PSYCH-UH 1004Q: Statistics for Psychology

Class 23: Factorial ANOVA: The main effects

Prof. Jon Sprouse
Psychology

Our example experiment

Let's use our course satisfaction theory as an example to better understand 2x2 designs.

DV		IV1		IV2
satisfaction with a course meeting time	←	time of the course	×	type of person
(1-7 scale)		(morning, afternoon)		(morning lark, night owl)

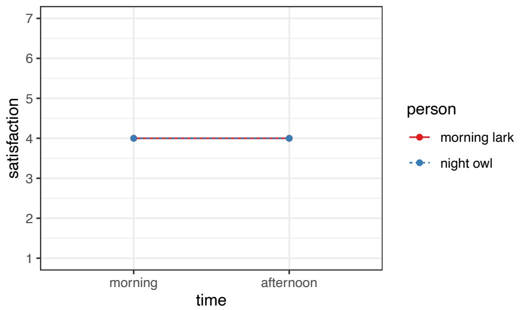
If we were to create an experiment to test this theory, how many groups (conditions) would we need?

	time	person
condition/group 1	morning	morning lark
condition/group 2	morning	night owl
condition/group 3	afternoon	morning lark
condition/group 4	afternoon	night owl

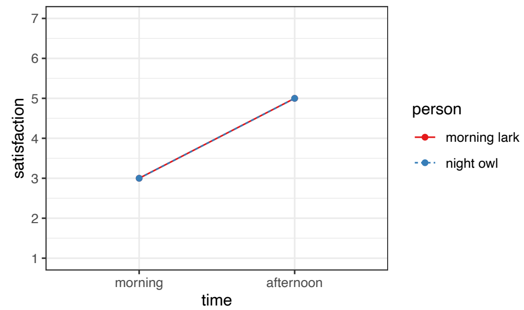
Defining Main effects

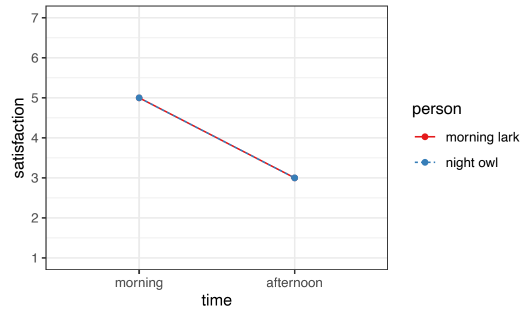
Let’s start with main effects

No effects

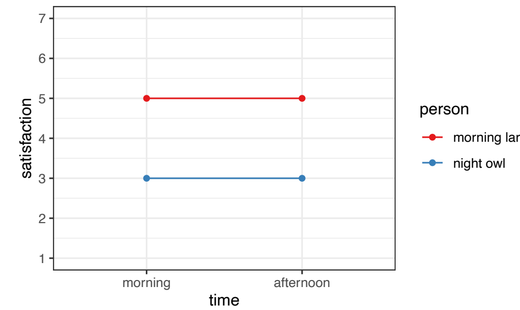


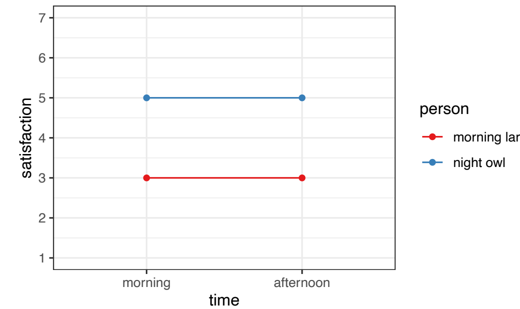
Main effect of time



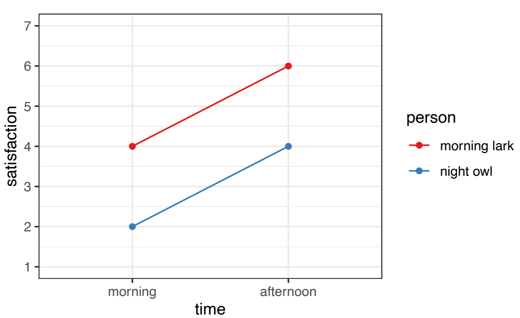


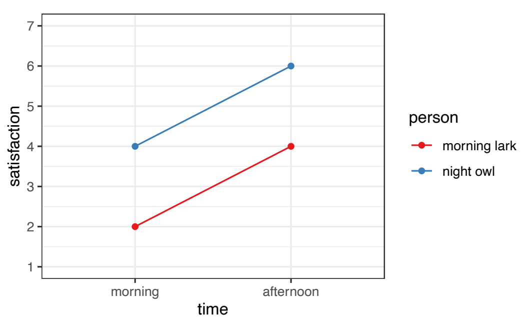
Main effect of person

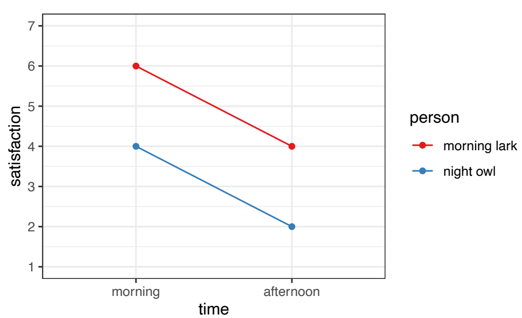


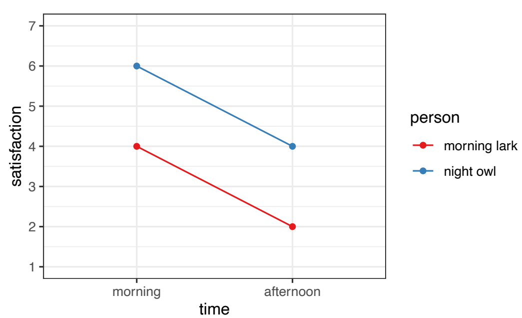


Main effect of time and main effect of person









What is a main effect?

