# Statistics
## Repeated-measures ANOVA worksheet

**Maarten Speekenbrink** *UCL Experimental Psychology*
**November, 2016**

## 1 Spss Exercises

1. A researcher wants to assess the effect of certain characteristics on people's ability to solve complex puzzles. The 12 participants in the study are each given four puzzles, differing in shape: round (R) or square (S), and use of colour: black and white (BW) or in colour (C). The dependent variable is the time taken to solve each puzzle, measured in minutes. The hypothesis is that round puzzles are more difficult than square puzzles, and that black-and-white puzzles are harder than colour pictures. You can find the (fictitious) data on Moodle as `puzzles.sav`.

   (a) Define a composite variable $W_1$ to code for the difference between round and square puzzles, as

   $$W_{1i} = \frac{Y_{\text{R,BW},i} + Y_{\text{R,C},i} - Y_{\text{S,BW},i} - Y_{\text{S,C},i}}{\sqrt{4}}$$

   and compare

   $$\text{MODEL C} : W_{1i} = 0 + \epsilon_i$$

   to

   $$\text{MODEL A} : W_{1i} = \beta_0 + \epsilon_i$$

   Make sure you write down SSE(C) and SSE(A). What null-hypothesis does this comparison test? Perform this test and interpret the results.

```
> # load the data
> library(foreign)
> dat <- as.data.frame(read.spss("puzzles.sav"))
> # compute composite variable
> dat$W1 <- (dat$R_BW + dat$R_C - dat$S_BW - dat$S_C)/sqrt(4)
> # estimate model A
> mod_A <- lm(W1~1,data=dat)
> summary(mod_A)


Call:
lm(formula = W1 ~ 1, data = dat)
```

```
Residuals:
   Min     1Q Median     3Q    Max 
-2.000 -1.125  0.000  1.125  1.500 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)   1.0000     0.3641   2.746    0.019 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.261 on 11 degrees of freedom

> # "estimate" model C
> mod_C <- lm(W1~-1,data=dat) # -1 means no intercept
> # compare to model C
> anova(mod_C,mod_A)

Analysis of Variance Table

Model 1: W1 ~ -1
Model 2: W1 ~ 1
  Res.Df  RSS Df Sum of Sq      F  Pr(>F)  
1     12 29.5                              
2     11 17.5  1        12 7.5429 0.01901 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*Answer:* SSE(C) = 29.5 and SSE(A) = 17.5. The null hypothesis is that there is no difference between round and square pictures, i.e. $H_0 : \beta_0 = 0$, or in terms of (marginal) means, $H_0 : \mu_{\mathrm{R},\cdot,i} = \mu_{\mathrm{S},\cdot,i}$. The test statistic is $F = \frac{(29.5-17.5)/(1-0)}{17.5/(12-1)} = 7.543$, and the probability of obtaining a value at least as large as this, if $H_0$ were actually true, is $P(F_{1,11} \geq 7.543) = .019$. As this is smaller than a significance level of $\alpha = .05$, $H_0$ can be rejected. The estimated intercept is $b_0 = 1$, which in this model is equal to the mean, $\overline{W}_1 = 1$. With the contrast coding used, that means that the average solution times for square puzzles are shorter than those for round puzzles. We should rescale this mean back to the scale of the dependent variable as $\frac{\overline{W}_1}{\sqrt{4}} = .5$, which gives us half of the difference in means. So round puzzles take on average 1 minute longer than square puzzles.

(b) Define a composite variable $W_2$ to code for the difference between black-and-white and colour puzzles. Test whether the mean of $W_2$ differs from 0 and interpret the results. Make sure you write down SSE(C) and SSE(A).

```
> # compute composite variable
> dat$W2 <- (dat$R_BW + dat$S_BW - dat$R_C - dat$S_C)/sqrt(4)
> # estimate model A
```

```
> mod_A <- lm(W2~1,data=dat)
> summary(mod_A)


Call:
lm(formula = W2 ~ 1, data = dat)

Residuals:
   Min     1Q Median     3Q    Max
-1.500 -0.625  0.000  0.625  1.500

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.0000     0.2683   3.728  0.00334 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9293 on 11 degrees of freedom

> # "estimate" model C
> mod_C <- lm(W2~-1,data=dat) # -1 means no intercept
> # compare to model C
> anova(mod_C,mod_A)

