# R Workshop Session 2

Nichole Bouffard

7/22/2020

## Description

This is a printable version of the script we'll be going over in today's session. Feel free to print it out and follow along. We will be covering how to run anovas and t tests for the single item recognition data. We will use the package 'afex' to run anovas and plot the effects.

## First load in the three packages below

Note, if you don't already have them downloaded you can uncomment the "install.packages" lines (delete #) and run those first, then load them using the library() line.

- tidyverse - useful datawrangling functions, piping, and ggplot
- afex - this is package I like to use to run my stats. It has a function "aov_ez" that I use instead of R's built in anova function "aov". It takes a more intutive input compared the aov formulas
- emmeans - for running the pairwise comparisons (t tests) after the anova

```
# install.pacakges(tidyverse)
```

```
library(tidyverse)
# install.packages(afex)
library(afex)
#install.packages(emmeans)
library(emmeans)
```

## Read in data

```
sirdat <-read_csv('/Users/nicholebouffard/Dropbox/Baycrest_Rworksh
```

```
## Parsed with column specification:
## cols(
##   subid = col_double(),
##   stimulus = col_character(),
##   group = col_character(),
##   DO_SS = col_double(),
##   SO_DS = col_double(),
##   SO_SS = col_double()
## )
```

I'm using read_csv here instead of read.csv. You must have the tidyverse loaded first before you can use read_csv, but it reads your data into a tibble way quicker and more efficient than reading your data into a data.frame using read.csv

## Data wrangle from wide to long format (long = one observation per row)

```
sirdat <- sirdat %>% #shortcut to pipe = command+shift+M
  gather(condition, hit_minus_fa,4:6)
```

## Look at how the different variables are categorized. Change grouping factors to factors

```
str(sirdat)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    354 obs. of  5 var
##  $ subid       : num  10 10 11 11 12 12 13 13 14 14 ...
##  $ stimulus    : chr  "object" "scene" "object" "scene" ...
##  $ group       : chr  "YA" "YA" "YA" "YA" ...
##  $ condition   : chr  "DO_SS" "DO_SS" "DO_SS" "DO_SS" ...
##  $ hit_minus_fa: num  0.912 0.715 0.632 0.562 0.824 ...
```

```
sirdat$subid <- as.factor(sirdat$subid)
sirdat$stimulus<- as.factor(sirdat$stimulus)
sirdat$group <- as.factor(sirdat$group)
sirdat$condition <- as.factor(sirdat$condition)
str(sirdat)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    354 obs. of  5 var
##  $ subid       : Factor w/ 59 levels "10","11","12",..: 1 1 2 2
##  $ stimulus    : Factor w/ 2 levels "object","scene": 1 2 1 2 1
##  $ group       : Factor w/ 2 levels "OA","YA": 2 2 2 2 2 2 2 2
##  $ condition   : Factor w/ 3 levels "DO_SS","SO_DS",..: 1 1 1 1
##  $ hit_minus_fa: num  0.912 0.715 0.632 0.562 0.824 ...
```

# DESCRIPTIVE STATISTICS

Calculate number of subjects

```
# Younger Adults
nYA <- length(unique(sirdat$subid[sirdat$group == 'YA'])) #31
nYA
```

