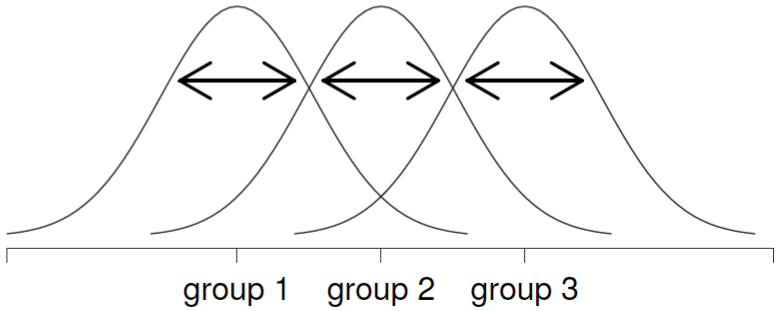
[1] SS_B stands for between groups sum of squares; SS_W stands for within group sum of squares

between group vs. within group variance

Between–group variation
(i.e., differences among group means)



Within–group variation
(i.e., deviations from group means)



One-way unrelated ANOVA

SS_{TOT}

$$SS_{TOT} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

in our example: n = 30 (the total number of scores in all groups) and $\bar{Y}$ is the grand mean (the mean of all scores, ignoring group membership); here $\bar{Y} = 552$

- NOTE: this is not the variance, because it is not divided by $n - 1$; in ANOVA, the variance, $\frac{SS_{TOT}}{n-1}$, is called mean squared error, or MS_{TOT}

One-way unrelated ANOVA

Variation within a group:

$$SS_W = \sum_{k=1}^{G} \sum_{i=1}^{N_k} (Y_{ik} - \bar{Y}_k)^2$$

in our example:
G = 3 (the number of groups, or levels of the factor),
$N_k = 10$ (the number of data points in a group),
$\bar{Y}_k$ is the mean of the kth level or group

One-way unrelated ANOVA

Variation explained by predictor variable, our factor: SS_B, or "between" sum of squares

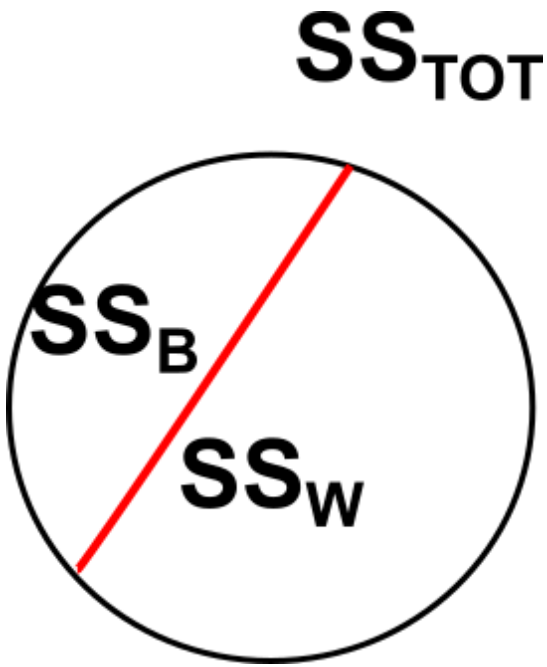
$$SS_B = \sum_{k=1}^{G} \sum_{i=1}^{N_k} (\bar{Y}_k - \bar{Y})^2$$
$$= \sum_{k=1}^{G} N_k (\bar{Y}_k - \bar{Y})^2$$

in our example:
G = 3 (the number of groups, or levels of the factor),
$\bar{Y}_k$ is the mean of the ith level or group
and $\bar{Y}$ is the overall mean

One-way unrelated ANOVA

$$SS_{TOT} = SS_W + SS_B$$

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{k=1}^{G}\sum_{i=1}^{N_k}(Y_{ik} - \bar{Y}_k)^2 + \sum_{k=1}^{G} N_k(\bar{Y}_k - \bar{Y})^2$$



One-way unrelated ANOVA - Step 4

We are now interested in quantifying how much of the variance is explained by our predictor variable (SS_B), vs. how much is just unexplained variation (SS_W).
The *F ratio* does this: it divides the variance between groups by the variance within groups:

$$F = \frac{MS_B}{MS_W} = \frac{\frac{SS_B}{k-1}}{\frac{SS_W}{n-k}}$$

Intuitively, this number increases with the portion of variation explained by the factor compared to the error variation.

One-way unrelated ANOVA - Step 4

We are now interested in quantifying how much of the variance is explained by our predictor variable (SS_B), vs. how much is just unexplained variation (SS_W).
The *F ratio* does this: it divides the variance between groups by the variance within groups:

