Class 8: Following up the Two-Way Analysis of Variance

Contents

1 Design Example for Last Class

Kardas & O'Brien (2018) recently reported that with the proliferation of YouTube videos, people seem to think they can learn by *seeing* rather than *doing.* They "hypothesized that the more people merely watch others, the more they believe they can perform the skill themselves." They compared extensively watching the "tablecloth trick" with extensively reading or thinking about it. Participants assessed their own abilities to perform the "tablecloth trick." Participants were asked to rate from 1 (I feel there's no chance at all I'd succeed on this attempt) to 7 (I feel I'd definitely succeed without a doubt on this attempt). The study was a 3 (type of exposure: watch, read, think) $\times$ 2 (amount of exposure: low, high) between-subjects design. They collected an N = 1,003 with small effect sizes ($n^2 < .06$), but to make this analysis doable by hand I will cut the same to 5 people per cell and inflate the effect size (but keep the pattern the same as their experiment 1).

Low Exposure Group	Watching	Reading	Thinking	Row Means
	1	3	3	
	3	2	3	
	3	3	2	
	4	1	3	
	2	4	1	
Cell Means	*2.6*	*2.6*	*2.4*	**2.53**
Cell SS	*5.2*	*5.2*	*3.2*	

High Exposure Group	Watching	Reading	Thinking	Row Means
	6	5	3	
	4	3	2	
	7	2	1	
	5	3	4	
	5	2	4	
Cell Means	*5.4*	*3.0*	*2.8*	**3.73**
Cell SS	*5.2*	*6.0*	*6.8*	
Column Means	**4.0**	**2.8**	**2.6**	G = 3.13

1.1 ANOVA Result

```
library(afex)

Anova.Results<-aov_car(Perceived.Ability~Level*Type + Error(ID),
            data=Example.Data)
Anova.Results
```

```
## Anova Table (Type 3 tests)
##
## Response: Perceived.Ability
##        Effect    df   MSE       F ges p.value
## 1       Level 1, 24 1.32 8.20 ** .25    .009
## 2        Type 2, 24 1.32  4.35 * .27     .02
## 3 Level:Type 2, 24 1.32  3.65 * .23     .04
```

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

2   Follow Up Logic

Just like the one-way ANOVA, the two-way ANOVA tells us which factors are different, but not which levels. The best approach to follow is the Hybrid approach: Do the Confirmatory approach (planned comparisons). Test anything exploratory as conservatively as you can (unplanned comparisons). You need to carefully distinguish what you are doing to your reader.

The challenge of the two-way ANOVA is **unpacking** a significant interaction. All interaction must be unpacked, meaning they must be explained (which cells have driven the effect). This is not always easy the interaction may not always come out as predicted.

2.1   Basic Rules

- **Do not follow up F-tests that were not significant**
- Use error term from ANOVA to do contrasts or protected-t
- Planned: no multiple comparison correction necessary if under 3 tests, but can be applied if you want to be conservative or have multiple planned tests
- Unplanned: apply multiple comparison correction
- Interactions: DO as FEW TESTS as possible to explain the interaction (don't follow it up both ways)

2.2   Hypothesis from Example data

Kardas and O'Brien (2018) hypothesized that spending more time watching YouTube videos would cause someone to think the can do the tablecloth trick. Therefore, from their hypothesis, we would predict that the degree of exposure would impact only the watched condition: the high exposure group to should have a higher mean on *watched* over the low exposure group, but no differences between low and high exposure on reading or thinking about the trick. In sum, they are predicting only 2 cells differ (Low vs High @ Watched). *Note: Non-significant mean no significant difference (cannot reject null). It does not mean the groups are the same. Be careful about talking about non-significant results.*