A **main effect** of a factor is the effect of that factor when you **ignore** all other factors. In other words, it is the result you would expect to obtain if you only included that one factor in your experimental design!

So, the main effect of TIME is the effect you'd get if you ran this experiment:

DV		IV1
satisfaction with a course meeting time (1-7 scale)	←	time of the course (morning, afternoon)

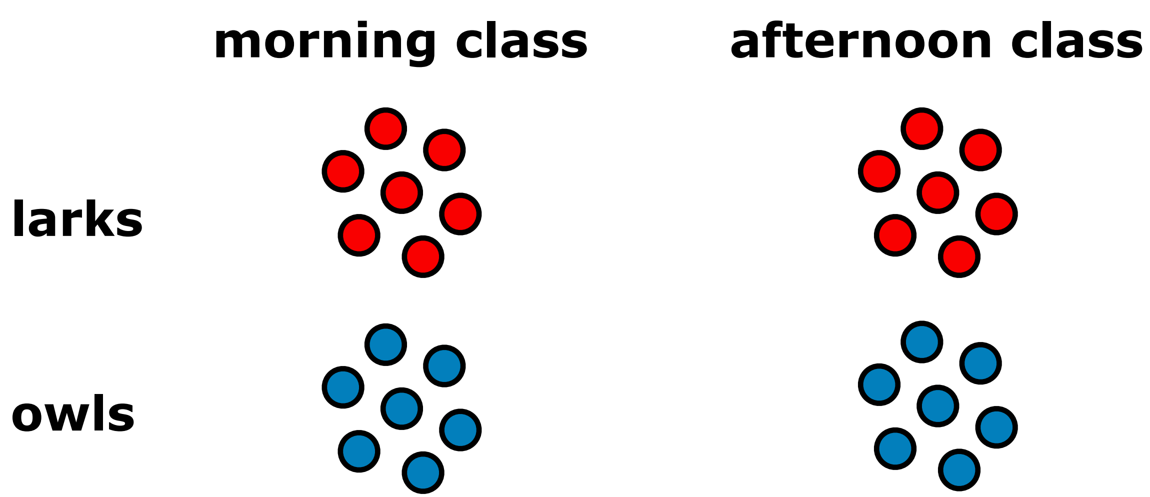
The main effect of PERSON is the effect you'd get if you ran this experiment:

DV		IV2
satisfaction with a course meeting time (1-7 scale)	←	type of person (morning lark, night owl)

We find the main effect by averaging

To find a main effect of a factor, we average together all of the conditions that share the same level of that factor.

To see why averaging has the consequence of “ignoring” the other factor, think about our 4 groups of people in our experiment:

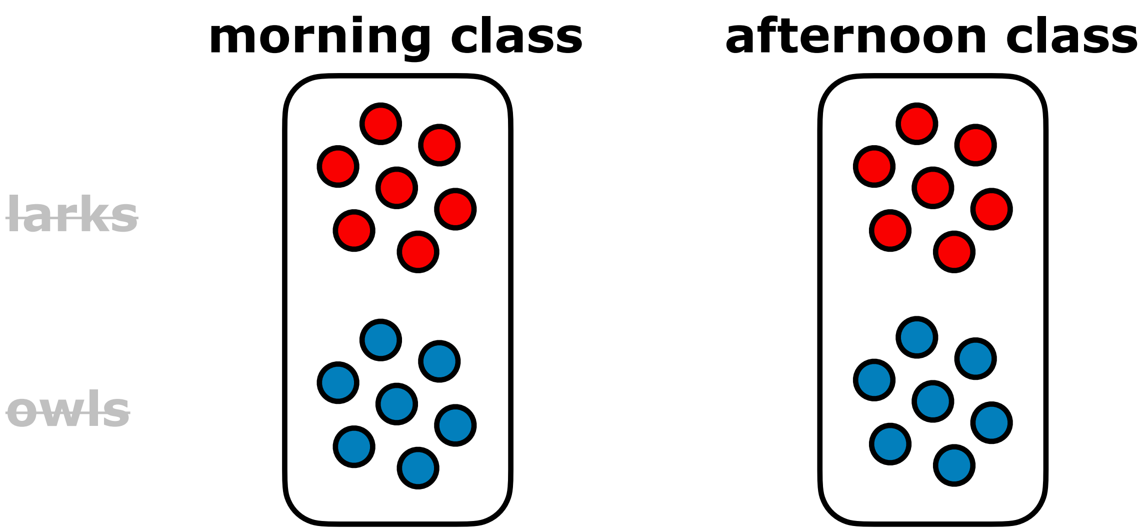


Arranging them in a grid (table!) like this makes it really easy to think about averaging - we can average across the rows or across the columns. Let’s see what each average does!