Analysis of Variance Table

Model 1: W2 ~ -1
Model 2: W2 ~ 1
  Res.Df  RSS Df Sum of Sq      F   Pr(>F)
1     12 21.5
2     11  9.5  1        12 13.895 0.003338 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*Answer:* I defined the composite variable as

$$W_{2i} = \frac{Y_{\text{R,BW},i} + Y_{\text{S,BW},i} - Y_{\text{R,C},i} - Y_{\text{S,C},i}}{\sqrt{4}}$$

which gives an SSE(C) $= 21.5$ and SSE(A) $= 9.5$. The null hypothesis is that there is no difference between black-and-white and colour pictures, i.e. $H_0 : \mu_{\cdot,\text{BW},i} = \mu_{\cdot,\text{C},i}$. The test statistic is $F = \frac{21.5-9.5}{9.5/(12-1)} = 13.894$, $P(F_{1,11} > 13.894) = .003$, so $H_0$ can be rejected. The mean $\overline{W}_2 = 1$, so the interpretation is similar to that above: black-and-white puzzles take on average 1 minute longer than colour puzzles.

(c) Define a composite variable $W_3$ to test for an interaction between puzzle colour and shape, and test whether the mean of $W_3$ differs from 0. Again, write down SSE(C) and SSE(A).

```
> # compute composite variable
> dat$W3 <- (dat$R_BW - dat$S_BW - dat$R_C + dat$S_C)/sqrt(4)
> # estimate model A
> mod_A <- lm(W3~1,data=dat)
> summary(mod_A)


Call:
lm(formula = W3 ~ 1, data = dat)

Residuals:
   Min     1Q Median     3Q    Max 
 -2.00  -1.50  -0.25   1.50   2.50 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.282e-16  4.807e-01       0        1

Residual standard error: 1.665 on 11 degrees of freedom

> # "estimate" model C
> mod_C <- lm(W3~-1,data=dat) # -1 means no intercept
> # compare to model C
> anova(mod_C,mod_A)

Analysis of Variance Table

Model 1: W3 ~ -1
Model 2: W3 ~ 1
  Res.Df  RSS Df Sum of Sq  F Pr(>F)
1     12 30.5                        
2     11 30.5  1         0  0      1
```

*Answer:* I defined the composite variable as

$$W_{3i} = \frac{Y_{\text{R,BW},i} - Y_{\text{S,BW},i} - Y_{\text{R,C},i} + Y_{\text{S,C},i}}{\sqrt{4}}$$

which gives an SSE(C) = 30.5 and SSE(A) = 30.5. The null-hypothesis is that the difference between round and square puzzles is identical for black-and-white and colour puzzles, i.e. $H_0 : \mu_{\text{R,BW},i} - \mu_{\text{S,BW},i} = \mu_{\text{R,C},i} - \mu_{\text{S,C},i}$. As the SSE for the two models are identical (as only really happens with fictitious data), we can be sure that $H_0$ will not be rejected (it gives an $F_{1,11} = 0$).