3
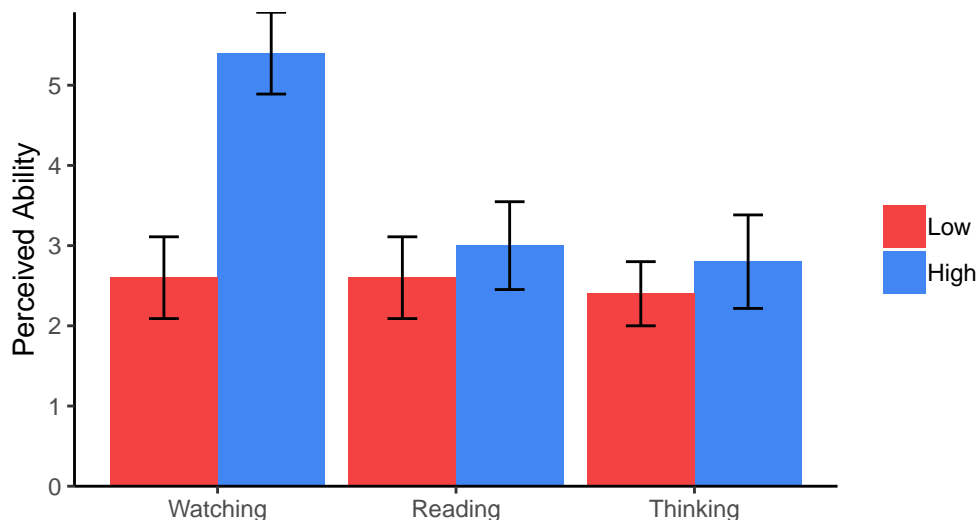
2.2.1 Main Effects vs Interaction

If the interaction is telling a conflicting story with the main effects, you may either not want to follow up the main effects or carefully explain them. You will notice in Kardas & O'Brien (2018) they did not follow up the main effects, only the interaction.

3 Planned Comparisons of Interaction

3.1 Simple Effects with 2-Levels

They want to test the **simple effects** (one factor at the level of another factor) of Level of Exposure at each Type of Exposure. AKA: Low vs. High @ Watching, Low vs. High @ Reading, Low vs. High @ Thinking.

Just like when we followed up the one-way ANOVA, we will use the df and error term from omnibus ANOVA (df_W & MS_W). To do this, we will use emmeans package. Note: These can be reported as F-tests (as basically, we are doing one-way ANOVAs) or as t-values.

3.1.1 Fit Model

First, we must **cut** the ANOVA we calculated the way we want to test the results:

```
library(emmeans)
Simple.Effects.By.Type<-emmeans(Anova.Results, ~Level|Type)
Simple.Effects.By.Type
```

```
## Type = Watching:
##  Level emmean        SE df lower.CL upper.CL
##  Low      2.6 0.5131601 24  1.54089  3.65911
##  High     5.4 0.5131601 24  4.34089  6.45911
##
## Type = Reading:
##  Level emmean        SE df lower.CL upper.CL
##  Low      2.6 0.5131601 24  1.54089  3.65911
##  High     3.0 0.5131601 24  1.94089  4.05911
##
## Type = Thinking:
##  Level emmean        SE df lower.CL upper.CL
##  Low      2.4 0.5131601 24  1.34089  3.45911
##  High     2.8 0.5131601 24  1.74089  3.85911
##
## Confidence level used: 0.95
```

3.1.2 As t-values (pairwise approach)

You see above; we have 3 subsets of the results. Now we can calculate the F-values, and we can use the `pairs` function. This will run 3 LSD tests (protected t-tests) using the DF and error terms from the two-way ANOVA automatically.

```
pairs(Simple.Effects.By.Type,adjust='none')
```

```
## Type = Watching:
##  contrast   estimate        SE df t.ratio p.value
```

```
##  Low - High     -2.8 0.725718 24  -3.858  0.0008
##
## Type = Reading:
##  contrast   estimate       SE df t.ratio p.value
##  Low - High     -0.4 0.725718 24  -0.551  0.5866
##
## Type = Thinking:
##  contrast   estimate       SE df t.ratio p.value
##  Low - High     -0.4 0.725718 24  -0.551  0.5866
```

Note: These are seen 3 families of 1 test. So even if you try to apply FWER you cannot. I will show you how to force that when we get to unplanned follow-up tests.