The main effect of TIME

To find a **main effect of TIME**, we average together the two morning groups and the two afternoon groups.

Notice that this <u>ignores the distinction of person</u>. It is as if we ran our experiment with two conditions: morning and afternoon (and ignored person).

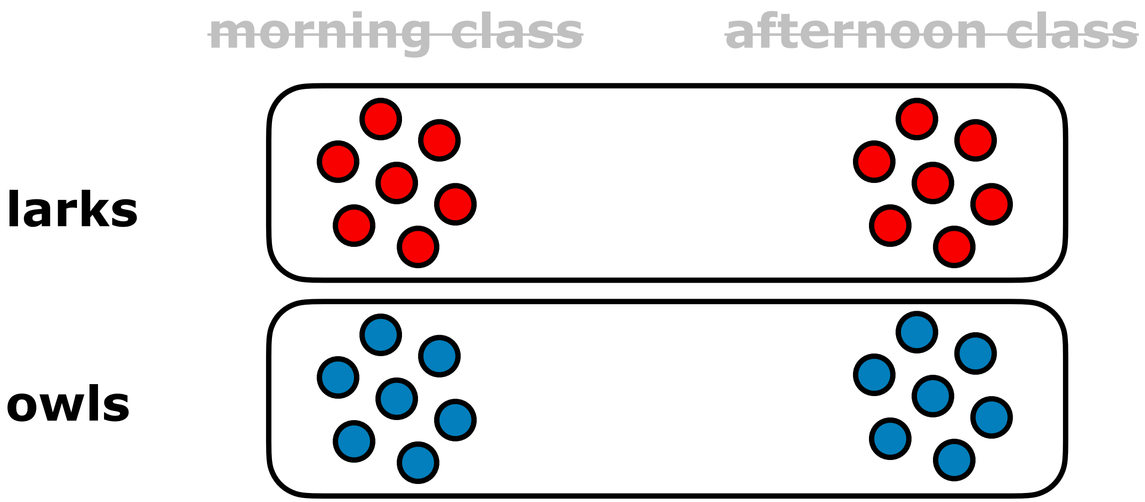


<u>DV</u>		<u>IV1</u>		<u>IV2</u>
satisfaction with a course meeting time	←	time of the course (morning, afternoon)	×	type of person (morning lark, night owl)

The main effect of PERSON

To find a **main effect of PERSON**, we average together the two lark groups and the two owl groups.

Notice that this ignores the distinction of time. It is as if we ran our experiment with two conditions: larks and owls (and ignored time).

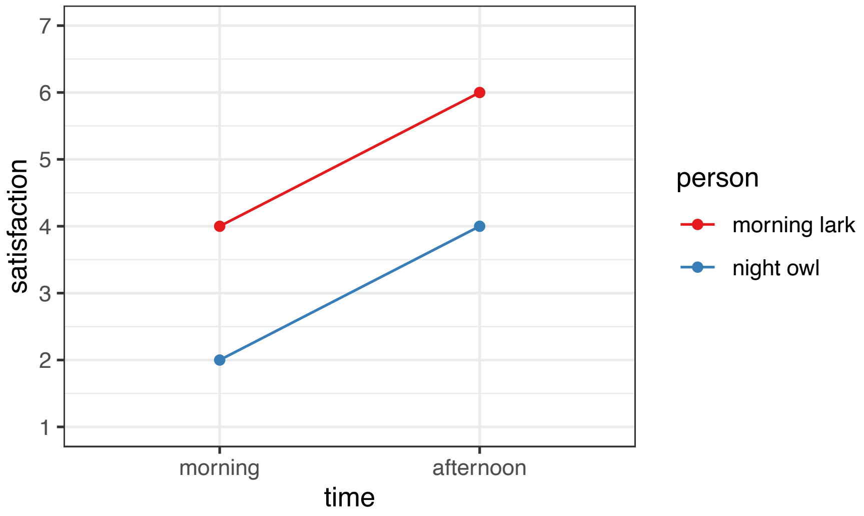


DV		IV1		IV2
satisfaction with a course meeting time	←	time of the course (morning, afternoon)	×	type of person (morning lark, night owl)

A table with means

It can be useful to put your condition means in a table, and then calculate the level means from that table. These show you the main effects.

