Assignment 3—Factorial ANOVA.

Yudhajit Ain

Student ID: 30182745

University of Calgary

PSYC 615—Lab

TAs: Benjamin Moon & Chris Davie

Assignment 3—Factorial ANOVA.

# Question 1

*For this assignment you will be designing, “collecting,” and analyzing a two-way factorial ANOVA.*

## Question 1a

*Using your research area, design an experimental 3x2 between-subjects ANOVA. State your research question, IVs, DV, and hypothesis; please also state what makes your design an experimental design. Your hypothesis must explicitly predict a significant interaction and a significant simple main effect at one level of an IV but not the other level(s) of that IV (5 marks). Hint: Visualizing how you want your results to look (with a graph or table) might help you figure out how to create your data so that it meets the hypothesis.*

*Answer*

Let us consider that we are interested in testing the functional double dissociation (mutual independence) between two brain networks - the default mode network (DMN), and the dorsal attention network (DAN) with respect to two cognitive functions they have been shown to subserve respectively - internal attention, and external attention.

To test this hypothesis, we recruit patients with lesions in either the DMN, the DAN, or non-lesion healthy controls. These three groups are levels of our first independent variable - **Lesion**. Within each group, participants are randomly assigned to performing only one of two cognitive attention tasks, one probing internal attention, and one probing external attention. These are the two levels of our second independent variable - **Task**. Our dependent variable for each participant will be their trial-averaged reaction time scores (RT) for each task. The fact that patients will be chosen for inclusion in the study based on localization of lesions within either the DMN or DAN only, and then they will be randomly assigned to either of the two types of Task, makes this an experimental design

Theoretically, we predict there will be a significant interaction between Task and Lesion variables, such that the RT for DMN-lesioned patients will be greater for the internal attention task compared to the external attention task. However, we expect external attention impairments for DAN-lesioned patients to uniformly increase the RT for both internal and external attention tasks since both tasks require top-down attention to external task stimuli. The experimental tasks were piloted for equal difficulty across tasks, so we expect RT to be the same for both tasks for healthy control participants.

## Question 1b

*Create a dataset that will successfully confirm your hypothesis. Upload a file with your dataset (saved as a .csv file) when you submit the assignment. Your data set must have 6 unique groups (i.e., cells; 3x2 design) with a total sample size of 150 (3 marks).*

between=list(  
 Lesion = c("DMN","DAN","Control"),  
 Task = c("Internal","External")  
)  
  
mu=list(  
 DMN\_Internal = 150,  
 DMN\_External = 100,  
 DAN\_Internal = 148,  
 DAN\_External = 150,  
 Control\_Internal = 99,  
 Control\_External = 101  
)  
  
st\_dev = 6  
  
set.seed(12)  
lesion\_task <- faux::sim\_design(  
 between = between,  
 n = 25,  
 mu = mu,  
 sd = st\_dev,  
 empirical = TRUE,  
 dv = "reaction\_time",  
 plot = FALSE  
)  
  
lesion\_task=tibble::as\_tibble(lesion\_task)  
write.csv(lesion\_task,here("assignments","assignment-3","Q1b.csv"))

*Answer*

The dataset is saved as Q1b.csv

## Question 1c

*Perform the analyses on your data and write an APA formatted results section (10 marks). This includes:*

* *A quick report on your assumptions (you are creating this data, so you can make sure it doesn’t violate assumptions!). Don’t worry about addressing independence. Please report tests of normality and homogeneity of variance (you do not need to include anything on skew, kurtosis, Q-Q plots, etc.) (2 marks)*
* *Reporting the omnibus ANOVA (2 marks)*
* *Reporting simple main effects and any follow-ups to simple main effects (2 marks)*
* *Reporting any corrections for multiple comparisons that you used (1 mark)*
* *Including and referencing an APA formatted figure depicting your interaction (2 marks)*
* *Hint: When reporting findings, you need to include practical significance! (1 mark)*