3.1.3 As t-values (contrast approach)

You can also call a contrast code and get the same exact results. Note: this will be useful as we add more levels.

```
Set1 <- list(
  H1 = c(-1,1))

contrast(Simple.Effects.By.Type,Set1,adjust='none')
```

```
## Type = Watching:
##  contrast estimate       SE df t.ratio p.value
##  H1            2.8 0.725718 24   3.858  0.0008
##
## Type = Reading:
##  contrast estimate       SE df t.ratio p.value
##  H1            0.4 0.725718 24   0.551  0.5866
##
## Type = Thinking:
##  contrast estimate       SE df t.ratio p.value
##  H1            0.4 0.725718 24   0.551  0.5866
```

3.1.4 As F-tests (one-way ANOVAs)

In this case the F-tests = squared(t-values). All we have to do is wrap the `pairs` or `contrast` in the `test` function (with `joint = TRUE` argument). This will run 3 one-way ANOVAs using the DF and error terms from the two-way ANOVA automatically.

```
test(pairs(Simple.Effects.By.Type), joint = TRUE)
```

```
##  Type      df1 df2 F.ratio p.value
##  Watching   1  24  14.886  0.0008
##  Reading    1  24   0.304  0.5866
##  Thinking   1  24   0.304  0.5866
```

3.1.5 Which to report?

F or t values? When you use the words "simple effects" or "contrasts" people are primed to see F-values (and people generally don't ask for FWER corrections). When you pairwise approach, they think you are doing unprotected t-tests and want to see FWER.

3.2 Following up the interaction the other way

First and foremost, ONLY UNPACK THE INTERACTION WITH ONE FAMILY OF SIMPLE EFFECTS. Do not think you now also have to test it the other way: Watching vs. Reading vs. Thinking @ Low, Watching vs. Reading vs. Thinking @ High. You will notice in their paper they did not do this as they did not hypothesize these type of follow-ups. Also, they already Unpacked their ANOVA. They know what caused the interaction. Any additional tests are merely a waste of time and will inflate FWER. Now, might you have a hypothesis that was very complex and required you to follow it both ways? Yes, you would limit which simple effects you would examine to those you need to see to understand the effects.

3.3 Simple Effects with 3-Levels

They did not hypothesize following up the other way, but we will do it anyway. The problem we have is that we have three level: Watching vs Reading vs Thinking @ Low, Watching vs Reading vs Thinking @ High. In the old school approach, we first have to run an F-test (one-way ANOVA) at each level of Level of Exposure and run pairwise or contrasts to follow up if the One-way was significant.

3.3.1 Fit Model

First, we must **cut** the ANOVA we calculated the way we want to test the results:

```
Simple.Effects.By.Level<-emmeans(Anova.Results, ~Type|Level)
Simple.Effects.By.Level
```

```
## Level = Low:
##  Type      emmean        SE df lower.CL upper.CL
##  Watching     2.6 0.5131601 24  1.54089  3.65911
##  Reading      2.6 0.5131601 24  1.54089  3.65911
##  Thinking     2.4 0.5131601 24  1.34089  3.45911
##
## Level = High:
##  Type      emmean        SE df lower.CL upper.CL
##  Watching     5.4 0.5131601 24  4.34089  6.45911
##  Reading      3.0 0.5131601 24  1.94089  4.05911
##  Thinking     2.8 0.5131601 24  1.74089  3.85911
##
## Confidence level used: 0.95
```

3.3.2 As F-tests (one-way ANOVAs)

Just like above, we run our one-way ANOVA using the df and error term from omnibus ANOVA (df_W & MS_W). (ignore any messages about linear dependence).

```
test(pairs(Simple.Effects.By.Level), joint = TRUE)
```

```
##  Level df1 df2 F.ratio p.value note
##  Low     2  24   0.051  0.9507  d
##  High    2  24   7.949  0.0022  d
##
## d: df1 reduced due to linear dependence
```

3.3.3 Pairwise follow up

The code will run these are 2 families of 3 tests. You need to ignore the results of the low exposure group.