(d) Repeat the analysis above using the SPSS Repeated Measures option. Compare the results to those obtained above (they should be identical).

```
> library(car)
> mod <- lm(cbind(R_BW,S_BW,R_C,S_C) ~ 1,data=dat)
> idata <- data.frame(
+   colour=factor(c(1,1,2,2),labels=c("Black-White","Colour")),
+   shape=factor(c(1,2,1,2),labels=c("Round","Square")))
> # make sure you use orthogonal contrasts
> contrasts(idata$colour) <- contr.helmert(2)
> contrasts(idata$shape) <- contr.helmert(2)
> puzzles_aov <- Anova(mod,idata=idata,idesign=~colour*shape,type=3)
> summary(puzzles_aov,multivariate=FALSE)


Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

                SS num Df Error SS den Df         F    Pr(>F)
(Intercept)  97200      1    226.5     11 4720.5298 7.708e-16 ***
colour          12      1      9.5     11   13.8947  0.003338 **
shape           12      1     17.5     11    7.5429  0.019012 *
colour:shape     0      1     30.5     11    0.0000  1.000000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*Answer:* Of course we get exactly the same results.

2. A study was conducted to compare the effect of two treatments for stress-related headache in children. Treatment 1 is a behavioural therapy focused on relaxation techniques, while treatment 2 is a cognitive therapy focused on stress-perception. Each treatment lasts for 20 weeks. Fourteen children each received one of the therapies. Every four weeks during treatment, the number of headache-free days in that period was recorded. This resulted in the following data:

| id | treatment | $t=1$ | $t=2$ | $t=3$ | $t=4$ | $t=5$ | $\overline{Y}_i$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 4 | 6 | 2 | 22 | 7.20 |
| 2 | 1 | 6 | 9 | 11 | 10 | 20 | 11.20 |
| 3 | 1 | 9 | 3 | 2 | 6 | 10 | 6.00 |
| 4 | 1 | 8 | 3 | 11 | 8 | 3 | 6.60 |
| 5 | 1 | 14 | 4 | 5 | 16 | 22 | 12.20 |
| 6 | 1 | 7 | 6 | 7 | 9 | 15 | 8.80 |
| 7 | 1 | 5 | 5 | 4 | 8 | 16 | 7.60 |
| 8 | 2 | 8 | 7 | 7 | 9 | 12 | 8.60 |
| 9 | 2 | 13 | 15 | 18 | 11 | 10 | 13.40 |
| 10 | 2 | 5 | 9 | 14 | 11 | 12 | 10.20 |
| 11 | 2 | 8 | 8 | 11 | 16 | 14 | 11.40 |
| 12 | 2 | 13 | 15 | 15 | 9 | 10 | 12.40 |
| 13 | 2 | 6 | 9 | 9 | 13 | 4 | 8.20 |
| 14 | 2 | 9 | 13 | 14 | 11 | 11 | 11.60 |
| $\overline{Y}_{\cdot j}$ | | 8.07 | 7.86 | 9.57 | 9.93 | 12.93 | 9.67 |

You can also find the data on Moodle as `headaches.sav`.

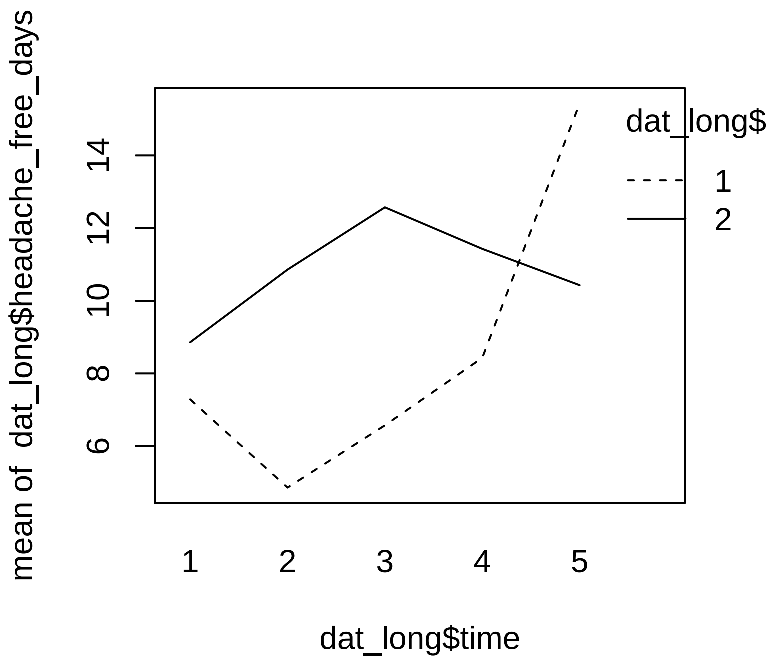
(a) Use the individual means ($\overline{Y}_i$) as dependent variable and test for an effect of treatment. What do you conclude?

```
> # load the data
> dat <- as.data.frame(read.spss("headaches.sav"))

Warning in read.spss("headaches.sav"):  headaches.sav:  File-indicated
value is different from internal value for at least one of the three
system values.  SYSMIS: indicated -1.79769e+308, expected -1.79769e+308;
HIGHEST: 1.79769e+308, 1.79769e+308; LOWEST: -1.79769e+308, -1.79769e+308
Warning in read.spss("headaches.sav"):  headaches.sav:  Unrecognized
record type 7, subtype 18 encountered in system file

> # compute the individual means
> dat$imean <- rowMeans(dat[,paste("t",1:5,sep="")])
> # do an anova on the individual means
> # first make treatmen a factor
> dat$treatmen <- as.factor(dat$treatmen)
> contrasts(dat$treatmen) <- contr.helmert(2)
> # now do the ANOVA
> summary(aov(imean~treatmen,data=dat))

            Df Sum Sq Mean Sq F value Pr(>F)
treatmen     1  18.75  18.746   4.038 0.0675 .
Residuals   12  55.70   4.642
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*Answer:* Using the individual means as dependent variable and performing a oneway ANOVA with treatment as independent variable gives the following test result for the effect of treatment: $F_{1,12} = 4.038$, $p = .068$. So the null hypothesis $H_0 : \mu_1 = \mu_2$ cannot be rejected, and the treatments seem equally effective in reducing headaches.