For example, let's imagine that these are the results of the experiment. This would yield the table to the right:



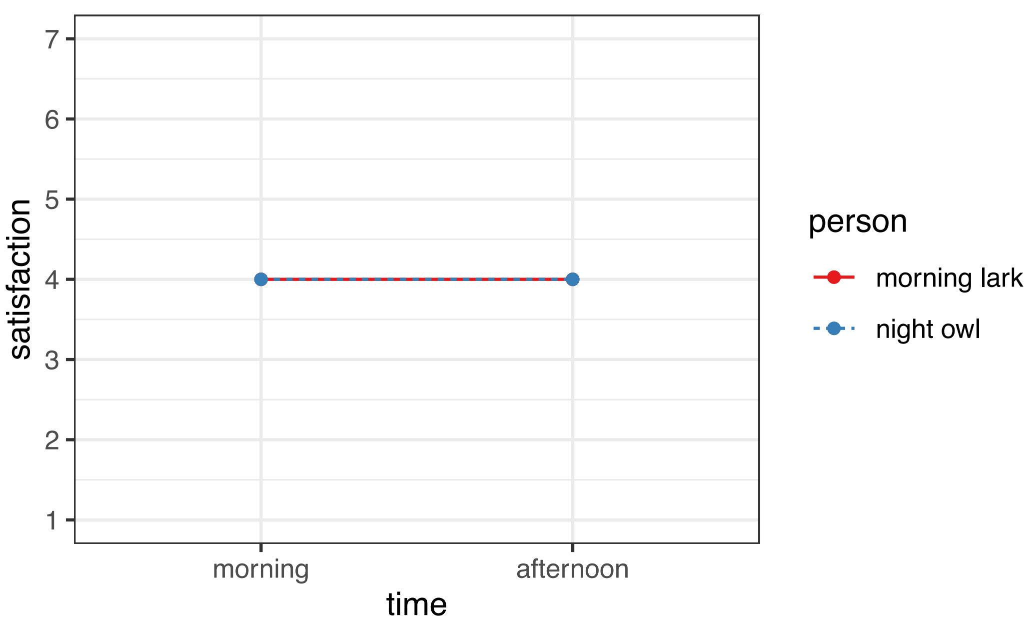
	morning	afternoon	mean
lark	4	6	5
owl	2	4	3
mean	3	5	grand mean: 4

Take a moment to make sure you can see how the "marginal means" (means in the margins) work!

Isolating main effects

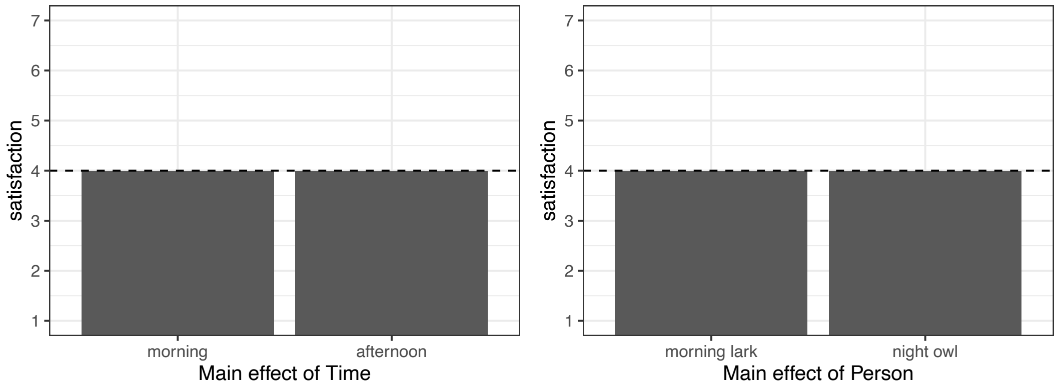
No main effects

When there is no effect, both the plot and table are pretty boring. Just the same score all the way through.



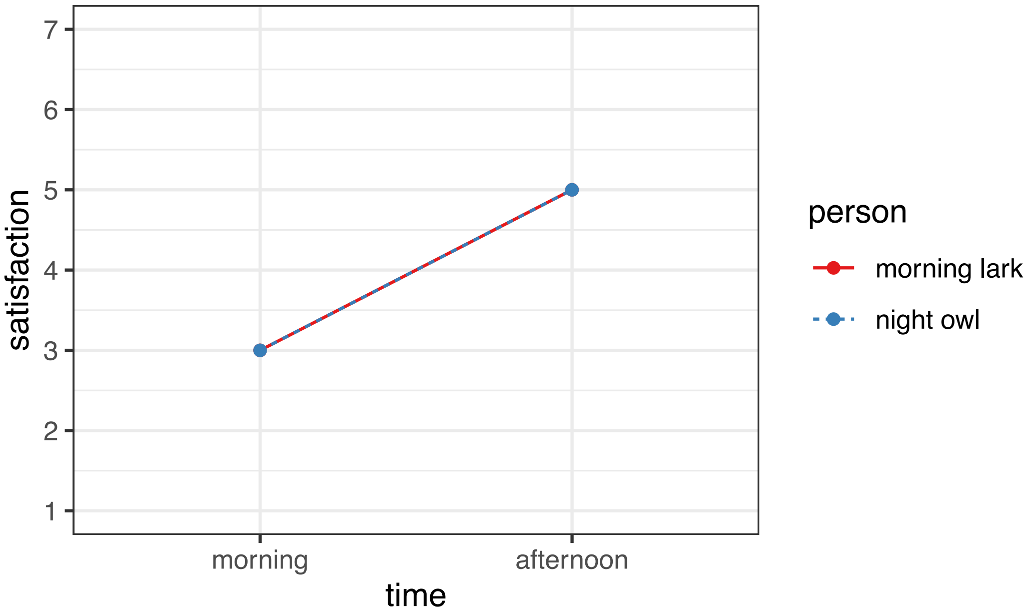
	morning	afternoon	mean
lark	4	4	4
owl	4	4	4
mean	4	4	grand mean: 4

We can plot those means if we want to see the main effects visually. We can add a line for the grand mean for comparison. This is not a plot you would normally make, but it is helpful to see the effects.