3.3.3.1 Fisher LSD

```
pairs(Simple.Effects.By.Level,adjust='none')
```

```
## Level = Low:
##  contrast                estimate       SE df t.ratio p.value
##  Watching - Reading  -1.054712e-15 0.725718 24   0.000  1.0000
##  Watching - Thinking  2.000000e-01 0.725718 24   0.276  0.7852
##  Reading - Thinking   2.000000e-01 0.725718 24   0.276  0.7852
##
## Level = High:
##  contrast                estimate       SE df t.ratio p.value
##  Watching - Reading   2.400000e+00 0.725718 24   3.307  0.0030
##  Watching - Thinking  2.600000e+00 0.725718 24   3.583  0.0015
##  Reading - Thinking   2.000000e-01 0.725718 24   0.276  0.7852
```

3.3.3.2 Tukey HSD

- It will correct for 3 tests per family

```
pairs(Simple.Effects.By.Level,adjust='tukey')
```

```
## Level = Low:
##  contrast                estimate       SE df t.ratio p.value
##  Watching - Reading  -1.054712e-15 0.725718 24   0.000  1.0000
##  Watching - Thinking  2.000000e-01 0.725718 24   0.276  0.9591
##  Reading - Thinking   2.000000e-01 0.725718 24   0.276  0.9591
##
## Level = High:
##  contrast                estimate       SE df t.ratio p.value
##  Watching - Reading   2.400000e+00 0.725718 24   3.307  0.0080
##  Watching - Thinking  2.600000e+00 0.725718 24   3.583  0.0041
##  Reading - Thinking   2.000000e-01 0.725718 24   0.276  0.9591
##
## P value adjustment: tukey method for comparing a family of 3 estimates
```

3.3.4 Contrasts

3.3.4.1 "Simple" Linear Contrast

This is indentical to LSD results.

```
Set2 <- list(
  H1 = c(-1,1,0),
  H2 = c(-1,0,1),
  H3 = c(0,-1,1))

contrast(Simple.Effects.By.Level,Set2,adjust='none')
```

```
## Level = Low:
```

```
## contrast        estimate       SE df t.ratio p.value
## H1         1.054712e-15 0.725718 24   0.000  1.0000
## H2        -2.000000e-01 0.725718 24  -0.276  0.7852
## H3        -2.000000e-01 0.725718 24  -0.276  0.7852
##
## Level = High:
## contrast        estimate       SE df t.ratio p.value
## H1        -2.400000e+00 0.725718 24  -3.307  0.0030
## H2        -2.600000e+00 0.725718 24  -3.583  0.0015
## H3        -2.000000e-01 0.725718 24  -0.276  0.7852
```

This coding system will only let you Bonferroni, Sidak, or FDR. If you type, Tukey the package will change it to Sidak. It will also correct for 3 tests within each family.

```
contrast(Simple.Effects.By.Level,Set2,adjust='sidak')
```

```
## Level = Low:
## contrast        estimate       SE df t.ratio p.value
## H1         1.054712e-15 0.725718 24   0.000  1.0000
## H2        -2.000000e-01 0.725718 24  -0.276  0.9901
## H3        -2.000000e-01 0.725718 24  -0.276  0.9901
##
## Level = High:
## contrast        estimate       SE df t.ratio p.value
## H1        -2.400000e+00 0.725718 24  -3.307  0.0089
## H2        -2.600000e+00 0.725718 24  -3.583  0.0045
## H3        -2.000000e-01 0.725718 24  -0.276  0.9901
##
## P value adjustment: sidak method for 3 tests
```

If you want a Tukey-like correction, you must use 'mvt' this is a Monte-Carlo corrected pvalue based on multivariate normal t-test distribution. Note this is a simulation and can be very slow with large amounts of data.

```
contrast(Simple.Effects.By.Level,Set2,adjust='mvt')
```

```
## Level = Low:
## contrast        estimate       SE df t.ratio p.value
## H1         1.054712e-15 0.725718 24   0.000  1.0000
## H2        -2.000000e-01 0.725718 24  -0.276  0.9982
## H3        -2.000000e-01 0.725718 24  -0.276  0.9982
##
## Level = High:
## contrast        estimate       SE df t.ratio p.value
## H1        -2.400000e+00 0.725718 24  -3.307  0.0158
## H2        -2.600000e+00 0.725718 24  -3.583  0.0082
## H3        -2.000000e-01 0.725718 24  -0.276  0.9982
##
## P value adjustment: mvt method for 3 tests
```