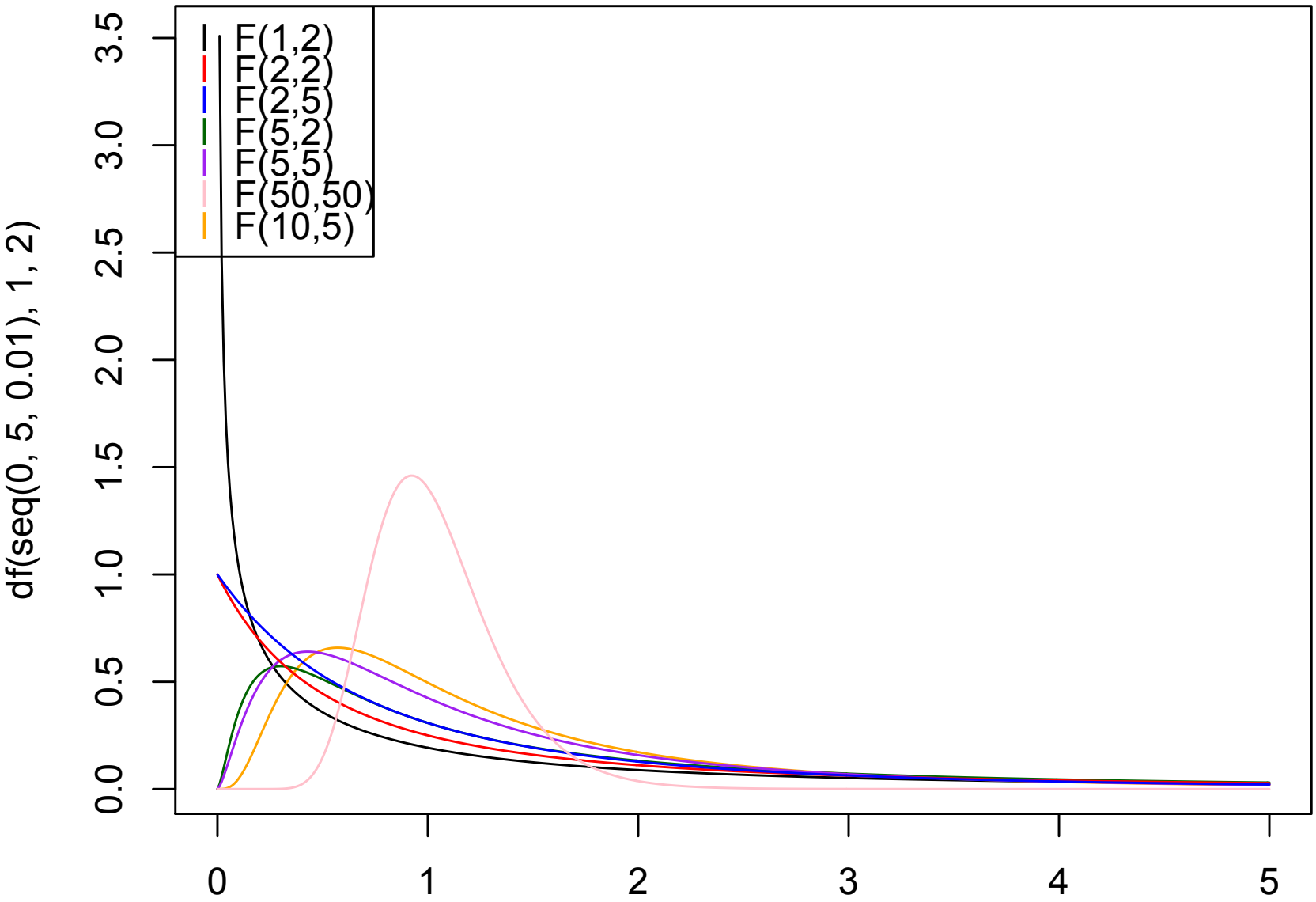
$$F = \frac{MS_B}{MS_W} = \frac{\frac{SS_B}{k-1}}{\frac{SS_W}{n-k}}$$

Intuitively, this number increases with the portion of variation explained by the factor compared to the error variation.

Oh boy – yet another distribution: F?

Table of Contents

The F distribution

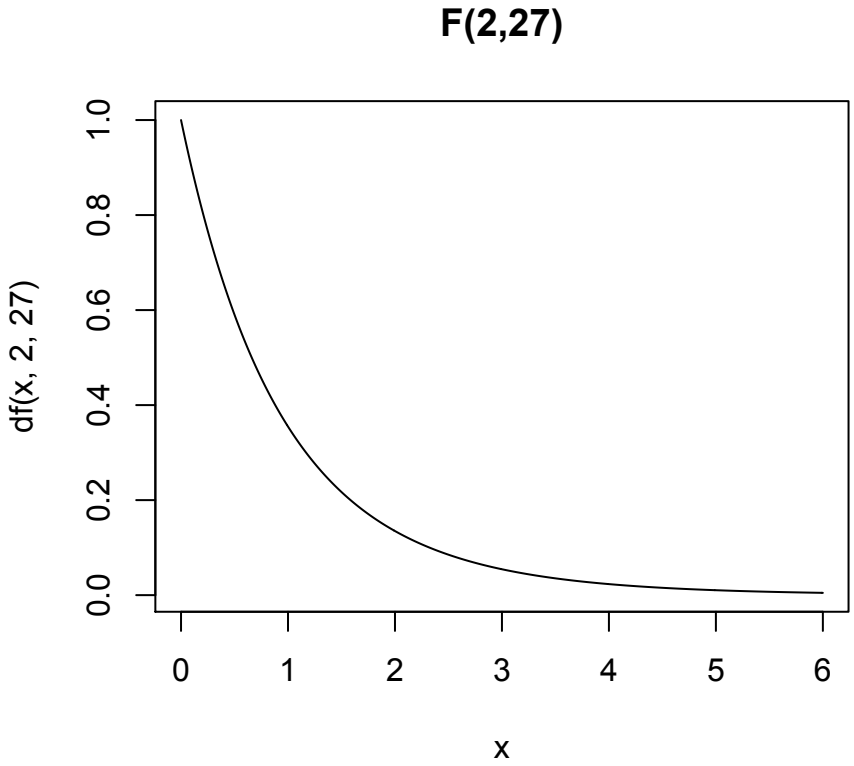


So, what about cognates?

How to report ANOVA results

F(2, 27) = 3.49, p <.05

The degrees of freedom are (k-1) and (n-k).




```
> pf(3.49, 2, 27, lower.tail=FALSE)
[1] 0.04486372
```

Doing it in R – the easy way

```
> summary(mdat)
  cognate          RT
 cog  :10   Min.   :470.0
 l2cog:10   1st Qu.:527.5
 l3cog:10   Median :557.5
            Mean   :551.7
            3rd Qu.:578.8
            Max.   :610.0
> summary(aov(mdat$RT~mdat$cognate))
             Df Sum Sq Mean Sq F value Pr(>F)
mdat$cognate  2   7607    3803   3.492 0.0448 *
Residuals    27  29410    1089
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

A note of caution on interpreting one-way ANOVA

- One-way ANOVA has a huge limitation
- it tells you whether the factor of interest has an effect on the D.V.
- put differently, it tells you whether the levels of the factor differ from one another, or not
- BUT it does not tell you where the difference lies!!
  - e.g., are L2-cognates processed differently from L3-cognates as well as from non-cognates?

Effect size

Again, we'd like to estimate also *effect size*. The most common thing for Anovas is to calculate η^2, which describes what proportion of the variance is explained by the factor.

$$\eta^2 = \frac{SS_b}{SS_{tot}} = \frac{7606.667}{7606.667 + 29410} = 0.205493$$