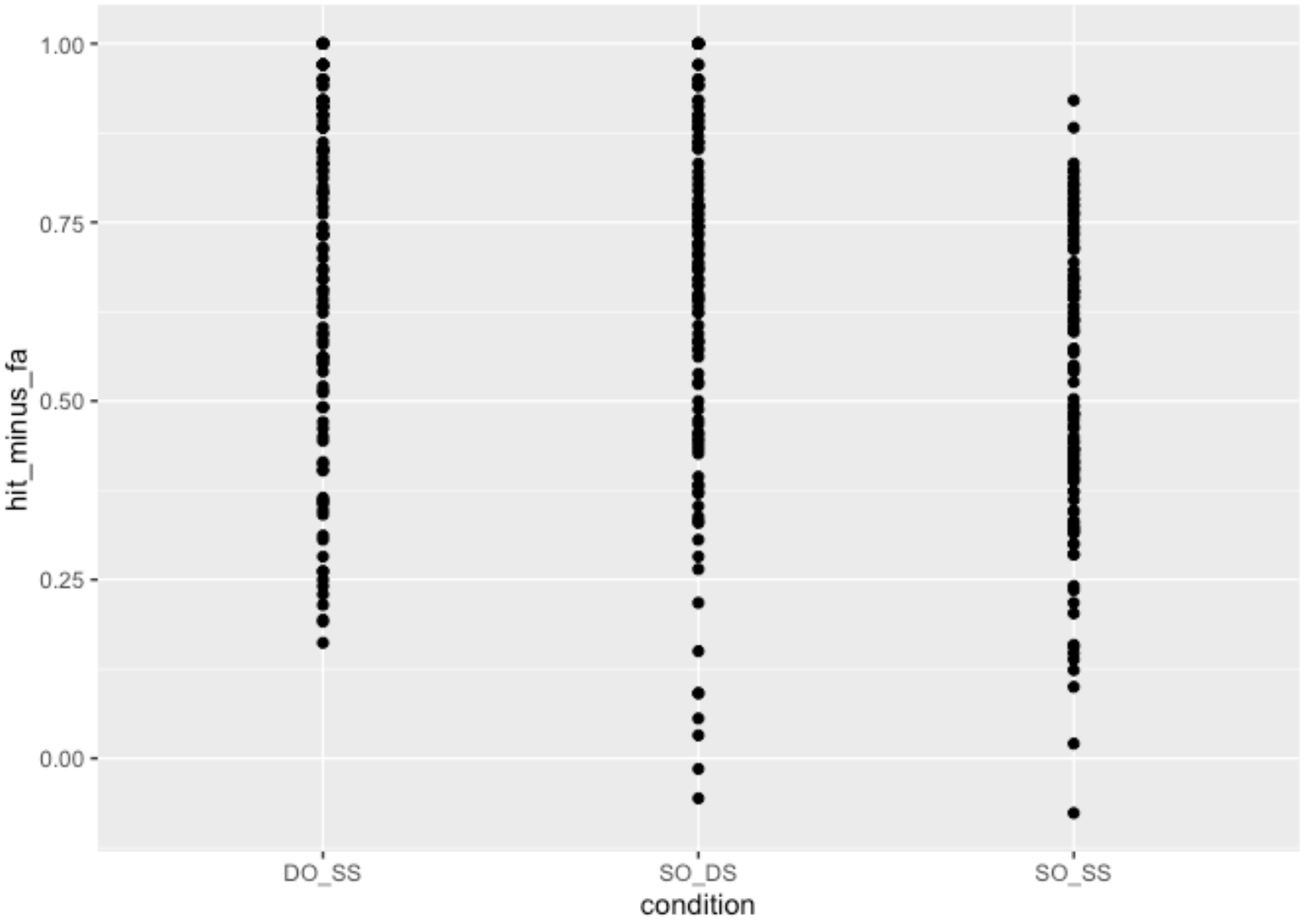
```
## [1] 31
```

```
# Older Adults
OA <- sirdat %>%
  filter(group == 'OA')
nOA <- length(unique(OA$subid)) #28
nOA
```