(b) Plot the means over time for each treatment. What can you see?

```
> # turn the data into the "long" format
> dat_long <- reshape(data = dat[,-8],
+                     varying = 3:7,
+                     v.names = "headache_free_days",
+                     timevar = "time",
+                     times =  1:5,
+                     idvar = c("id","treatmen"),
+                     new.row.names=1:1000,
+                     direction = "long")
> interaction.plot(x.factor=dat_long$time,
+                  trace.factor = dat_long$treatmen,
+                  response=dat_long$headache_free_days)
```




*Answer:* There seems to be a different pattern over time: Behavioural therapy first reduces, but then increases the number of headache-free days, while the cognitive therapy first increases, but then decreases and in general seems more constant over time (it may have no effect).

(c) Perform a repeated measures ANOVA. Don't forget to check the assumptions underlying the repeated measures analysis. Does this change your conclusion about the treatments? What else can you conclude?

```
> mod <- lm(cbind(t1,t2,t3,t4,t5) ~ treatmen,data=dat)
> idata <- data.frame(time = ordered(1:5))
> contrasts(idata$time) # this should be polynomial, so orthogonal :-)

             .L         .Q            .C         ^4
[1,] -0.6324555  0.5345225 -3.162278e-01  0.1195229
[2,] -0.3162278 -0.2672612  6.324555e-01 -0.4780914
[3,]  0.0000000 -0.5345225 -4.095972e-16  0.7171372
[4,]  0.3162278 -0.2672612 -6.324555e-01 -0.4780914
[5,]  0.6324555  0.5345225  3.162278e-01  0.1195229

> head_aov <- Anova(mod,idata=idata,idesign=~time,data=dat,type=3)
> summary(head_aov,multivariate=FALSE)


Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

                  SS num Df Error SS den Df        F    Pr(>F)
(Intercept)   6547.6      1   278.51     12 282.1065  1.06e-09 ***
treatmen        93.7      1   278.51     12   4.0384 0.0675114 .
time           231.5      4   609.77     48   4.5561 0.0033608 **
treatmen:time  285.9      4   609.77     48   5.6267 0.0008508 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Mauchly Tests for Sphericity

              Test statistic  p-value
time                 0.19869 0.054085
treatmen:time        0.19869 0.054085


Greenhouse-Geisser and Huynh-Feldt Corrections
 for Departure from Sphericity

               GG eps Pr(>F[GG])
time          0.63155   0.012876 *
treatmen:time 0.63155   0.005145 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                 HF eps  Pr(>F[HF])
```

```
time           0.8138331 0.006592563
treatmen:time  0.8138331 0.002099941
```

*Answer:* Mauchley's test for Sphericity is not significant, although pretty close. It may be wise to be cautious and look at the Huynh-Feldt (or Greenhouse-Geisser) corrected tests. You can also save the residuals of the repeated-measures ANOVA and run SPSS explore to test for normality, etc. However, as there are not many observations in each condition, this will be of limited use here.

The test for the between-subjects effect of treatment is the same as the test performed above, so no overall difference between treatments. Looking at the omnibus tests for the within-subjects effects, we see a significant main effect of time, $F_{4,48} = 4.556$, $p = .003$. Using the Huynh-Feldt correction for the degrees of freedom also gives a significant result, $F_{3.522,42.462} = 4.556$, $p = .005$. There is also a significant interaction between treatment and time, $F_{4,48} = 5.627$, $p = .001$. So the number of headache-free days varies during each treatment (significant effect of time), and the way in which it does differs between the treatments (interaction between time and treatment).

(d) How many orthogonal contrasts (for both within and between participants effects) can you test? What are the assumptions underlying each test?

*Answer:* There are five time points, so we can use four composite variables (contrasts) for time. There are 2 groups, so we can use one contrast to code for groups. Note that we can also tests for interactions of time with treatment, using $1 \times 4 = 4$ contrast. So in total, there are $4 + 4 + 1 = 9$ possible contrasts. The assumptions underlying each contrast are mild (identical to the GLM). In particular, we need no sphericity when we test each contrast separately (these are 1 degree of freedom tests).