```
> etaSquared(aov(mdat$RT~mdat$cognate), anova=TRUE)
                eta.sq eta.sq.part        SS df       MS       F
mdat$cognate 0.205493    0.205493  7606.667  2 3803.333 3.49167
Residuals    0.794507          NA 29410.000 27 1089.259      NA
                     p
mdat$cognate 0.04480425
Residuals            NA
```

ANOVA vs. t-test

So what's the relation between an ANOVA and a t-test?

ANOVA vs. t-test

So what's the relation between an ANOVA and a *t*-test?

Actually, an ANOVA with a single binary predictor variable does exactly the same thing as a t test.

ANOVA vs. t-test

So what's the relation between an ANOVA and a *t*-test?

Actually, an ANOVA with a single binary predictor variable does exactly the same thing as a t test.

Let's take a look at an example by changing our data to only include cognate and L2 cognate.

```
> summary(aov(ddat$RT~ddat$cognate))
             Df Sum Sq Mean Sq F value Pr(>F)
ddat$cognate  1   7605    7605   7.074  0.016 *
Residuals    18  19350    1075
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> t.test(ddat$RT~ddat$cognate, var.equal=TRUE)

    Two Sample t-test

data:  ddat$RT by ddat$cognate
t = 2.6598, df = 18, p-value = 0.01596
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  8.194436 69.805564
sample estimates:
  mean in group cog mean in group l2cog
                571                 532


> 2.6598^2
[1] 7.074536
```

Post-hoc comparisons and planned contrasts

For the data with three levels of the factor `cognate`, the ANOVA told us that there is a significant difference between the cognate types. But between which ones?

- to answer such questions, you need to carry out additional tests
- if you thought about testing for differences between specific conditions **before** carrying out your ANOVA, and you have specific expectations about where the differences will be, then you need to specify **planned contrasts**
- if you did not think about this beforehand, or have no specific expectation, then you should go for **post-hoc tests**
- we will say more about planned comparisons when we talk about linear regression

Pairwise post-hoc tests in R

The function `pairwise.t.test` automatically runs $t-$tests for all levels.
First example: no correction!

```
> pairwise.t.test(mdat$RT, mdat$cognate, p.adjust.method="none")

    Pairwise comparisons using t tests with pooled SD

data:  mdat$RT and mdat$cognate

      cog   l2cog
l2cog 0.014 -
l3cog 0.209 0.187

P value adjustment method: none
```

Post-hoc comparisons

Bonferroni correction for multiple testing

- basic idea (very conservative; i.e., with high Type II error) is to carry out as many t-tests as needed and then divide the α level by the number of tests
  - if you have 3 levels, you need to carry out 3 t-tests, so your α level should be set to $\frac{.05}{3} = .0167$ (or you multiply the p value by 3 before comparing it to your usual α level of .05)
- this is done to avoid inflating the Type I error rate

For alternatives, e.g. Tukey's correction, see also the DataCamp exercise.

t-test with Bonferroni correction

without correction:

```
> pairwise.t.test(mdat$RT, mdat$cognate, p.adjust.method="none")

    Pairwise comparisons using t tests with pooled SD

data:  mdat$RT and mdat$cognate

      cog   l2cog
l2cog 0.014 -
l3cog 0.209 0.187

P value adjustment method: none
```

with correction:

```
> pairwise.t.test(mdat$RT, mdat$cognate, p.adjust.method="bonferroni")

    Pairwise comparisons using t tests with pooled SD

data:  mdat$RT and mdat$cognate

      cog   l2cog
l2cog 0.041 -
l3cog 0.627 0.560

P value adjustment method: bonferroni
```

Assumptions of one-way ANOVA

- Normality: the residuals are assumed to be normally distributed (QQ-plot or Shapiro-Wilk test).
- Homogeneity of variance: population standard deviation is the same for all groups $\rightarrow$ see also DataCamp Tutorial.
- Independence: no repeated measures.

Table of Contents

Two factors

- test for the effect of two factors simultaneously
- similar problem as when testing for the effect of a single factor (one-way ANOVA), except that we now have two, so we need to compute two **main effects**
- additionally, we need to take into account the fact that the two factors might interact

Example

Case study:
imagine we have a clinical trial where we test for the effect of an antidepressant. We compare the effect of three different medications, and also compare what happens with vs. without additional therapy. (3×2 factorial design)
We have 18 patients in total:

```
> xtabs( ~ drug + therapy, clin.trial )
           therapy
drug        no.therapy CBT
  placebo             3   3
  anxifree            3   3
  joyzepam            3   3
```

We have an equal number of people in all cells, so it's a *balanced design*.

We want to see whether any of the medications are helpful, and whether being in therapy makes a difference: compare mood gain for each medication and therapy.

```
> aggregate( mood.gain ~ drug + therapy, clin.trial, mean )
      drug    therapy mood.gain
1  placebo no.therapy  0.300000
2 anxifree no.therapy  0.400000
3 joyzepam no.therapy  1.466667
4  placebo        CBT  0.600000
5 anxifree        CBT  1.033333
6 joyzepam        CBT  1.500000
```

```
> aggregate( mood.gain ~ drug, clin.trial, mean )
      drug mood.gain
1  placebo 0.4500000
2 anxifree 0.7166667
3 joyzepam 1.4833333
```