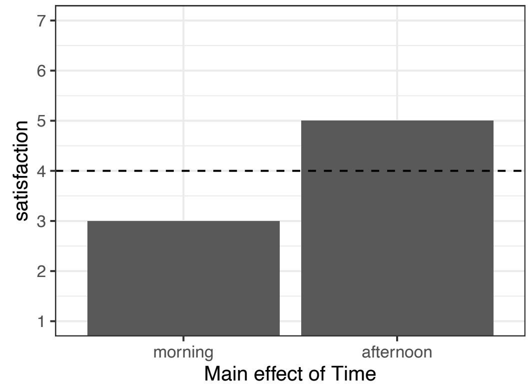
A main effect of TIME, but not PERSON

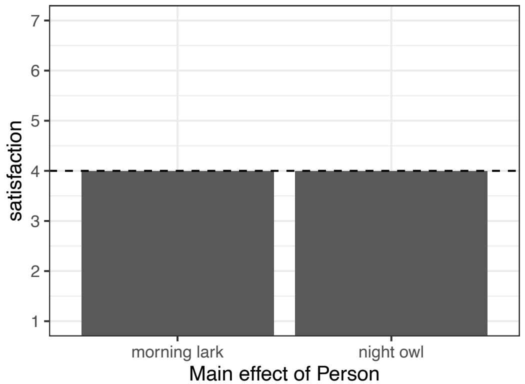
For a main effect of TIME but not PERSON, we see a difference in the means between morning and afternoon, but no difference between lark and owl.



	morning	afternoon	mean
lark	3	5	4
owl	3	5	4
mean	3	5	grand mean: 4

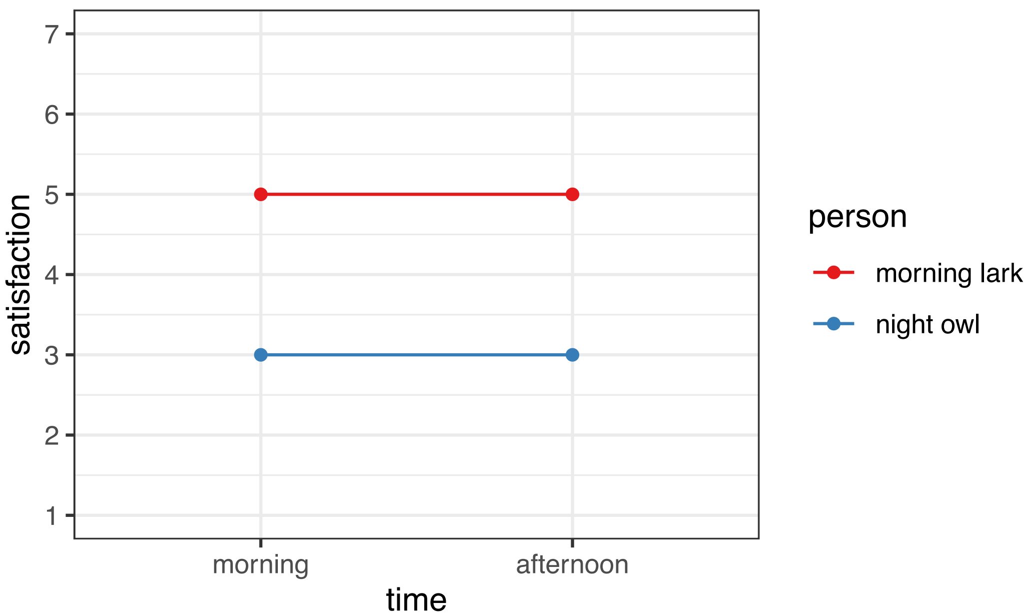
We can plot those means if we want to see the main effects visually. We can add a line for the grand mean for comparison. This is not a plot you would normally make, but it is helpful to see the effects.





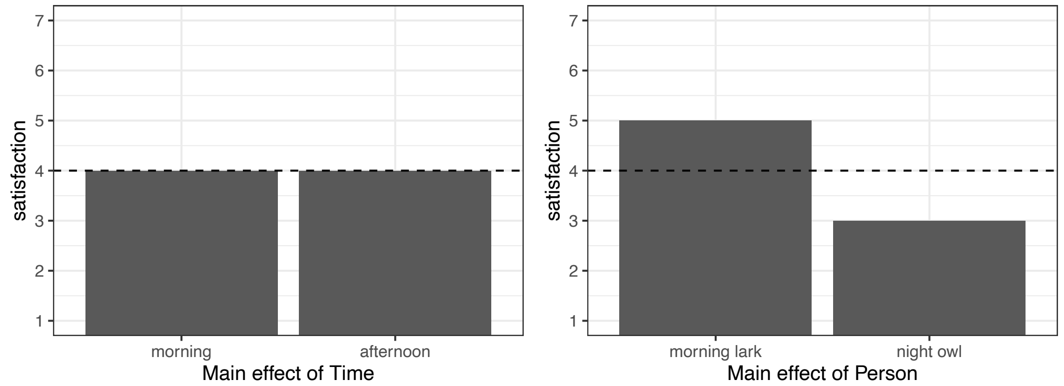
A main effect of PERSON, but not TIME

For a main effect of PERSON but not TIME, we see a difference in the means between larks and owls, but no difference between morning and afternoon.