(e) By default, SPSS will apply a polynomial contrast code to time. Look at the results of the within-subjects contrasts and interpret the results.

```
> # To get the results of the individual contrasts, we will fit the models
> # set up contrast codes
> dat$lin <- -2*dat$t1 + -1*dat$t2 + 0*dat$t3 + 1*dat$t4 + 2*dat$t5
> dat$quad <- 2*dat$t1 + -1*dat$t2 + -2*dat$t3 + -1*dat$t4 + 2*dat$t5
> dat$cub <-  -1*dat$t1 + 2*dat$t2 + 0*dat$t3 + -2*dat$t4 + 1*dat$t5
> dat$fourth <-  1*dat$t1 + -4*dat$t2 + 6*dat$t3 + -4*dat$t4 + 1*dat$t5
> # linear contrast (intercept) (and interaction (treatmen)
> summary(lm(lin~treatmen,data=dat))


Call:
lm(formula = lin ~ treatmen, data = dat)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-24.857 -11.214   2.143   8.893  18.143

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   11.786      3.661   3.219  0.00737 **
treatmen1     -8.071      3.661  -2.204  0.04776 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.7 on 12 degrees of freedom
Multiple R-squared:  0.2882,Adjusted R-squared:  0.2289
F-statistic: 4.859 on 1 and 12 DF,  p-value: 0.04776

> # quadratic contrast (intercept) (and interaction (treatmen)
> summary(lm(quad~treatmen,data=dat))


Call:
lm(formula = quad ~ treatmen, data = dat)

Residuals:
    Min      1Q  Median      3Q     Max
-30.000  -6.643  -1.143   6.643  23.000

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.071      3.688   1.375  0.19421
treatmen1    -13.929      3.688  -3.777  0.00264 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.8 on 12 degrees of freedom
Multiple R-squared:  0.5431,Adjusted R-squared:  0.505
F-statistic: 14.26 on 1 and 12 DF,  p-value: 0.002639

> # cubic contrast (intercept) (and interaction (treatmen)
> summary(lm(cub~treatmen,data=dat))


Call:
lm(formula = cub ~ treatmen, data = dat)

Residuals:
    Min      1Q  Median      3Q     Max
-17.000  -9.321   1.786   5.321  23.000
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.7143     3.0839   0.232    0.821
treatmen1    -0.2857     3.0839  -0.093    0.928

Residual standard error: 11.54 on 12 degrees of freedom
Multiple R-squared:  0.0007148,Adjusted R-squared:  -0.08256
F-statistic: 0.008584 on 1 and 12 DF,  p-value: 0.9277
```

```
> # fourth-order contrast (intercept) (and interaction (treatmen)
> summary(lm(fourth~treatmen,data=dat))


Call:
lm(formula = fourth ~ treatmen, data = dat)

Residuals:
    Min      1Q  Median      3Q     Max
-29.571 -13.893  -1.286  14.429  27.000

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    7.286      5.065   1.438    0.176
treatmen1     -1.714      5.065  -0.338    0.741

Residual standard error: 18.95 on 12 degrees of freedom
Multiple R-squared:  0.009456,Adjusted R-squared:  -0.07309
F-statistic: 0.1146 on 1 and 12 DF,  p-value: 0.7409
```

*Answer:* There is a significant effect of the linear contrast for time (average number of headache-free days increases linearly over time). More importantly, there is a significant interaction with treatment. Also, there is a significant interaction between the quadratic contrast and treatment, which is due to the difference in shape (Behavioural goes down and then up, while cognitive goes up and then down).

3. Segal, Bogaards, Becker, & Chatman (1999) conducted a study on the effectiveness of verbally disclosing thoughts and feelings about the loss of a spouse. Participants participated in four 20-minute sessions within a 2-week period. In each session the participant was asked to talk about the loss of their spouse and to express their deepest thoughts and feelings for the entire 20 minutes. The experimental sessions were conducted in the participants home. Immediately before and after each session, participants completed the Positive and Negative Affect Schedule (PANAS), which provides measures of participants' positive and negative affect. You can find the data on Moodle as `VerbalDisclosure.sav`. This data was also discussed in the lecture.

