MSc Applied Information and Data Science

HSLU Lucerne University of Applied Sciences and Arts

Data Analytics for Data Scientists

Design of Experiments (DoE)

Lecture 09: Factorial designs

2025

Prof. Dr. Jürg Schwarz

Program: 16:15 until 17:55

16:15	**Begin of the lesson**
	Lecture: Jürg Schwarz ◦ Design of experiments – A look back ◦ Full factorial designs with control ◦ Fractional factorial designs ◦ Preview of Lecture 10
	Tutorial: Students / Jürg Schwarz / Assistants ◦ Working on the exercise ◦ Support by Jürg Schwarz / Assistants
17:55	**End of the lesson**

Detailed table of contents → Slide 30

Designs of experiments – A look back

Factorial designs

Review of Lecture 03

Slide 3

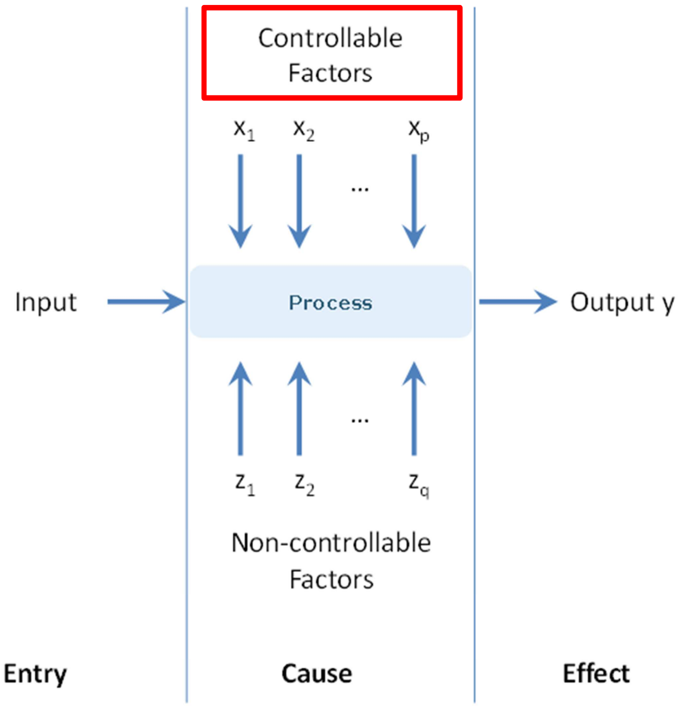
Why conduct trials / experiments?

Cause and effect (causality)

A **trial** / **experiment** is carried out to discover a cause-and-effect relationship in a process.

The lady tasting tea
A lady declares that by tasting a cup of tea made with milk she can discriminate whether the milk or the tea infusion was first added to the cup. We will consider the problem of designating an experiment by means of which this assertion can be tested.

Fisher in *The Design of Experiments* (1935)



Review of Lecture 04

Summary

General DoE

Trial and error
- Changing many factors per run

DoE according to the principles of One-factor-at-a-time (OFAT)

One-factor-at-a-time (OFAT)
- Changing one factor per run

Principles of DoE

Design according to the principles of DoE

Full factorial designs – e.g. **2^k factorial** designs
- The factor levels are determined before the experiment.
- All possible factor combinations are varied.
- Variants of full factorial designs → Dean et al. (2017)

Fractional factorial designs – e.g. **2^{k-1} factorial** designs
- The factor levels are determined before the experiment.
- Only a (balanced) part of the possible combinations of factors are varied.
- Variants, for example «Latin Square» → Dean et al. (2017) / Lecture 09

Example of a full factorial design

Average dwell time on a website

How do factors IV1 and IV2 affect the dwell time DV, measured in seconds?