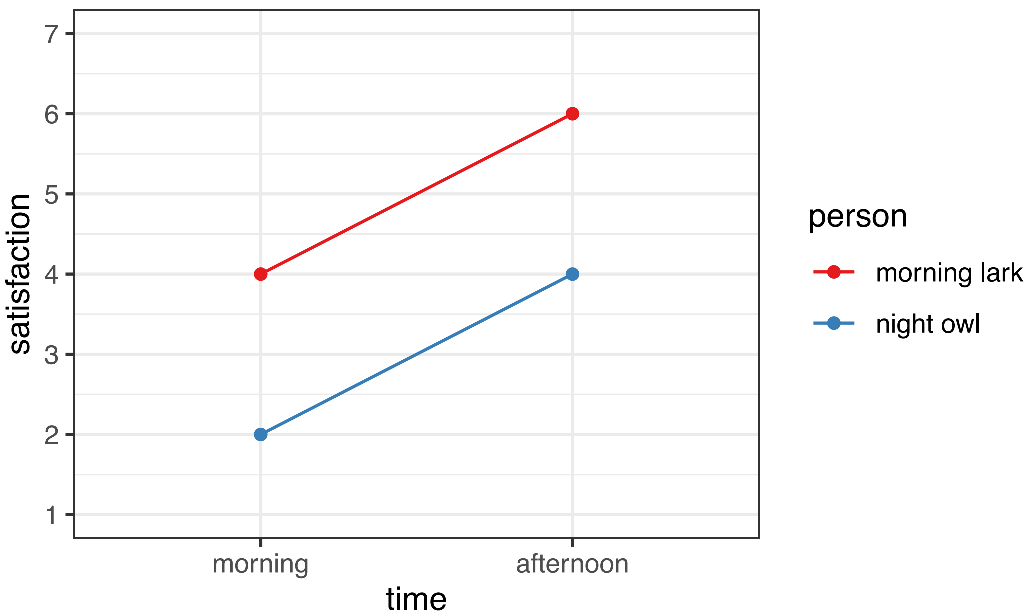
	morning	afternoon	mean
lark	5	5	5
owl	3	3	3
mean	4	4	grand mean: 4

We can plot those means if we want to see the main effects visually. We can add a line for the grand mean for comparison. This is not a plot you would normally make, but it is helpful to see the effects.



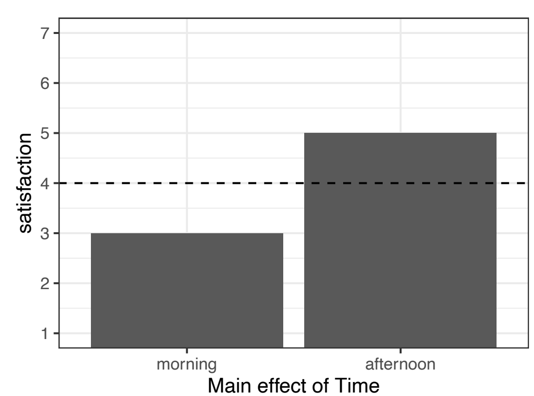
Main effects of TIME and PERSON

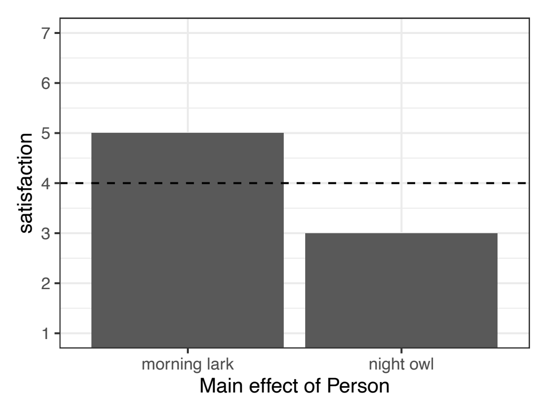
When both main effects are present, we see differences in both pairs of means: a difference between morning and afternoon and a difference between lark and owl.



	morning	afternoon	mean
lark	4	6	5
owl	2	4	3
mean	3	5	grand mean: 4

We can plot those means if we want to see the main effects visually. We can add a line for the grand mean for comparison. This is not a plot you would normally make, but it is helpful to see the effects.





Calculating the main effects in a factorial ANOVA

The general logic of factorial ANOVA

The general logic of factorial ANOVA is that there is a separate F-ratio for each of the effects. In a two-way ANOVA, there are three effects:

Main effect of factor 1: $F_{factor1} = \dfrac{MS_{B1}}{MS_W}$

Main effect of factor 2: $F_{factor2} = \dfrac{MS_{B2}}{MS_W}$

The interaction of factor1 x factor2: $F_{inter} = \dfrac{MS_{inter}}{MS_W}$

So we need to calculate 4 terms: MS_{B1}, MS_{B2}, MS_{inter}, MS_W

Calculating MS_W

MS_W is exactly the same for factorial ANOVA as it is for one-way ANOVA. You simply pool the variance for the number of groups in your experiment:

$$MS_W = \frac{\sum (n_i - 1)\, s_i^2}{n_{total} - k}$$

You already know this equation: n_{total} is the sum of the sample sizes for all of the groups, and k is the number of groups. (And notice that we already have a different n for each group, so this equation works for both equal and unequal sample sizes.)