```
> aggregate( mood.gain ~ therapy, clin.trial, mean )
     therapy mood.gain
1 no.therapy 0.7222222
2        CBT 1.0444444
```

Calculating means for different groupings

	no therapy	CBT	total
placebo	0.30	0.60	0.45
anxifree	0.40	1.03	0.72
joyzepam	1.47	1.50	1.48
total	0.72	1.04	0.88

	no therapy	CBT	total
placebo	μ_{11}	μ_{12}	$\mu_{1.}$
anxifree	μ_{21}	μ_{22}	$\mu_{2.}$
joyzepam	μ_{31}	μ_{32}	$\mu_{3.}$
total	$\mu_{.1}$	$\mu_{.2}$	$\mu_{..}$

Calculating means for different groupings

	no therapy	CBT	total
placebo	μ_{11}	μ_{12}	$\mu_{1.}$
anxifree	μ_{21}	μ_{22}	$\mu_{2.}$
joyzepam	μ_{31}	μ_{32}	$\mu_{3.}$
total	$\mu_{.1}$	$\mu_{.2}$	$\mu_{..}$

Hypotheses:

regarding drugs:

H_0: average changes in mood gain are the same, i.e. $\mu_{1.} = \mu_{2.} = \mu_{3.}$

H_1: the mean mood change for at least one drug is different

regarding therapy:

H_0: average changes in mood gain are the same, i.e. $\mu_{.1} = \mu_{.2}$

H_1: the mean mood change for therapy vs. no therapy is different, i.e. $\mu_{.1} \neq \mu_{.2}$

Running this in R

```
> model.2 <- aov( mood.gain ~ drug + therapy, clin.trial )
> summary( model.2 )
            Df Sum Sq Mean Sq F value  Pr(>F)
drug         2   3.45   1.727   26.15 1.9e-05 ***
therapy      1   0.47   0.467    7.08   0.019 *
Residuals   14   0.92   0.066
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Single Anova with 2 predictors $\neq$ 2 simple Anovas

```
> model.2 <- aov( mood.gain ~ drug + therapy, clin.trial )
> summary( model.2 )
            Df Sum Sq Mean Sq F value  Pr(>F)
drug         2   3.45   1.727   26.15 1.9e-05 ***
therapy      1   0.47   0.467    7.08   0.019 *
Residuals   14   0.92   0.066
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

```
> summary( my.anova )
            Df Sum Sq Mean Sq F value  Pr(>F)
drug         2   3.45   1.727    18.6 8.6e-05 ***
Residuals   15   1.39   0.093
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

```
> summary( aov( mood.gain ~ therapy, data = clin.trial ))
            Df Sum Sq Mean Sq F value Pr(>F)
therapy      1   0.47   0.467    1.71   0.21
Residuals   16   4.38   0.274
```

Why is this different?

Remember that ANOVA compares the amount of variance explained by a factor to the unexplained variance.

If another factor explains a lot of the variance, this therefore affects the outcome of the ANOVA. Therefore, it's important to always include all *relevant* factors!

Why is this different?

Remember that ANOVA compares the amount of variance explained by a factor to the unexplained variance.

If another factor explains a lot of the variance, this therefore affects the outcome of the ANOVA. Therefore, it's important to always include all *relevant* factors!

Remember also our example from the first lecture: Gender balance in Berkeley data lead to completely wrong results if not taking into account the subject studied!

Two factors: interactions

What we just saw was a simple (Type 1) factorial Anova.

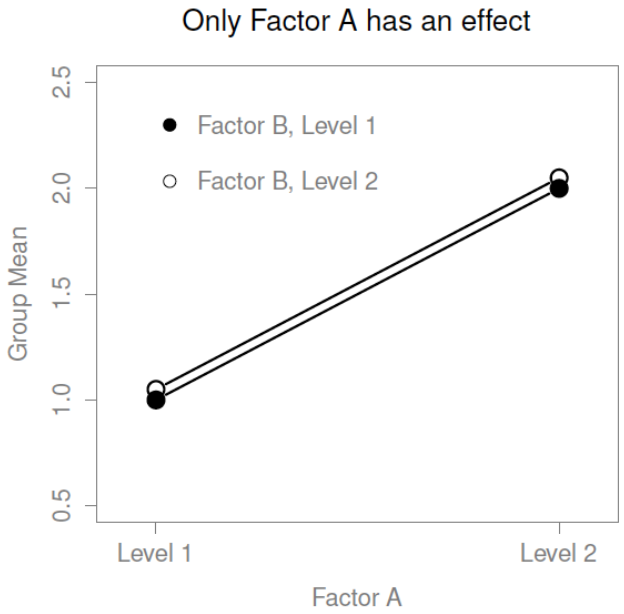
- tested for the effect of two factors simultaneously
- similar to testing for the effect of a single factor (one-way ANOVA), except that we now have two, so we need to compute two **main effects**
- however, we need to take into account the fact that the two factors might interact

Interaction

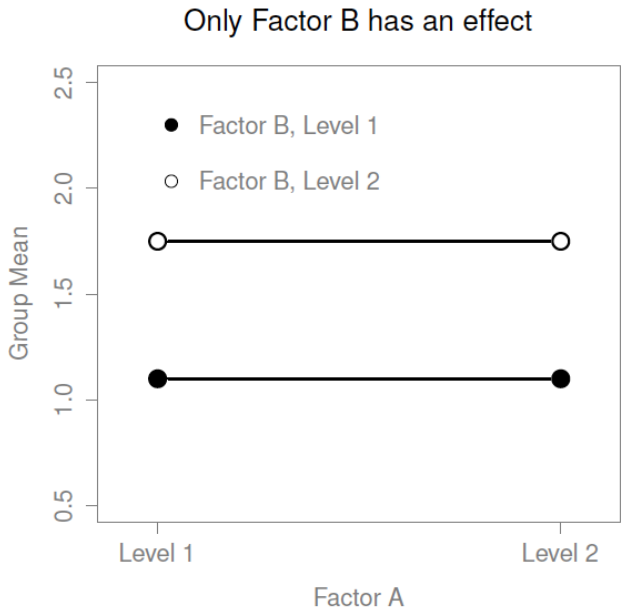
The combined effect of two predictor variables on the outcome variable; in general terms, the presence of a significant interaction means that the effect of one factor is NOT the same across all levels of the other factor.