DV = dwell time [s] → output with metric scaling

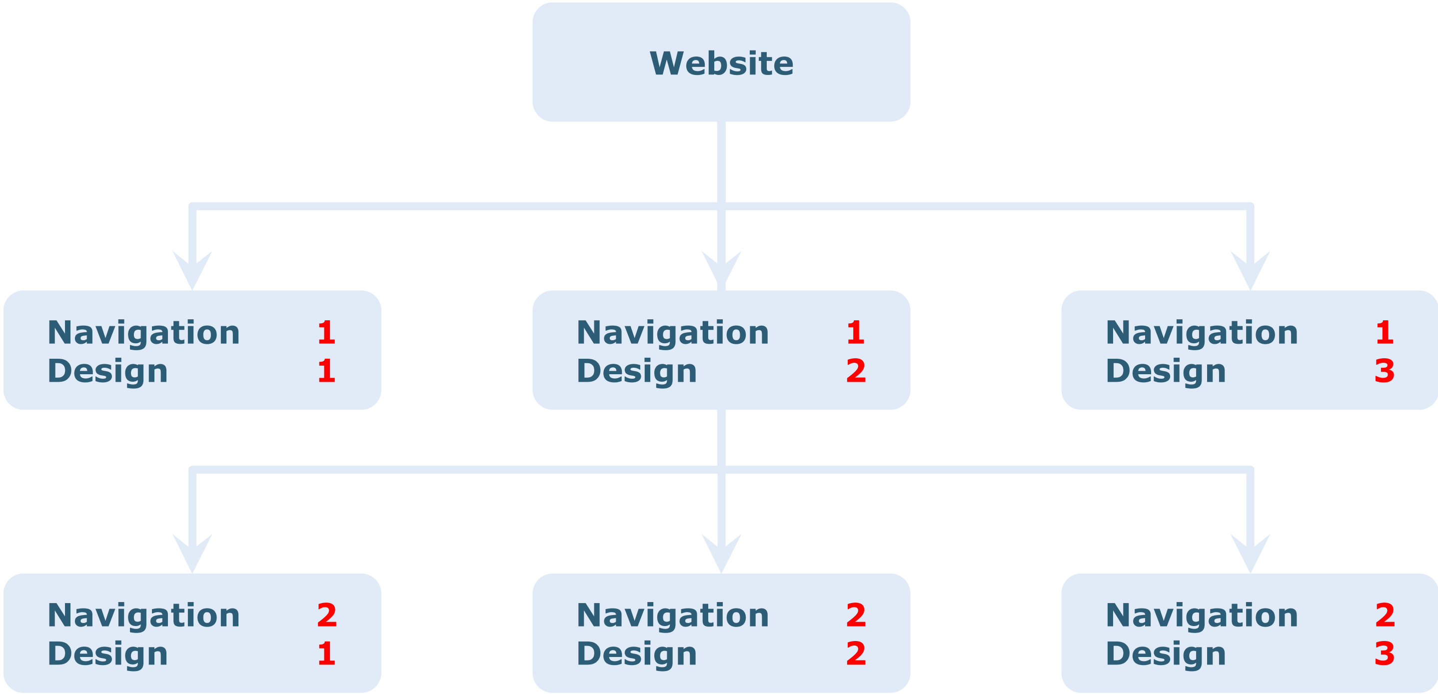
IV1 = design variants [1,2,3] → factor with 3 levels

IV2 = navigation variants [1,2] → factor with 2 levels

$\overline{A}_{ij}$ = mean of the dwell time in group ij

		IV1		
		1	2	3
IV2	1	$\overline{A}_{11}$	$\overline{A}_{12}$	$\overline{A}_{13}$
	2	$\overline{A}_{21}$	$\overline{A}_{22}$	$\overline{A}_{23}$

Note: See also "Lecture 07: Addendum: Introduction to Analysis of Variance (ANOVA)"