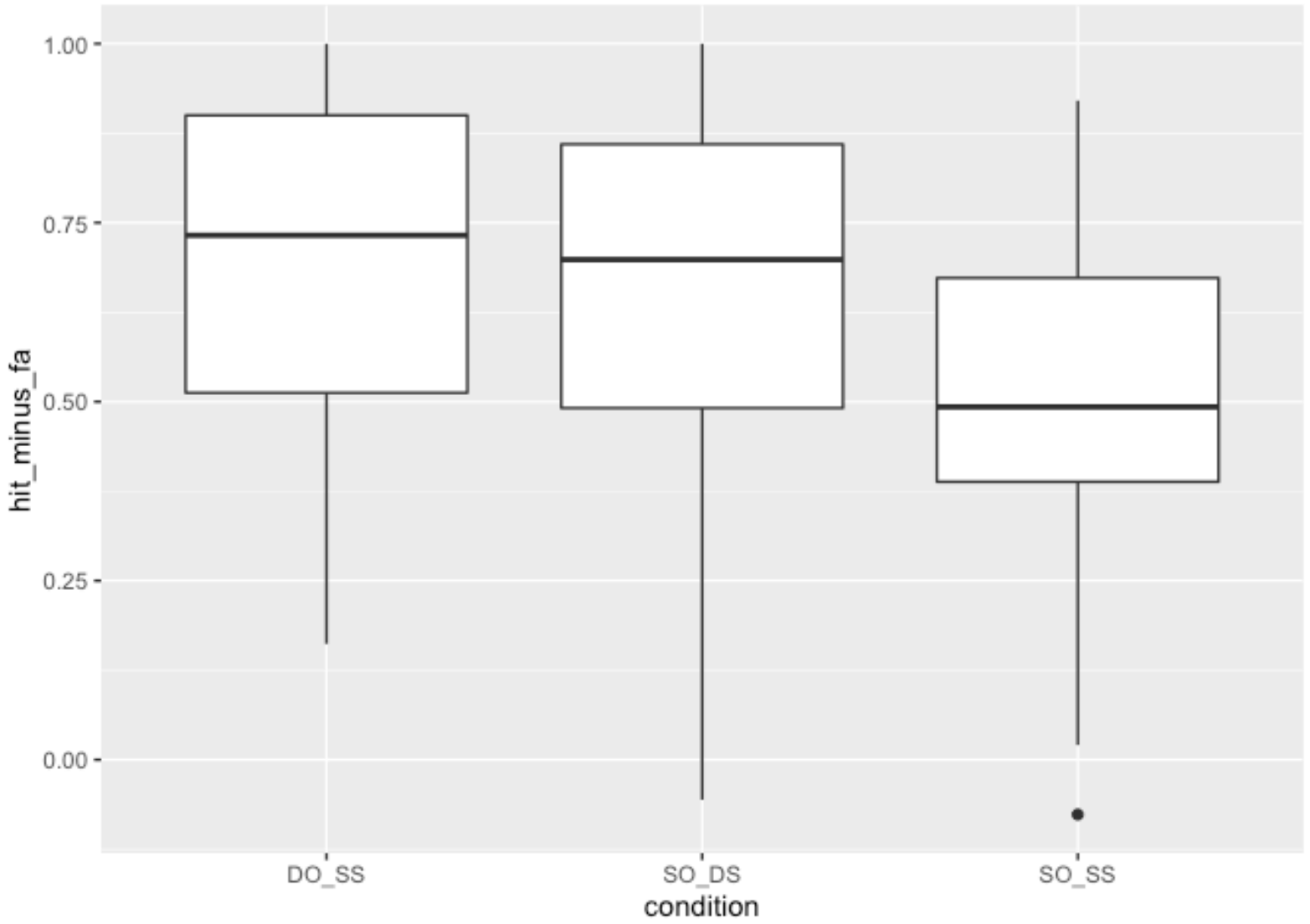
```
## [1] 28
```

## Plot Data

```
#Qplot
qplot(x=condition, y=hit_minus_fa, data=sirdat)
```



```
# Add boxplot
qplot(x=condition, y=hit_minus_fa, data=sirdat, geom='boxplot')
```




## Compute means, standard deviation, standard error, range, median

```
sirdat_summary<- sirdat %>%
  group_by(condition, stimulus, group) %>%
  summarise(mean = mean(hit_minus_fa), sd = sd(hit_minus_fa))

# Compute standard error
sirdat_summary <- sirdat_summary %>%
  mutate(n = ifelse(group == 'YA', nYA, nOA)) %>%
  mutate(se = sd/sqrt(n))

# Range, median
range(sirdat$hit_minus_fa)
```