Example

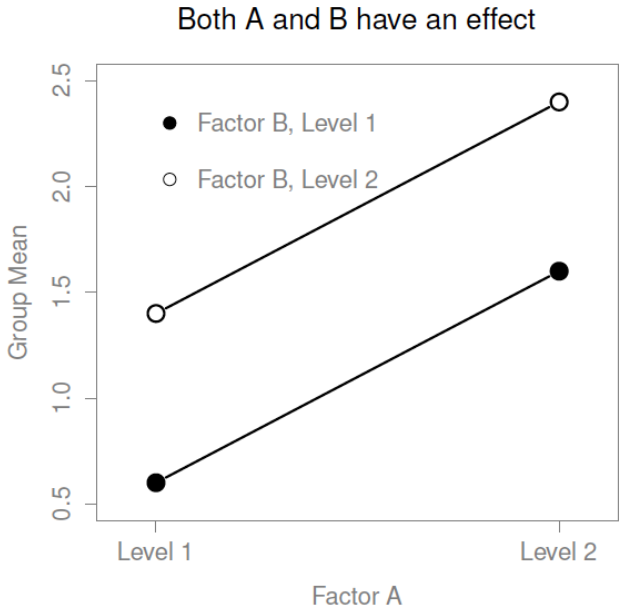
Suppose that the operation of Anxifree and Joyzepam is governed by quite different physiological mechanisms, and one consequence of this is that while Joyzepam has more or less the same effect on mood regardless of whether one is in therapy, Anxifree is actually much more effective when administered in conjunction with therapy.


Only Factor A has an effect
Factor B, Level 1
Factor B, Level 2
Group Mean
0.5
1.0
1.5
2.0
2.5
Level 1
Level 2
Factor A

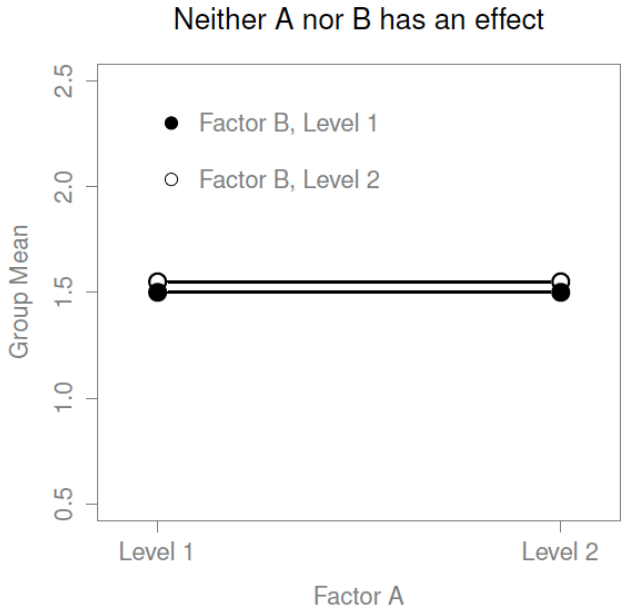
(a)


Only Factor B has an effect
Factor B, Level 1
Factor B, Level 2
Group Mean
0.5
1.0
1.5
2.0
2.5
Level 1
Level 2
Factor A

(b)

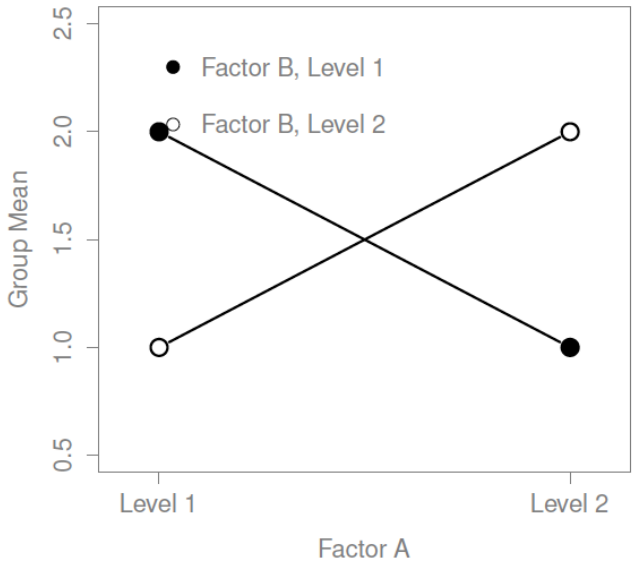

Both A and B have an effect
Factor B, Level 1
Factor B, Level 2
Group Mean
0.5
1.0
1.5
2.0
2.5
Level 1
Level 2
Factor A

(c)


Neither A nor B has an effect
Factor B, Level 1
Factor B, Level 2
Group Mean
0.5
1.0
1.5
2.0
2.5
Level 1
Level 2
Factor A

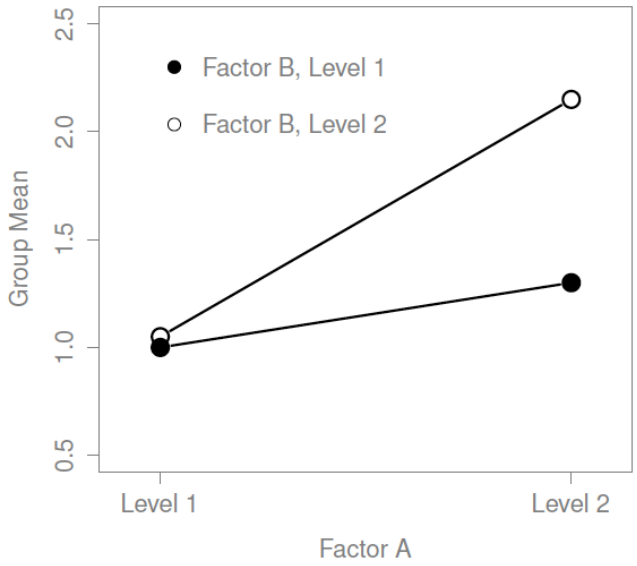
(d)