(a) Perform a repeated-measures ANOVA on the negative affect measures, using `gender` as a between-subjects factor, and `time` and `prepost` as within subject measures. For `time`, choose a polynomial contrast, and for `prepost` and `gender` you can choose e.g. a Helmert contrast. Analyse the results, both in terms of the omnibus tests as well as the tests for the contrasts and interactions.

```
> # load the data
> dat <- as.data.frame(read.spss("VerbalDisclosure.sav"))
> # get rid of observation with only missing values
> dat <- dat[-28,]
> contrasts(dat$gender) <- contr.helmert(2)
> # estimate a model for the negative affect
> mod_neg <-lm(cbind(panpren1,panpren2,panpren3,panpren4,
+                    panpstn1,panpstn2,panpstn3,panpstn4)~gender,data=dat)
> # set up a data.frame with within-subjects design
> idata <- data.frame(time=factor(c(1:4,1:4)),
+                     prepost=factor(rep(c(1,2),each=4),
+                                    labels=c("pre","post")))
> contrasts(idata$time) <-contr.poly(4)
> contrasts(idata$prepost) <- contr.helmert(2)
> aov_neg <- Anova(mod_neg,idata=idata,idesign=~time*prepost,type=3)
> summary(aov_neg,multivariate=FALSE)

Warning in summary.Anova.mlm(aov_neg, multivariate = FALSE): HF eps >
1 treated as 1


Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

                       SS num Df Error SS den Df        F    Pr(>F)
(Intercept)         87227      1  13713.7     27 171.7348 3.224e-13 ***
gender                252      1  13713.7     27   0.4969 0.4868811
time                   60      3   1489.1     81   1.0945 0.3562932
gender:time           419      3   1489.1     81   7.5997 0.0001536 ***
prepost               479      1    775.3     27  16.6975 0.0003521 ***
gender:prepost        164      1    775.3     27   5.6940 0.0242972 *
time:prepost           67      3    842.2     81   2.1431 0.1011810
gender:time:prepost   114      3    842.2     81   3.6394 0.0161466 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Mauchly Tests for Sphericity

                    Test statistic p-value
time                       0.81134 0.37188
```

```
gender:time                 0.81134 0.37188
time:prepost                0.60978 0.02623
gender:time:prepost         0.60978 0.02623


Greenhouse-Geisser and Huynh-Feldt Corrections
 for Departure from Sphericity

                     GG eps Pr(>F[GG])
time                0.90129  0.3531674
gender:time         0.90129  0.0002838 ***
time:prepost        0.77972  0.1177859
gender:time:prepost 0.77972  0.0258098 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                       HF eps    Pr(>F[HF])
time                1.0112499 0.3562932105
gender:time         1.0112499 0.0001535854
time:prepost        0.8582653 0.1115979175
gender:time:prepost 0.8582653 0.0218201286

> # using the contrasts we have already defined
> # the following is a shortcut to compute
> # the "W" variables, using a little matrix
> # algebra...
>
> contrast_matrix <- model.matrix(~time*prepost,data=idata)[,-1]
> within_dependents <- model.response(model.frame(mod_neg))
> W <- within_dependents%*%contrast_matrix
> # the ordeer of the following tests is:
> colnames(W)

[1] "time.L"          "time.Q"          "time.C"          "prepost1"
[5] "time.L:prepost1" "time.Q:prepost1" "time.C:prepost1"

> # perform the tests using a loop
> for(i in 1:7) {
+   # print the name  of the effect we're looking at:
+   print(colnames(W)[i])
+   # print the results
+   print(summary(lm(W[,i]~dat$gender)))
+ }

[1] "time.L"
```

```
Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
    Min      1Q  Median      3Q     Max
-15.256  -4.299   1.291   5.092   7.539

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.9489     1.4531  -0.653 0.519292
dat$gender1  -5.7085     1.4531  -3.928 0.000535 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.697 on 27 degrees of freedom
Multiple R-squared:  0.3637,Adjusted R-squared:  0.3401
F-statistic: 15.43 on 1 and 27 DF,  p-value: 0.0005347

[1] "time.Q"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
     Min       1Q   Median       3Q      Max
-12.7955  -3.2955  -0.2955   2.2045  17.8571

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.969      1.370   1.438    0.162
dat$gender1   -1.674      1.370  -1.222    0.232

Residual standard error: 6.312 on 27 degrees of freedom
Multiple R-squared:  0.05242,Adjusted R-squared:  0.01732
F-statistic: 1.494 on 1 and 27 DF,  p-value: 0.2322

[1] "time.C"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
     Min       1Q   Median       3Q      Max
-15.7033  -1.8397   0.8305   2.1852   8.6698
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.9518     1.0980   0.867    0.394
dat$gender1  -2.0190     1.0980  -1.839    0.077 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.061 on 27 degrees of freedom
Multiple R-squared:  0.1113,Adjusted R-squared:  0.07837
F-statistic: 3.381 on 1 and 27 DF,  p-value: 0.07698

[1] "prepost1"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
    Min      1Q  Median      3Q     Max
-24.591  -6.591  -3.591   3.714  49.714

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   13.438      3.289   4.086 0.000352 ***
dat$gender1   -7.847      3.289  -2.386 0.024297 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.16 on 27 degrees of freedom
Multiple R-squared:  0.1742,Adjusted R-squared:  0.1436
F-statistic: 5.694 on 1 and 27 DF,  p-value: 0.0243

[1] "time.L:prepost1"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
    Min      1Q  Median      3Q     Max
-8.0803 -1.8193 -0.2541  0.8639 17.1858

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.267      1.121   2.023  0.05303 .
dat$gender1   -3.131      1.121  -2.794  0.00945 **
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.164 on 27 degrees of freedom
Multiple R-squared:  0.2243,Adjusted R-squared:  0.1956
F-statistic: 7.809 on 1 and 27 DF,  p-value: 0.009449

[1] "time.Q:prepost1"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
     Min       1Q   Median       3Q      Max
-11.5714  -2.4318   0.4286   1.5682  14.4286

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.7484     0.9972  -0.750    0.459
dat$gender1   0.1802     0.9972   0.181    0.858

Residual standard error: 4.596 on 27 degrees of freedom
Multiple R-squared:  0.001208,Adjusted R-squared:  -0.03578
F-statistic: 0.03265 on 1 and 27 DF,  p-value: 0.858

[1] "time.C:prepost1"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
    Min      1Q  Median      3Q     Max
-8.1210 -2.8111  0.1525  1.4941  7.5315

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.7703     0.8288   0.929    0.361
dat$gender1  -0.9227     0.8288  -1.113    0.275

Residual standard error: 3.82 on 27 degrees of freedom
Multiple R-squared:  0.0439,Adjusted R-squared:  0.008486
F-statistic:  1.24 on 1 and 27 DF,  p-value: 0.2754
```