shapiro.test(dplyr::filter(.data=lesion\_task,Lesion=="DMN" & Task=="Internal")$reaction\_time)  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: dplyr::filter(.data = lesion\_task, Lesion == "DMN" & Task == "Internal")$reaction\_time  
#> W = 0.97427, p-value = 0.7535  
shapiro.test(dplyr::filter(.data=lesion\_task,Lesion=="DMN" & Task=="External")$reaction\_time)  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: dplyr::filter(.data = lesion\_task, Lesion == "DMN" & Task == "External")$reaction\_time  
#> W = 0.96516, p-value = 0.5264  
shapiro.test(dplyr::filter(.data=lesion\_task,Lesion=="DAN" & Task=="Internal")$reaction\_time)  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: dplyr::filter(.data = lesion\_task, Lesion == "DAN" & Task == "Internal")$reaction\_time  
#> W = 0.96781, p-value = 0.5903  
shapiro.test(dplyr::filter(.data=lesion\_task,Lesion=="DAN" & Task=="External")$reaction\_time)  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: dplyr::filter(.data = lesion\_task, Lesion == "DAN" & Task == "External")$reaction\_time  
#> W = 0.95222, p-value = 0.2812  
shapiro.test(dplyr::filter(.data=lesion\_task,Lesion=="Control" & Task=="Internal")$reaction\_time)  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: dplyr::filter(.data = lesion\_task, Lesion == "Control" & Task == "Internal")$reaction\_time  
#> W = 0.96053, p-value = 0.4252  
shapiro.test(dplyr::filter(.data=lesion\_task,Lesion=="Control" & Task=="External")$reaction\_time)  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: dplyr::filter(.data = lesion\_task, Lesion == "Control" & Task == "External")$reaction\_time  
#> W = 0.96077, p-value = 0.43  
  
car::leveneTest(reaction\_time ~ Lesion \* Task,center = mean,data = lesion\_task)  
#> Levene's Test for Homogeneity of Variance (center = mean)  
#> Df F value Pr(>F)  
#> group 5 0.087 0.9942  
#> 144  
  
lesion\_task\_model=aov(reaction\_time ~ Lesion \* Task, data = lesion\_task)  
broom::tidy(lesion\_task\_model)  
#> # A tibble: 4 × 6  
#> term df sumsq meansq statistic p.value  
#> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
#> 1 Lesion 2 60033. 30017. 834. 6.63e-80  
#> 2 Task 1 8817. 8817. 245. 7.14e-33  
#> 3 Lesion:Task 2 22533. 11267. 313. 3.78e-53  
#> 4 Residuals 144 5184. 36.0 NA NA  
#lesion\_task\_model  
effectsize::omega\_squared(lesion\_task\_model)  
#> # Effect Size for ANOVA (Type I)  
#>   
#> Parameter | Omega2 (partial) | 95% CI  
#> ---------------------------------------------  
#> Lesion | 0.92 | [0.90, 1.00]  
#> Task | 0.62 | [0.54, 1.00]  
#> Lesion:Task | 0.81 | [0.76, 1.00]  
#>   
#> - One-sided CIs: upper bound fixed at [1.00].  
  
emmeans::joint\_tests(lesion\_task\_model,by="Lesion")  
#> Lesion = DMN:  
#> model term df1 df2 F.ratio p.value  
#> Task 1 144 868.056 <.0001  
#>   
#> Lesion = DAN:  
#> model term df1 df2 F.ratio p.value  
#> Task 1 144 1.389 0.2405  
#>   
#> Lesion = Control:  
#> model term df1 df2 F.ratio p.value  
#> Task 1 144 1.389 0.2405  
  
lesion\_task\_emm = emmeans::emmeans(lesion\_task\_model,specs="Task",by="Lesion")  
  
emmeans::eff\_size(lesion\_task\_emm,sigma=sigma(lesion\_task\_model),edf=df.residual(lesion\_task\_model))  
#> Lesion = DMN:  
#> contrast effect.size SE df lower.CL upper.CL  
#> Internal - External 8.333 0.567 144 7.213 9.453  
#>   
#> Lesion = DAN:  
#> contrast effect.size SE df lower.CL upper.CL  
#> Internal - External -0.333 0.284 144 -0.894 0.227  
#>   
#> Lesion = Control:  
#> contrast effect.size SE df lower.CL upper.CL  
#> Internal - External -0.333 0.284 144 -0.894 0.227  
#>   
#> sigma used for effect sizes: 6   
#> Confidence level used: 0.95