Crossover interaction



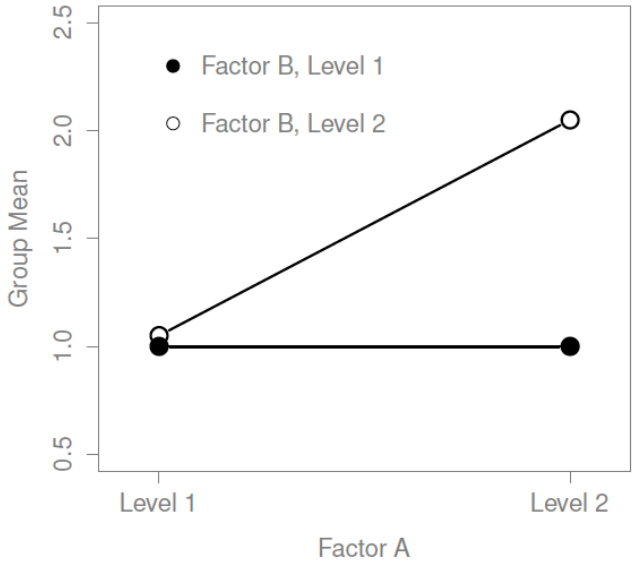
(a)

Effect for one level of Factor A

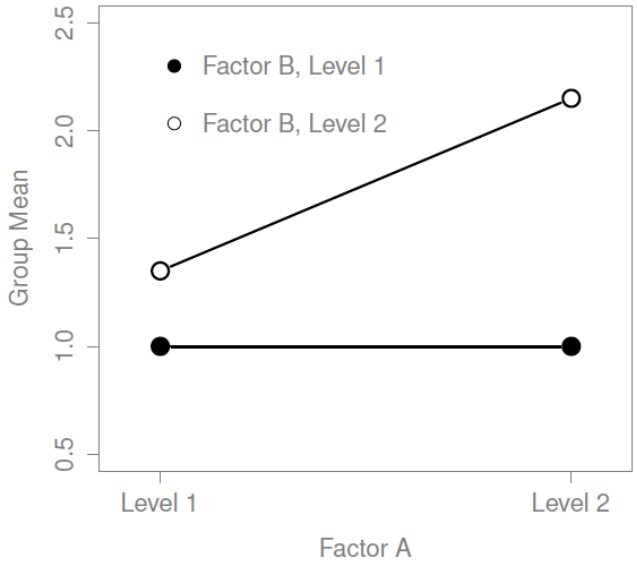


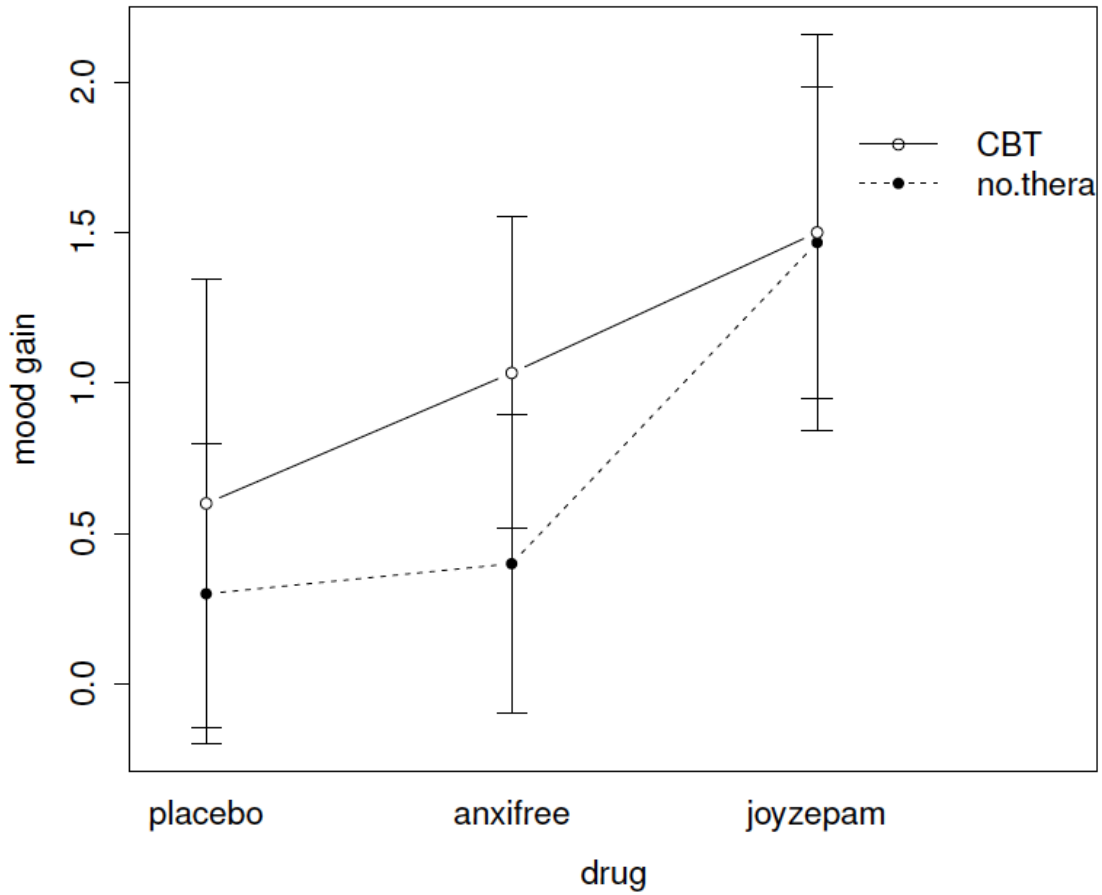
(b)