*Answer:* First of all, note that Mauchly's test indicates that the sphericity assumption does not hold for the `prepost` $\times$ `time` interaction, so for any tests involving this interaction, we should use e.g. the Huynh-Feldt correction on the degrees of freedom. Looking at the omnibus tests, there is a significant main effect of `prepost`. There is also a significant `prepost` × `gender` interaction, a significant `time` × `gender` interaction, and a significant threeway interaction between `prepost`, `time` and `gender`, for which we look at the Huynh-Feldt corrected test. To interpret these significant effects, we can look at the individual contrasts and plots. The main effect of `prepost` shows that negative feelings generally increase immediately after a session. The `prepost` × `gender` appears mainly due to the moderation of the linear effect of `time` by `gender`: for females, negative feelings decrease over sessions, while for males, negative feelings increase (the treatment doesn't appear to be very effective for males). The three-way interaction is the hardest to interpret. Again, this appears to be mainly due to a moderation of the linear effect of `time`. Looking at the plots, what appears to happen is that the increase in negative feelings from pre- to posttest gets larger at each session for males. For females, on the other hand, the increase seems to get a little bit smaller at each session.

(b) Repeat the analysis above for the positive affect measures.

```
> # estimate a model for the positive affect
> mod_pos <-lm(cbind(panprep1,panprep2,panprep3,panprep4,
+                     panpstp1,panpstp2,panpstp3,panpstp4)~gender,data=dat)
> aov_pos <- Anova(mod_pos,idata=idata,idesign=~time*prepost,type=3)
> summary(aov_pos,multivariate=FALSE)


Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

                        SS num Df Error SS den Df        F  Pr(>F)
(Intercept)         151670      1  12581.9     27 325.4744 < 2e-16 ***
gender                  91      1  12581.9     27   0.1957 0.66175
time                    54      3   1404.5     81   1.0296 0.38400
gender:time            107      3   1404.5     81   2.0584 0.11222
prepost                 52      1    402.5     27   3.4576 0.07389 .
gender:prepost           1      1    402.5     27   0.0363 0.85037
time:prepost            11      3    696.4     81   0.4318 0.73082
gender:time:prepost      1      3    696.4     81   0.0561 0.98243
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Mauchly Tests for Sphericity

                    Test statistic  p-value
time                       0.71310 0.122026
gender:time                0.71310 0.122026
```

```
time:prepost                0.56852 0.012664
gender:time:prepost         0.56852 0.012664


Greenhouse-Geisser and Huynh-Feldt Corrections
 for Departure from Sphericity

                     GG eps Pr(>F[GG])
time                0.85215     0.3767
gender:time         0.85215     0.1230
time:prepost        0.79085     0.6851
gender:time:prepost 0.79085     0.9645

                       HF eps Pr(>F[HF])
time                0.9488622  0.3816737
gender:time         0.9488622  0.1158317
time:prepost        0.8720796  0.7042433
gender:time:prepost 0.8720796  0.9730566

> # using the contrasts we have already defined
> # the following is a shortcut to compute
> # the "W" variables, using a little matrix
> # algebra
> within_dependents <- model.response(model.frame(mod_pos))
> W <- within_dependents%*%contrast_matrix
> # the ordeer of the following tests is:
> colnames(W)

[1] "time.L"          "time.Q"          "time.C"          "prepost1"
[5] "time.L:prepost1" "time.Q:prepost1" "time.C:prepost1"

> # perform the tests
> for(i in 1:7) {
+   print(colnames(W)[i])