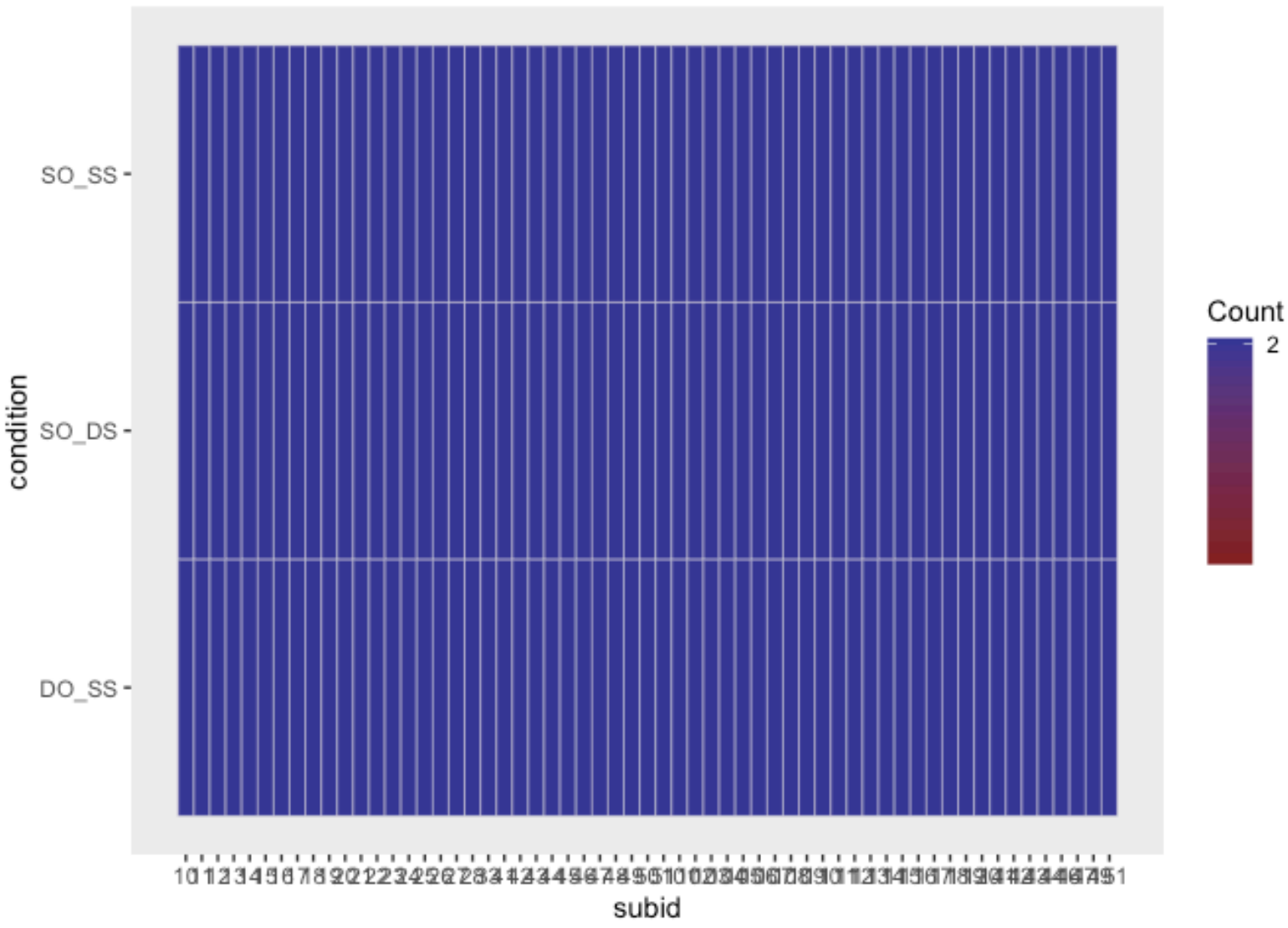
```
## [1] -0.07647  1.00000
```

```
median(sirdat$hit_minus_fa)
```

```
## [1] 0.6485295
```

## Check balance with ezDesign

```
#install.packages('ez')
library(ez)
ezDesign(sirdat, x=subid, y=condition)
```




## ANOVA 2x2x3 (group, stimulus, condition)

```r
# This is something I like to run at the top of my scripts. It bas
set_sum_contrasts()
```

```r
## setting contr.sum globally: options(contrasts=c('contr.sum', 'c
```

```r
# Model 1
mdl1<-aov_ez(id = 'subid',
             dv = 'hit_minus_fa',
             data = sirdat,
             between = 'group',
             within = c('condition', 'stimulus'))

# print the output of the anova
mdl1
```

```
## Anova Table (Type 3 tests)
##
## Response: hit_minus_fa
##                     Effect           df  MSE          F    ges
## 1                    group         1, 57 0.12       0.01 <.0001
## 2                condition 1.95, 111.06 0.02  56.96 ***    .14
## 3          group:condition 1.95, 111.06 0.02       0.94   .003
## 4                 stimulus         1, 57 0.02     3.61 +   .007
## 5           group:stimulus         1, 57 0.02       0.80   .002
## 6       condition:stimulus 1.92, 109.71 0.02 206.11 ***    .36
## 7 group:condition:stimulus 1.92, 109.71 0.02       1.74   .005
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
##
## Sphericity correction method: GG
```