One cell is different

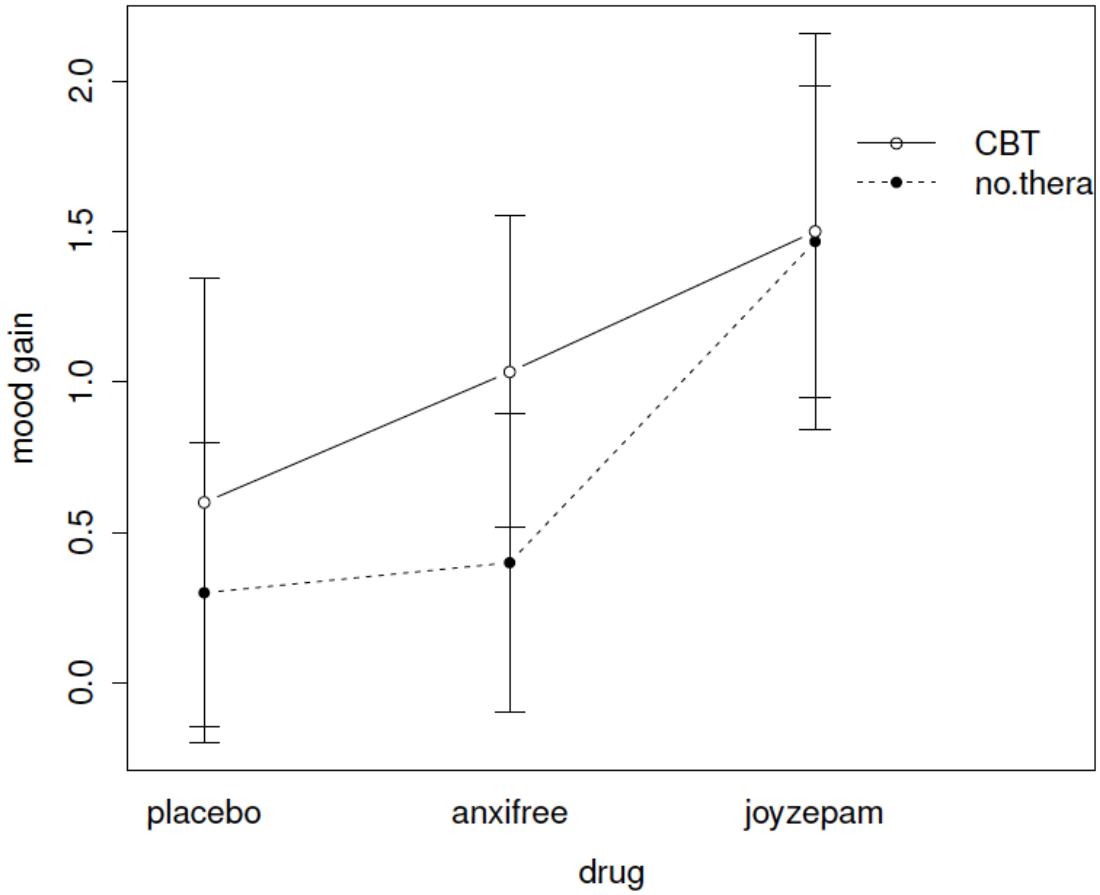


Effect for one level of Factor B

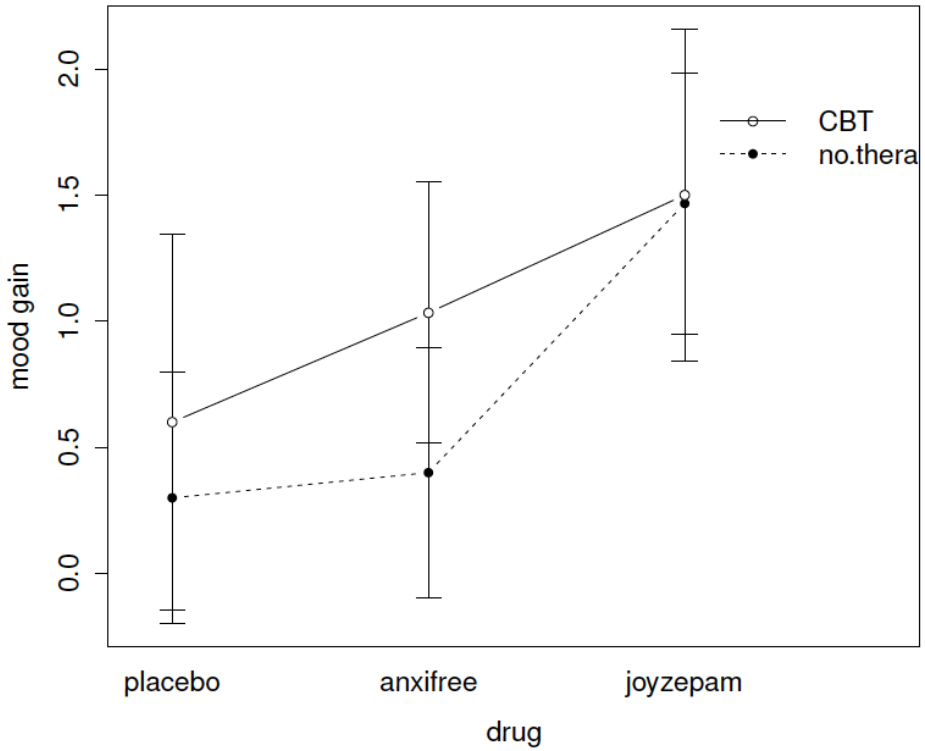




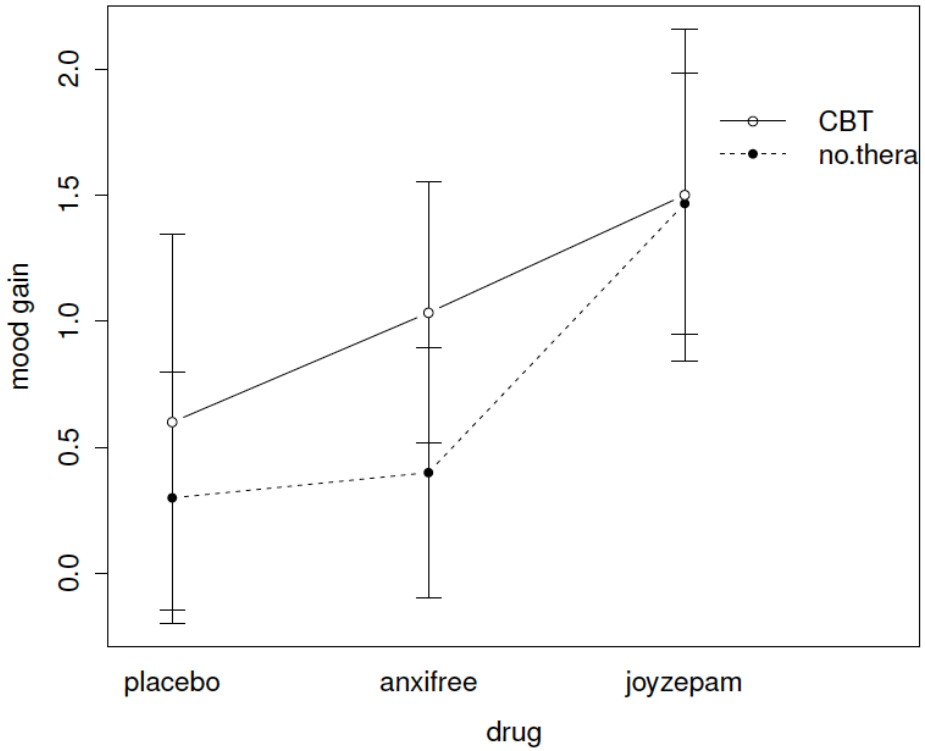
There is no effect of therapy when Joyzepam is administered, but we do find a large beneficial effect of therapy when Anxifree is used.




```
> lineplot.CI( x.factor = clin.trial$drug,
+              response = clin.trial$mood.gain,
+              group = clin.trial$therapy,
+              ci.fun = ciMean,
+              xlab = "drug",
+              ylab = "mood gain" )
```



hm, can we see this interaction effect in our model?



hm, can we see this interaction effect in our model?
(No, we haven't told the model to calculate this)

```
> model.2 <- aov( mood.gain ~ drug + therapy, clin.trial )
> summary( model.2 )
            Df Sum Sq Mean Sq F value  Pr(>F)
drug         2   3.45   1.727   26.15 1.9e-05 ***
therapy      1   0.47   0.467    7.08   0.019 *
Residuals   14   0.92   0.066
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Including an Interaction term in R

instead of a type 1 anova without interactions:

```
aov(mood.gain ~ drug + therapy , data=clin.trial)
```

Including an Interaction term in R

instead of a type 1 anova without interactions:

```
aov(mood.gain ~ drug + therapy , data=clin.trial)
```

we can add an interaction term:

```
aov(mood.gain ~ drug + therapy +drug:therapy ,
data=clin.trial)
```

Including an Interaction term in R

instead of a type 1 anova without interactions:

```
aov(mood.gain ~ drug + therapy , data=clin.trial)
```

we can add an interaction term:

```
aov(mood.gain ~ drug + therapy +drug:therapy ,
data=clin.trial)
```

a shorthand for this in R is:

```
aov(mood.gain ~ drug * therapy , data=clin.trial)
```

Interaction model in R

```
> model.3 <- aov( mood.gain ~ drug * therapy, clin.trial )
> summary( model.3 )
             Df Sum Sq Mean Sq F value  Pr(>F)
drug          2   3.45   1.727   31.71 1.6e-05 ***
therapy       1   0.47   0.467    8.58   0.013 *
drug:therapy  2   0.27   0.136    2.49   0.125
Residuals    12   0.65   0.054
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Interaction model in R

```
> model.3 <- aov( mood.gain ~ drug * therapy, clin.trial )
> summary( model.3 )
             Df Sum Sq Mean Sq F value  Pr(>F)