3.3.4.2 "Complex" Linear Contrast

Here you can merge and compare groups as before. You can apply corrections if you lots of tests, but they will be within the family: Bonferroni, Sidak, mvt, or FDR.

```
Set3 <- list(
  WvsRT = c(2,-1,-1))
```

```
contrast(Simple.Effects.By.Level,Set3,adjust='none')
```

```
## Level = Low:
##  contrast estimate       SE df t.ratio p.value
##  WvsRT         0.2 1.256981 24   0.159  0.8749
##
## Level = High:
##  contrast estimate       SE df t.ratio p.value
##  WvsRT         5.0 1.256981 24   3.978  0.0006
```

3.3.4.3 "Consecutive" Contrast

If the data were ordinal, you could test them in some kind of logical order. You can apply corrections if you lots of tests, but they will be within the family: Bonferroni, Sidak, mvt, or FDR.

```
Consec.Set <- list(
  WvsR = c(-1,1,0),
  RvsT = c(0,1,0))

contrast(Simple.Effects.By.Level,Consec.Set,adjust='none')
# Also
# contrast(Simple.Effects.By.Level,'consec',adjust='none')
```

```
## Level = Low:
##  contrast      estimate        SE df t.ratio p.value
##  WvsR     1.054712e-15 0.7257180 24   0.000  1.0000
##  RvsT     2.600000e+00 0.5131601 24   5.067  <.0001
##
## Level = High:
##  contrast      estimate        SE df t.ratio p.value
##  WvsR    -2.400000e+00 0.7257180 24  -3.307  0.0030
##  RvsT     3.000000e+00 0.5131601 24   5.846  <.0001
```

3.3.4.4 "Polynomial" Contrast

If the data were ordinal, you could test to see if they have a linear slope or curvilinear slope. You can apply corrections if you lots of tests, but they will be within the family: Bonferroni, Sidak, mvt, or FDR.

```
Poly.Set <- list(
  Linear = c(-1,0,1),
  Quad = c(-1,2,-1))

contrast(Simple.Effects.By.Level,Poly.Set,adjust='none')
# Also
# contrast(Simple.Effects.By.Level,'poly',adjust='none')
```

```
## Level = Low:
##  contrast estimate       SE df t.ratio p.value
##  Linear       -0.2 0.725718 24  -0.276  0.7852
##  Quad          0.2 1.256981 24   0.159  0.8749
##
## Level = High:
##  contrast estimate       SE df t.ratio p.value
```

```
## Linear        -2.6 0.725718 24  -3.583  0.0015
## Quad          -2.2 1.256981 24  -1.750  0.0929
```

4 Main-effect Follow up

Remember, we would not follow this one up because of the hypothesis we are testing, but you may want to follow up the main effects in other cases. If this is planned, go with less stringent correction. If its unplanned use a more conservative correction.

Code wise it the same as we did for one-way ANOVA, here we can only follow up the Type of Exposure. Again we will use the df and error term from the two-way interaction.

```
Main.Effects.Type<-emmeans(Anova.Results, ~Type)
Main.Effects.Type
```

```
## Type       emmean       SE df lower.CL upper.CL
## Watching      4.0 0.362859 24 3.251096 4.748904
## Reading       2.8 0.362859 24 2.051096 3.548904
## Thinking      2.6 0.362859 24 1.851096 3.348904
##
## Results are averaged over the levels of: Level
## Confidence level used: 0.95
```

You can run the pairwise or any of the contrast approaches I showed you for interactions.

```
pairs(Main.Effects.Type,adjust='mvt')
```

```
## contrast            estimate       SE df t.ratio p.value
## Watching - Reading       1.2 0.5131601 24   2.338  0.0694
## Watching - Thinking      1.4 0.5131601 24   2.728  0.0302
## Reading - Thinking       0.2 0.5131601 24   0.390  0.9200
##
## Results are averaged over the levels of: Level
## P value adjustment: mvt method for 3 tests
```

5 Unplanned Comparisons

This is all the same code as above, but you should use conservative corrections. There are also ways to force R to recognize there are more tests within the family but we will not cover that today.

6 References

Kardas, M., & O'Brien, E. (2018). Easier Seen Than Done: Merely Watching Others Perform Can Foster an Illusion of Skill Acquisition. *Psychological science*, 29(4), 521-536.