# Results

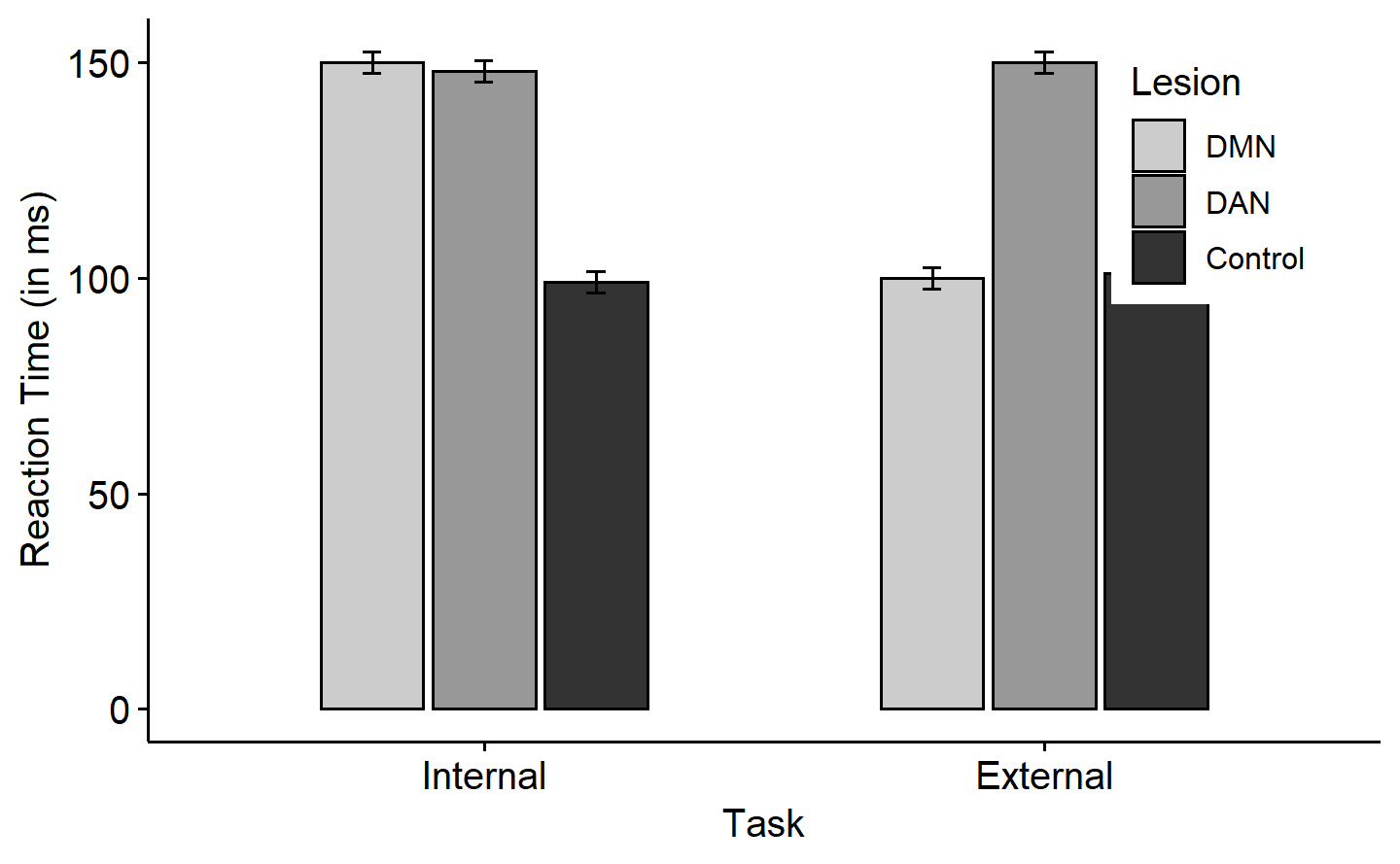
Before running an omnibus ANOVA, we first check for validity of the assumptions of within-group normality and between-group homogeneity of variance. The Shapiro-Wilk’s test confirms that all groups are normally distributed: DMN-lesion,Internal-task , DMN-lesion,External-task , DAN-lesion,Internal-task , DAN-lesion,External-task , Control,Internal-task , and Control,External-task . The Levene’s test was non-significant , confirming homogeneity of variance across all groups.