drug          2   3.45   1.727   31.71 1.6e-05 ***
therapy       1   0.47   0.467    8.58   0.013 *
drug:therapy  2   0.27   0.136    2.49   0.125
Residuals    12   0.65   0.054
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

We find a significant main effect of drug ($F(2, 12) = 31.7, p < .001$) and therapy type ($F(1, 12) = 8.6, p = .013$), but there is no significant interaction between the two ($F(2, 12) = 2.5, p = 0.125$).

Two-way unrelated ANOVA - Interaction

- the variation explained by the interaction is the residual model variation (i.e., the residual non-error variation)
- we compute it by subtracting the sum of squares associated with the two main effects from the total sum of squares explained by the model

NOTE: For now, we will assume that the design is balanced: i.e., the number of scores in each cell is the same

Two-way unrelated ANOVA - Interaction

- the total variation explained by the model is computed as:

$$SS_M = \sum_{i=1}^{t} N_i * (Y_i - \bar{Y})^2$$

- where t is the number of cells in the design (for a 3X2 design, this is 6)
- and N_i is the number of scores in each cell

Two-way unrelated ANOVA - Interaction

- the variation explained by the interaction is then computed as:

$$SS_{A:B} = SS_M - SS_A - SS_B$$

- the degrees of freedom associated with the interaction sum of squares are

$$\begin{aligned} df_{A:B} &= (R \times C - 1) - (R - 1) - (C - 1) \\ &= RC - R - C + 1 \\ &= (R - 1)(C - 1) \end{aligned}$$

- residual degrees of freedom:

$$\begin{aligned} df_{resid} &= N - 1 - (R - 1) - (C - 1) - (R - 1) \times (C - 1) \\ &= N - R - C + 1 - (R \times C - R - C + 1) \\ &= N - (R \times C) \end{aligned}$$

Summary

- ANOVA = Analysis of variance: compare explained to unexplained variance
- F distribution: two parameters for degrees of freedom of chisq distributions which model the sum of squares.
- simple ANOVA: equivalent to t-test if done for factor with 2 levels
- simple ANOVA allows to test for factors with more than 2 levels
- post-hoc t-tests to test for paired effects: need to adjust for multiple testing
- factorial ANOVA allows to account for more than a single effect: two effects at the same time, plus also their interaction
- Calculations differ slightly for whether design is balanced or not $\rightarrow$ Navarro chapter 16