You will use the same MS_W for all of the *F*-ratios in your factorial ANOVA, so that is very nice, as it means less work for each of the effects.

Calculating MS_{B1} and MS_{B2}

You also already know how to calculate the between-group variances for each of the main effects. It is the MS_B equation that we already know!

$$MS_B: \quad n\frac{\sum(\bar{x}_i - \bar{x}_G)^2}{k-1}$$

For MS_{B1}, the group means ($\bar{x}_i$) are the result of averaging to see the main effect of factor 1. Then calculate their differences from the grand mean.

For MS_{B2}, the group means ($\bar{x}_i$) are the result of averaging to see the main effect of factor 2. Then calculate their differences from the grand mean.

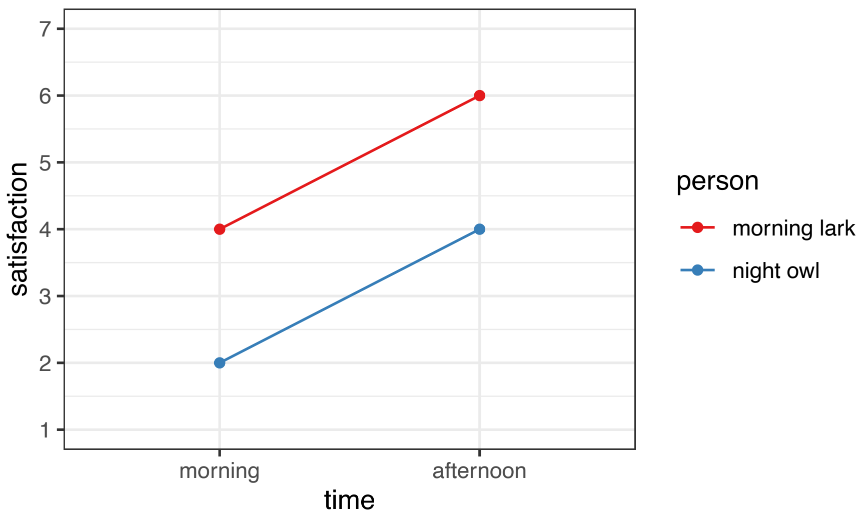
This is exactly the same as calculating an ANOVA for a two-condition experiment.

All that has changed is that we are doing it twice, and we are doing it on averages across the conditions that show us the main effects of our factors.

	morning	afternoon	mean
lark	4	6	5
owl	2	4	3
mean	3	5	grand mean: 4

Let's do a mini-example for main effects

Let's imagine that our experiment gave us a result that looks like two main effects and no interaction.



		morning	afternoon	mean
lark	raw	3,4,5	5,6,7	
	mean	4	6	5
	s^2	1	1	
owl	raw	1,2,3	3,4,5	
	mean	2	4	3
	s^2	1	1	
mean		3	5	grand mean: 4

The first thing we can do is calculate the descriptive statistics for each of our groups and the full experiment. We can calculate each of the means, each of the variances, and the grand mean (the mean of the means). Notice that the n_{total} for this experiment is 12, with n=3 in each group.

MS_W

I suggest always starting with MS_W. You need it for all of the effects. And it is usually straightforward to calculate.

$$MS_W = \frac{\sum (n_i - 1)\, s_i^2}{n_{total} - k}$$

$$\frac{(3-1)1 + (3-1)1 + (3-1)1 + (3-1)1}{12-4}$$

$MS_W = 1$

		morning	afternoon	mean
lark	raw	3,4,5	5,6,7	
	mean	4	6	5
	s^2	1	1	
owl	raw	1,2,3	3,4,5	
	mean	2	4	3
	s^2	1	1	
mean		3	5	grand mean: 4

Notice that for designs where the groups have an equal number of participants, this is also just the mean of the variances. (This is because the weights are equal.)

MS_B for the main effect of TIME

Next, let's calculate the MS_B for the main effect of TIME.

MS_B: $$n\frac{\sum(\bar{x}_i - \bar{x}_G)^2}{k-1}$$

$$\frac{(6)(3-4)^2 + (6)(5-4)^2}{2-1}$$

$MS_B = 12$

		morning	afternoon	mean
lark	raw	3,4,5	5,6,7	
	mean	4	6	5
	s^2	1	1	
owl	raw	1,2,3	3,4,5	
	mean	2	4	3
	s^2	1	1	
mean		3	5	grand mean: 4

There are two tricky things to remember here. The first is that the n for each group is now 6. Why is that? Because there are 3 in each of the four groups, and we are combining two groups together (3+3=6) to form the new groups for the main effect. The second is that k is the number of levels in the factor we are working on (so 2 here); it is not the total k for the experiment.

MS_B for the main effect of PERSON

Next, let's calculate the MS_B for the main effect of PERSON.