2 Factors
6 Groups

n_g = 384
n = 2'304

$\bar{A}_{ij}$ = mean of the dwell time in group ij

		IV1		
		1	2	3
IV2	1	35.5	37.2	39.2
	2	36.9	40.3	45.0

ANOVA with data set *dwelltime* and R file *average_time*

Only main effects, no interaction

```
library(readxl)
dwelltime <- read_excel("<YOUR PATH>/dwelltime.xlsx")

fit  <-  aov(DV  ~  factor(IV1)  +  factor(IV2),  data  =  dwelltime)
summary(fit)
```

```
              Df Sum Sq Mean Sq F value Pr(>F)    
factor(IV1)    2  13256    6628   80.71 <2e-16 ***
factor(IV2)    1   6829    6829   83.16 <2e-16 ***
Residuals   2300 188879      82                   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

There is a main effect of IV1 (levels 1, 2, 3) on DV, $F(2, 2300) = 80.71$, $p = .000$.
There is a main effect of IV2 (levels 1, 2) on DV, $F(1, 2300) = 83.16$, $p = .000$.

Design variants (1, 2, 3) and the navigation variants (1, 2) have a significant effect on dwell time.

Average dwell time with interaction

How do factors IV1 and IV2 affect the dwell time DV, measured in seconds?

DV	= dwell time [s]	→ output with metric scaling
IV1	= design variants [1,2,3]	→ factor with 3 levels
IV2	= navigation variants [1,2]	→ factor with 2 levels

The two factors IV1 and IV2 could influence each other in their effect.

→ Interaction between IV1 and IV2

The combination of factor levels creates an autonomous effect.

For this purpose, a virtual third factor is «automatically» introduced in the model

→ Interaction term [in the R output → factor(IV1):factor(IV2)

The number of factors and thus the number of groups remain the same →

2 Factors
6 Groups

ANOVA with data set *dwelltime* and R file *average_time*

With interaction

```
factor(IV1)  +  factor(IV2) → without interaction
factor(IV1)  *  factor(IV2) → with interaction
```

```
library(readxl)
dwelltime <- read_excel("<YOUR PATH>/dwelltime.xlsx")

fit  <-  aov(DV  ~  factor(IV1)  *  factor(IV2),  data  =  dwelltime)
summary(fit)
```

```
                          Df Sum Sq Mean Sq F value   Pr(>F)    
factor(IV1)                2  13256    6628   81.48  < 2e-16 ***
factor(IV2)                1   6829    6829   83.95  < 2e-16 ***
factor(IV1):factor(IV2)    2   1946     973   11.96 6.79e-06 ***
Residuals               2298 186933      81                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

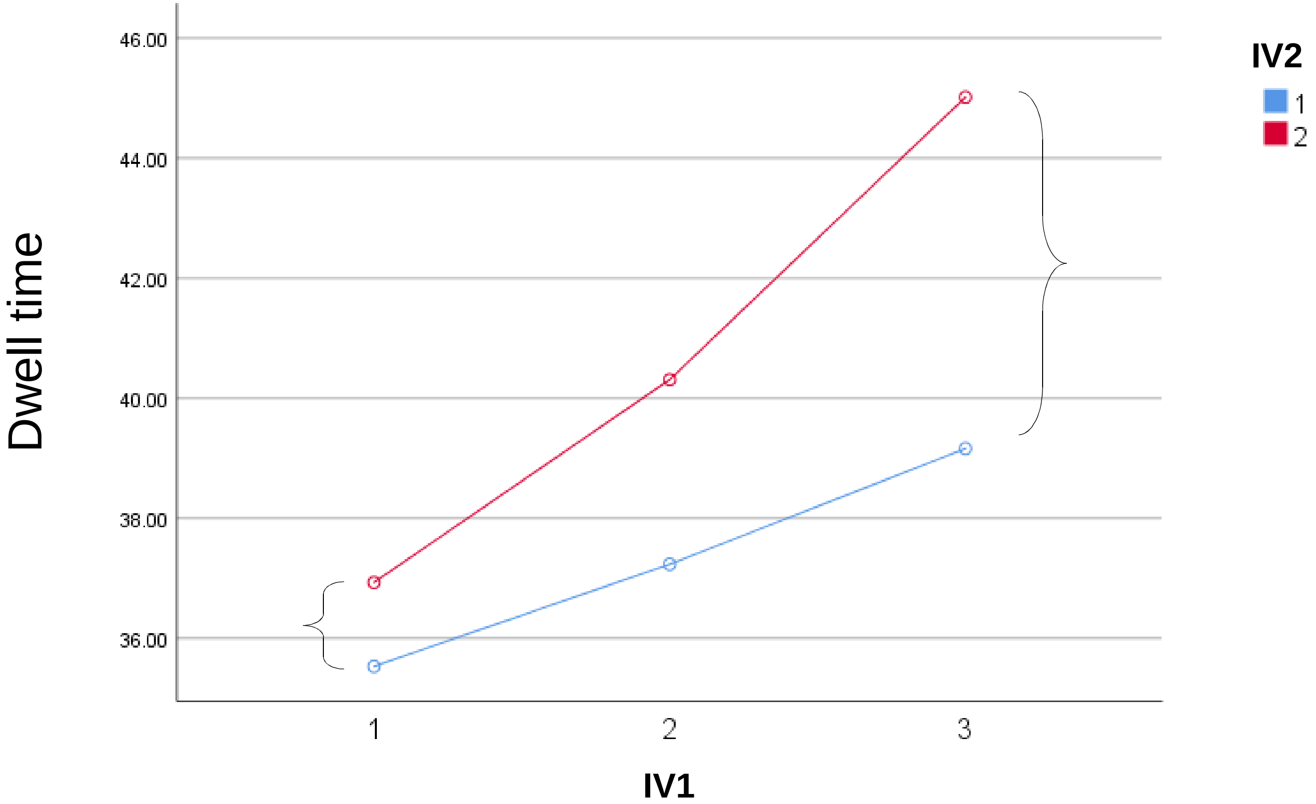
There is a main effect of IV1 (levels 1, 2, 3) on DV, $F(2, 2298) = 81.48$, $p = .000$.
There is a main effect of IV2 (levels 1, 2) on DV, $F(1, 2298) = 83.95$, $p = .000$.