## Post hoc t tests

```r
# You can play around with this and add different variables to cor
```

```
pairs(emmeans(mdl1, c('condition')))
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##  contrast      estimate     SE  df t.ratio p.value
##  DO_SS - SO_DS   0.0317 0.0172 114  1.837  0.1624
##  DO_SS - SO_SS   0.1727 0.0172 114 10.023  <.0001
##  SO_DS - SO_SS   0.1411 0.0172 114  8.187  <.0001
##
## Results are averaged over the levels of: group, stimulus
## P value adjustment: tukey method for comparing a family of 3 es
```

```
pairs(emmeans(mdl1, 'stimulus'))
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##  contrast       estimate     SE df t.ratio p.value
##  object - scene   0.0307 0.0162 57 1.899   0.0626
##
## Results are averaged over the levels of: group, condition
```

```
pairs(emmeans(mdl1, c('condition','stimulus')))
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##  contrast                   estimate     SE  df t.ratio p.valu
##  DO_SS,object - SO_DS,object  0.36942 0.0240 228  15.421 <.0001
##  DO_SS,object - SO_SS,object  0.33650 0.0240 228  14.047 <.0001
##  DO_SS,object - DO_SS,scene   0.36508 0.0251 166  14.532 <.0001
##  DO_SS,object - SO_DS,scene   0.05896 0.0255 207   2.310 0.1945
##  DO_SS,object - SO_SS,scene   0.37408 0.0255 207  14.658 <.0001
##  SO_DS,object - SO_SS,object -0.03293 0.0240 228  -1.374 0.7423
```

```
##  SO_DS,object - DO_SS,scene  -0.00434 0.0255 207  -0.170 1.0000
##  SO_DS,object - SO_DS,scene  -0.31046 0.0251 166 -12.358 <.0001
##  SO_DS,object - SO_SS,scene   0.00466 0.0255 207   0.182 1.0000
##  SO_SS,object - DO_SS,scene   0.02859 0.0255 207   1.120 0.8726
##  SO_SS,object - SO_DS,scene  -0.27753 0.0255 207 -10.875 <.0001
##  SO_SS,object - SO_SS,scene   0.03758 0.0251 166   1.496 0.6673
##  DO_SS,scene - SO_DS,scene   -0.30612 0.0240 228 -12.779 <.0001
##  DO_SS,scene - SO_SS,scene    0.00900 0.0240 228   0.375 0.9990
##  SO_DS,scene - SO_SS,scene    0.31512 0.0240 228  13.154 <.0001
##
## Results are averaged over the levels of: group
## P value adjustment: tukey method for comparing a family of 6 es
```

```
# Bonferroni adjustment
pairs(emmeans(mdl1, c('condition')), adjust='bonferroni')
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##  contrast      estimate     SE  df t.ratio p.value
##  DO_SS - SO_DS   0.0317 0.0172 114  1.837  0.2067
##  DO_SS - SO_SS   0.1727 0.0172 114 10.023  <.0001
##  SO_DS - SO_SS   0.1411 0.0172 114  8.187  <.0001
##
## Results are averaged over the levels of: group, stimulus
## P value adjustment: bonferroni method for 3 tests
```

```
# Can see how the contrasts were coded using coef()
coef(pairs(emmeans(mdl1, c('condition'))))
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##       condition c.1 c.2 c.3
## DO_SS     DO_SS   1   1   0
## SO_DS     SO_DS  -1   0   1
```

```
## SO_SS     SO_SS   0  -1  -1
```