+   print(summary(lm(W[,i]~dat$gender)))
+ }

[1] "time.L"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
     Min       1Q   Median       3Q      Max
-11.6784  -4.5230  -0.0508   3.3033  13.1420
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.230      1.357   1.643    0.112
dat$gender1   -2.179      1.357  -1.606    0.120

Residual standard error: 6.254 on 27 degrees of freedom
Multiple R-squared:  0.08716,Adjusted R-squared:  0.05335
F-statistic: 2.578 on 1 and 27 DF,  p-value: 0.12

[1] "time.Q"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
    Min      1Q  Median      3Q     Max
-9.1136 -3.1136 -0.1136  3.8864 10.8864

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.2354     1.0521   0.224    0.825
dat$gender1   1.3782     1.0521   1.310    0.201

Residual standard error: 4.849 on 27 degrees of freedom
Multiple R-squared:  0.05976,Adjusted R-squared:  0.02494
F-statistic: 1.716 on 1 and 27 DF,  p-value: 0.2012

[1] "time.C"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
     Min       1Q   Median       3Q      Max
-20.5413  -3.4180   0.4777   3.9610  11.7873

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.1285     1.3963   0.092    0.927
dat$gender1   1.8535     1.3963   1.327    0.195

Residual standard error: 6.435 on 27 degrees of freedom
Multiple R-squared:  0.06126,Adjusted R-squared:  0.02649
F-statistic: 1.762 on 1 and 27 DF,  p-value: 0.1955
```

```
[1] "prepost1"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
    Min      1Q  Median      3Q     Max 
-36.143  -5.045   1.955   6.955  13.955 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)  -4.4058     2.3694  -1.859   0.0739 .
dat$gender1   0.4513     2.3694   0.190   0.8504  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.92 on 27 degrees of freedom
Multiple R-squared:  0.001342,Adjusted R-squared:  -0.03565 
F-statistic: 0.03628 on 1 and 27 DF,  p-value: 0.8504

[1] "time.L:prepost1"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
   Min     1Q Median     3Q    Max 
-9.320 -3.059 -1.047  1.860 10.581 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.7151     0.9603   0.745    0.463
dat$gender1   0.3318     0.9603   0.345    0.732

Residual standard error: 4.426 on 27 degrees of freedom
Multiple R-squared:  0.004401,Adjusted R-squared:  -0.03247 
F-statistic: 0.1194 on 1 and 27 DF,  p-value: 0.7324

[1] "time.Q:prepost1"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
```

```
   Min     1Q Median     3Q    Max
-10.25  -2.25   0.75   1.75   6.75

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.73214    0.74381   0.984    0.334
dat$gender1  0.01786    0.74381   0.024    0.981

Residual standard error: 3.428 on 27 degrees of freedom
Multiple R-squared:  2.135e-05,Adjusted R-squared:  -0.03701
F-statistic: 0.0005764 on 1 and 27 DF,  p-value: 0.981

[1] "time.C:prepost1"

Call:
lm(formula = W[, i] ~ dat$gender)

Residuals:
    Min      1Q  Median      3Q     Max
-7.8567 -1.8193 -0.2541  2.2056 12.9387

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.03267    0.97618  -0.033    0.974
dat$gender1 -0.16045    0.97618  -0.164    0.871

Residual standard error: 4.499 on 27 degrees of freedom
Multiple R-squared:  0.0009995,Adjusted R-squared:  -0.036
F-statistic: 0.02701 on 1 and 27 DF,  p-value: 0.8707
```

*Answer:* As for the negative affect measure, Mauchly's test indicates that the sphericity assumption does not hold for the `prepost` × `time` interaction, so for any tests involving this interaction, we should use e.g. the Huynh-Feldt correction on the degrees of freedom. Looking at the omnibus tests, there is a trend towards a main effect of `prepost`, $F_{1,27} = 3.458$, $p = .074$. There are no significant main effects or interactions.