MS_B: $$n\frac{\sum(\bar{x}_i - \bar{x}_G)^2}{k-1}$$

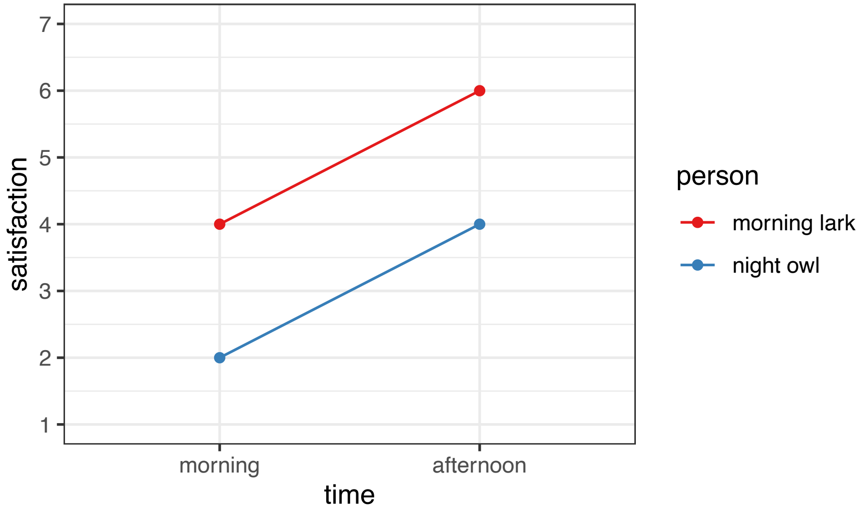
$$\frac{(6)(5-4)^2 + (6)(3-4)^2}{2-1}$$

$MS_B = 12$

		morning	afternoon	mean
lark	raw	3,4,5	5,6,7	
	mean	4	6	5
	s^2	1	1	
owl	raw	1,2,3	3,4,5	
	mean	2	4	3
	s^2	1	1	
mean		3	5	grand mean: 4

There are two tricky things to remember here. The first is that the n for each group is now 6. Why is that? Because there are 3 in each of the four groups, and we are combining two groups together (3+3=6) to form the new groups for the main effect. The second is that k is the number of levels in the factor we are working on (so 2 here); it is not the total k for the experiment.

Now we can see the *F*s for our two main effects



		morning	afternoon	mean
lark	raw	3,4,5	5,6,7	
	mean	4	6	5
	s^2	1	1	
owl	raw	1,2,3	3,4,5	
	mean	2	4	3
	s^2	1	1	
mean		3	5	grand mean: 4

Main effect of TIME: $F_{time} = \frac{MS_{B1}}{MS_W} = \frac{12}{1} = 12$

Main effect of PERSON: $F_{person} = \frac{MS_{B2}}{MS_W} = \frac{12}{1} = 12$

Calculating the *p*-values for each *F*

The final step is to calculate the *p*-values for each of the main effects. These are *F*-ratios, so you need to look up the *p*-value based on the two degrees of freedom df_B and df_W for each of them.

Luckily, the degrees of freedom for the main effects follow the same logic as one-ANOVA. The df_B is the number of groups (k) minus 1. And the df_W is the total n minus the total number of groups.

Main effect of TIME: $F_{time} = 12$

$df_B = k\text{-}1 = 2\text{-}1 = 1$

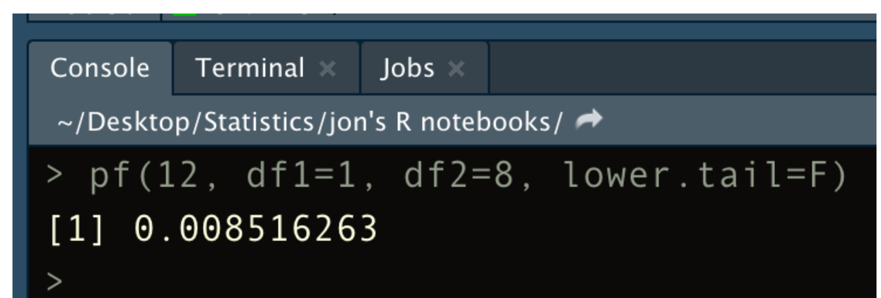
$df_W = n_{total}\text{-}k = 12\text{-}4 = 8$

Main effect of PERSON: $F_{person} = 12$

$df_B = k\text{-}1 = 2\text{-}1 = 1$

$df_W = n_{total}\text{-}k = 12\text{-}4 = 8$

Then we can use the function pf() in R to look up the *p*-value.



A factorial ANOVA table

We can also create an ANOVA table for a factorial design:

	df	SS	MS	*F*	*p*
between-cells	3	24			
time	1	12	12	12	.0085
person	1	12	12	12	.0085
interaction	1	0	0	0	1
within-cells	8	8	1		
total	11	32			

As before, R's is very similar to the human one from the book. The two differences are no totals, and within is called "residuals".

```
Console   Terminal ×   Jobs ×
~/Desktop/Statistics/jon's R notebooks/
> m=aov(satisfaction~time*person, data=data)
> summary(m)
            Df Sum Sq Mean Sq F value  Pr(>F)
time         1     12      12      12 0.00852 **
person       1     12      12      12 0.00852 **
time:person  1      0       0       0 1.00000
Residuals    8      8       1
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

Next time: Interactions!