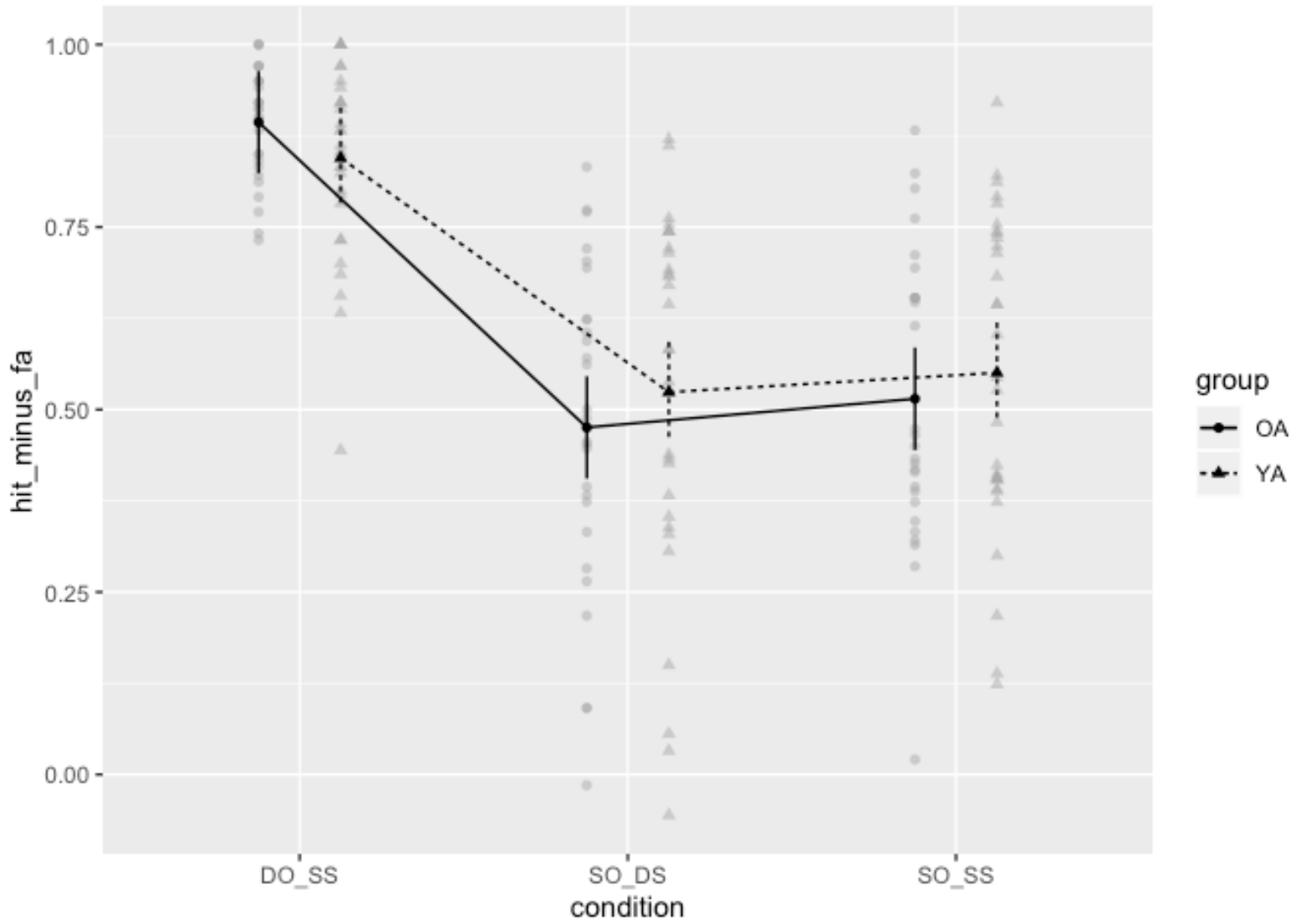
## Object Recognition ANOVA 2x3 (group, condition)

```
# Filter only object data
sirdat_object <- sirdat %>%
  filter(stimulus == 'object')

mdl2<-aov_ez(id = 'subid',
             dv = 'hit_minus_fa',
             data = sirdat_object,
             between = 'group',
             within = c('condition'),
             anova_table= list(es = "pes"))

# print the table
mdl2
```

```
## Anova Table (Type 3 tests)
##
## Response: hit_minus_fa
##           Effect           df  MSE          F  pes p.value
## 1          group        1, 57 0.07       0.08 .001     .77
## 2      condition 1.90, 108.43 0.02 130.14 ***  .70  <.0001
## 3 group:condition 1.90, 108.43 0.02       2.18  .04     .12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
##
## Sphericity correction method: GG
```

```
# Plot effects
afex_plot(mdl2, "condition","group")
```

```
## Warning: Panel(s) show a mixed within-between-design.
## Error bars do not allow comparisons across all means.
## Suppress error bars with: error = "none"
```