Finally, we run the omnibus ANOVA, and it reveals a significant main effect of both Lesion and Task on reaction time, as well as a significant interaction between Lesion and Task . Given that there is a significant interaction, we do not interpret the main effects directly, rather proceeding to perform tests for simple main effects of Task, at individual levels of Lesion. Since there are 3 levels of Lesion for which the test for simple main effect of Task is repeated, each test is evaluated at an -level of , and significance of simple main effect is inferred only if the observed -value is less than . This keeps the family-wise Type-I error rate at .

As predicted, only for patients with DMN lesions, reaction times for Internal attention tasks were significantly higher than the reaction times for External attention tasks , . For patients with DAN lesions, there was no significant difference in reaction times between Internal attention and External attention tasks, . There was also as no significant difference in reaction times between Internal attention and External attention tasks in healthy control participants .

The observed simple main effects (and interaction) is illustrated in figure 1.

source(here::here("R", "apa\_legend.R"))  
  
BarP = ggpubr::ggbarplot(data=lesion\_task,xlab="Task",x="Task",ylab="Reaction Time (in ms)",y="reaction\_time",add="mean\_ci",fill="Lesion",palette = "grey", position = position\_dodge(0.6),width=0.55)  
  
BarP + apa\_legend(position = "topright")



*Figure* *1.*  Mean reaction times (in ms) for all conditions. Error bars represent 95% CI around mean

## Question 1d

*State why you chose to investigate the specific simple main effects that you reported in your results section (i.e., why you chose to hold that IV constant as opposed to the other). State one advantage or disadvantage of your analytic decision, compared to if you had alternatively decided to hold the other IV constant. (2 marks)*

*Answer*

Three reasons behind why I hold the levels of Lesion constant and compare between Task conditions, and not the other way round, are:

1. This helps me test whether the behavioral efficiency differs across the two types of Tasks (internal vs external attention) due to damage in one particular brain network (DMN, or DAN, or none) at a time. This approach directly relates to my research hypothesis, and helps test for the relative necessity of an intact DMN (and DAN, separately) for internal vs external attention tasks.
2. The Lesion groups are not randomly assigned, and so this is only a quasi-experimental study. The Task groups however, are randomly assigned, and so comparison between Task groups will be free from any confounds (physiological etc.) dependent on membership to Lesion group (which cannot be balanced by randomisation of assignment).
3. Comparing across the levels of Task for simple main effects does not require post-hoc tests because there are only two levels of Task, and the most stringent -level used is , as opposed to three post-hoc pairwise comparisons for each of two simple main effects the other way round, setting the -level for all pairwise comparisons at , which has much less power than the former option.

# Question 2

*State why you would typically want to pool variance for your follow-up tests. Give an example of when you would not want to pool variance (i.e., provide a reason – you do not need to come up with a scenario). (2 marks)*

*Answer*

Using pooled variance (across groups) for follow-up tests gives a more accurate estimate of the variance. Moreover, using pooled variance for the comparisons makes the denominator degrees of freedom larger, and can often lead to greater power for our post-hoc tests (which might already be underpowered by corrections/adjustments).

If the assumption of homogeneity of variances across groups is not met, we should not be pooling the variances, since now group membership itself determines the within-group variance estimate, and any comparisons using pooled variances in such a scenario may significantly inflate the Type-I or Type-II error rate.