## Posthoc t tests

```
pairs(emmeans(md12, c('condition')))
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##  contrast      estimate     SE  df t.ratio p.value
##  DO_SS - SO_DS   0.3694 0.0253 114 14.576  <.0001
##  DO_SS - SO_SS   0.3365 0.0253 114 13.277  <.0001
##  SO_DS - SO_SS  -0.0329 0.0253 114 -1.299  0.3986
##
## Results are averaged over the levels of: group
```

```
## P value adjustment: tukey method for comparing a family of 3 es
```

```
pairs(emmeans(mdl2, c('group')))
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##  contrast estimate     SE df t.ratio p.value
##  OA - YA   -0.0116 0.0402 57 -0.290  0.7728
##
## Results are averaged over the levels of: condition
```

```
pairs(emmeans(mdl2, c('condition','group')))
```

```
##  contrast            estimate     SE  df t.ratio p.value
##  DO_SS,OA - SO_DS,OA  0.41807 0.0367 114 11.378  <.0001
##  DO_SS,OA - SO_SS,OA  0.37878 0.0367 114 10.309  <.0001
##  DO_SS,OA - DO_SS,YA  0.04897 0.0497 117  0.986  0.9217
##  DO_SS,OA - SO_DS,YA  0.36975 0.0497 117  7.441  <.0001
##  DO_SS,OA - SO_SS,YA  0.34318 0.0497 117  6.907  <.0001
##  SO_DS,OA - SO_SS,OA -0.03929 0.0367 114 -1.069  0.8926
##  SO_DS,OA - DO_SS,YA -0.36910 0.0497 117 -7.428  <.0001
##  SO_DS,OA - SO_DS,YA -0.04832 0.0497 117 -0.972  0.9258
##  SO_DS,OA - SO_SS,YA -0.07488 0.0497 117 -1.507  0.6604
##  SO_SS,OA - DO_SS,YA -0.32981 0.0497 117 -6.638  <.0001
##  SO_SS,OA - SO_DS,YA -0.00903 0.0497 117 -0.182  1.0000
##  SO_SS,OA - SO_SS,YA -0.03560 0.0497 117 -0.716  0.9796
##  DO_SS,YA - SO_DS,YA  0.32078 0.0349 114  9.186  <.0001
##  DO_SS,YA - SO_SS,YA  0.29421 0.0349 114  8.426  <.0001
##  SO_DS,YA - SO_SS,YA -0.02657 0.0349 114 -0.761  0.9734
##
## P value adjustment: tukey method for comparing a family of 6 es
```

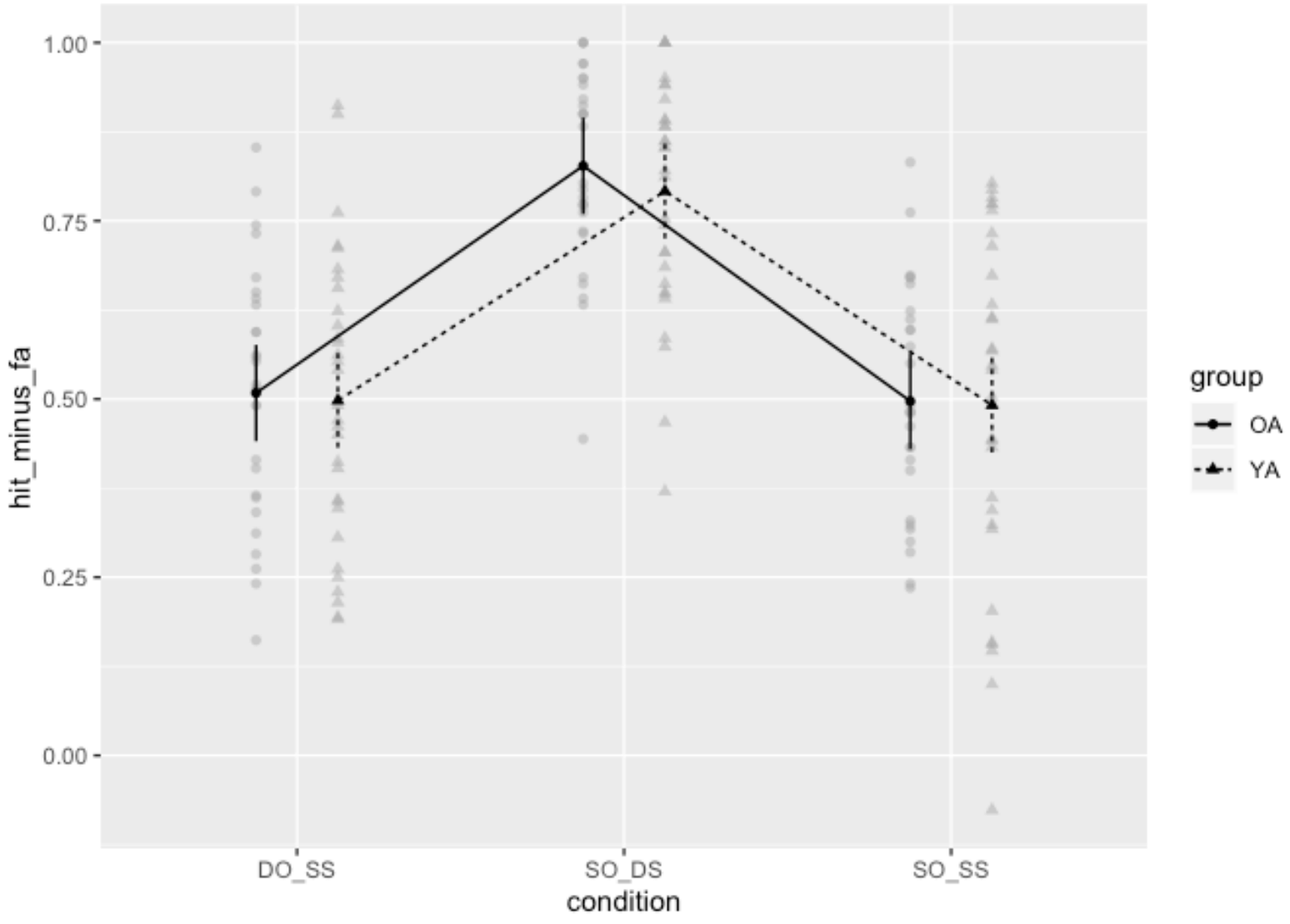
## Scene Recognition ANOVA 2x3 (group, condition)

```r
# Filter only scene data
sirdat_scene <- sirdat %>%
  filter(stimulus == 'scene')

mdl3<-aov_ez(id = 'subid',
             dv = 'hit_minus_fa',
             data = sirdat_scene,
             between = 'group',
             within = c('condition'),
             anova_table= list(es = "pes"))

# print table
mdl3
```

```
## Anova Table (Type 3 tests)
##
## Response: hit_minus_fa
##            Effect           df  MSE         F  pes p.value
## 1           group         1, 57 0.07      0.18 .003     .67
## 2       condition 1.97, 112.15 0.02 127.35 ***  .69  <.0001
## 3 group:condition 1.97, 112.15 0.02      0.25 .004     .77
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
##
## Sphericity correction method: GG
```

```r
# Plot effects
afex_plot(mdl3, "condition","group")
```

```
## Warning: Panel(s) show a mixed within-between-design.
## Error bars do not allow comparisons across all means.
## Suppress error bars with: error = "none"
```




## Posthoc t tests

```
pairs(emmeans(mdl3, c('condition')))
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##  contrast      estimate     SE  df t.ratio p.value
##  DO_SS - SO_DS   -0.306 0.0225 114 -13.617 <.0001
##  DO_SS - SO_SS    0.009 0.0225 114   0.400 0.9156
##  SO_DS - SO_SS    0.315 0.0225 114  14.017 <.0001
##
## Results are averaged over the levels of: group
## P value adjustment: tukey method for comparing a family of 3 es
```

```
pairs(emmeans(mdl3, c('group')))
```

```
## NOTE: Results may be misleading due to involvement in interacti
```

```
##  contrast estimate     SE df t.ratio p.value
##  OA - YA    0.0173 0.0403 57 0.429   0.6698
##
## Results are averaged over the levels of: condition
```

```
pairs(emmeans(mdl3, c('condition','group')))
```

```
##  contrast             estimate     SE  df t.ratio p.value
##  DO_SS,OA - SO_DS,OA -0.318697 0.0326 114 -9.778  <.0001
##  DO_SS,OA - SO_SS,OA  0.011345 0.0326 114  0.348  0.9993
##  DO_SS,OA - DO_SS,YA  0.010442 0.0479 105  0.218  0.9999
##  DO_SS,OA - SO_DS,YA -0.283102 0.0479 105 -5.908  <.0001
##  DO_SS,OA - SO_SS,YA  0.017088 0.0479 105  0.357  0.9992
##  SO_DS,OA - SO_SS,OA  0.330042 0.0326 114 10.127  <.0001
##  SO_DS,OA - DO_SS,YA  0.329140 0.0479 105  6.869  <.0001
##  SO_DS,OA - SO_DS,YA  0.035596 0.0479 105  0.743  0.9760
##  SO_DS,OA - SO_SS,YA  0.335785 0.0479 105  7.008  <.0001
##  SO_SS,OA - DO_SS,YA -0.000902 0.0479 105 -0.019  1.0000
##  SO_SS,OA - SO_DS,YA -0.294446 0.0479 105 -6.145  <.0001
##  SO_SS,OA - SO_SS,YA  0.005743 0.0479 105  0.120  1.0000
##  DO_SS,YA - SO_DS,YA -0.293544 0.0310 114 -9.477  <.0001
##  DO_SS,YA - SO_SS,YA  0.006646 0.0310 114  0.215  0.9999
##  SO_DS,YA - SO_SS,YA  0.300190 0.0310 114  9.691  <.0001
##
## P value adjustment: tukey method for comparing a family of 6 es
```