# Bonus

*If you had a factorial ANOVA with two IVs, and both main effects plus the interaction were significant, how would you proceed with the analysis and why? (0.5 BONUS marks, not to exceed 100%)*

*Answer*

If the interaction is significant, we would not interpret the main effects, since a significant interaction already tells us that the effect of one IV (IV1) depends on the level of the other IV (IV2), and a main effect of either IV by themselves collapses across levels of the other IV, and is therefore simply not informative/specific enough about where actually the group differences are between levels of that IV, and how they depend on levels of the other IV.

If there is an interaction, therefore, we choose IV1 to hold constant at one of its levels, and then look at the simple main effect of IV2 at that particular level of IV1. If the simple main effect of IV2 at this levels of IV1 is significant, we proceed to do follow-up tests to determine which levels of the IV2 differ from each other, at that particular level of IV1. This entire process is iteratively repeated to test for the simple main effect of IV2 (and follow-ups, if significant) at each individual level of IV1.

In order to control for the infaltion of Type-I error rates by the multiple comparisons we make, in this process, appropriate corrections/adjustments are made to the tests depending on factors like whether they are a-priori vs post-hoc hypotheses, or whether the comparisons involve simple vs complex contrasts.

# All R Output

mget(ls())  
#> $BarP  
#>   
#> $between  
#> $between$Lesion  
#> [1] "DMN" "DAN" "Control"  
#>   
#> $between$Task  
#> [1] "Internal" "External"  
#>   
#>   
#> $lesion\_task  
#> # A tibble: 150 × 4  
#> id Lesion Task reaction\_time  
#> <chr> <fct> <fct> <dbl>  
#> 1 S001 DMN Internal 142.  
#> 2 S002 DMN Internal 161.  
#> 3 S003 DMN Internal 145.  
#> 4 S004 DMN Internal 145.  
#> 5 S005 DMN Internal 139.  
#> 6 S006 DMN Internal 149.  
#> 7 S007 DMN Internal 149.  
#> 8 S008 DMN Internal 147.  
#> 9 S009 DMN Internal 151.  
#> 10 S010 DMN Internal 154.  
#> # … with 140 more rows  
#>   
#> $lesion\_task\_emm  
#> Lesion = DMN:  
#> Task emmean SE df lower.CL upper.CL  
#> Internal 150 1.2 144 147.6 152  
#> External 100 1.2 144 97.6 102  
#>   
#> Lesion = DAN:  
#> Task emmean SE df lower.CL upper.CL  
#> Internal 148 1.2 144 145.6 150  
#> External 150 1.2 144 147.6 152  
#>   
#> Lesion = Control:  
#> Task emmean SE df lower.CL upper.CL  
#> Internal 99 1.2 144 96.6 101  
#> External 101 1.2 144 98.6 103  
#>   
#> Confidence level used: 0.95   
#>   
#> $lesion\_task\_model  
#> Call:  
#> aov(formula = reaction\_time ~ Lesion \* Task, data = lesion\_task)  
#>   
#> Terms:  
#> Lesion Task Lesion:Task Residuals  
#> Sum of Squares 60033.33 8816.67 22533.33 5184.00  
#> Deg. of Freedom 2 1 2 144  
#>   
#> Residual standard error: 6  
#> Estimated effects may be unbalanced  
#>   
#> $mu  
#> $mu$DMN\_Internal  
#> [1] 150  
#>   
#> $mu$DMN\_External  
#> [1] 100  
#>   
#> $mu$DAN\_Internal  
#> [1] 148  
#>   
#> $mu$DAN\_External  
#> [1] 150  
#>   
#> $mu$Control\_Internal  
#> [1] 99  
#>   
#> $mu$Control\_External  
#> [1] 101  
#>   
#>   
#> $params  
#> $params$firstname  
#> [1] "Yudhajit"  
#>   
#> $params$lastname  
#> [1] "Ain"  
#>   
#> $params$studentid  
#> [1] 30182745  
#>   
#> $params$TAs  
#> [1] "Benjamin Moon & Chris Davie"  
#>   
#> $params$assignment  
#> [1] 3  
#>   
#> $params$show\_output  
#> [1] TRUE  
#>   
#>   
#> $st\_dev  
#> [1] 6