There is a significant interaction, $F(2, 2298) = 11.96$, $p = .000$.

Interpretation of the interaction

Do the levels of the design variants (IV1) and the navigation variants (IV2) influence each other? Is the influence of one variable different depending on the characteristics of the other variable?

Yes, if IV1 takes the values 2 or 3, the influence of IV2 increases.



Simplified: Lines are not parallel

Interpretation: Navigation variant 2 has more influence in design variants 2 and 3.

Full factorial designs

Single factorial (One)

One independent variable **IV** (factor) acts on the dependent variable **DV**
Two groups are compared

1 IV → **one IV with two levels** (dichotomous)
1 DV → metric scaled

Examples: A/B test (→ Lecture 08), RCT with treatment and control or two treatments

Analysis: One-way ANOVA (also: t-test as a special case of univariate ANOVA)

One independent variable **IV** (factor) acts on the dependent variable **DV**
Several groups are compared

1 IV → **one IV with several levels**
1 DV → metric scaled

Examples: RCT with combinations of more than two treatments and control

Analysis: One-way ANOVA

Multifactorial (Several)

Several factors act on the dependent variable **DV**

Several groups are compared

X IV → **several IV with two or more levels each**

1 DV → metric scaled

Examples: Dwell time with two IV (→ Slide 5)

Analysis: Multi-factorial ANOVA (two-way ANOVA, three-way ANOVA, ...)

Fractional factorial designs → Slide 21ff

In principle multifactorial (Several)

Several factors act on the dependent variable **DV**

Several groups are compared

X IV → **several IV with two or more levels each**

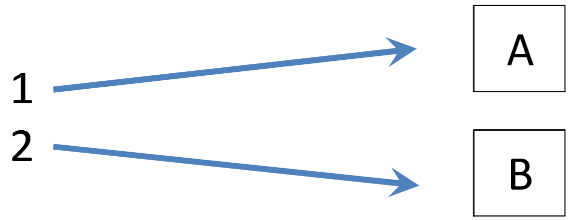
1 DV → metric scaled

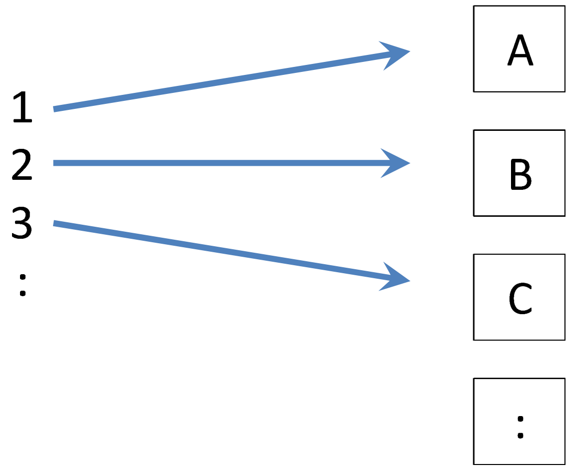
BUT **Not all possible combinations are considered**

→ Latin squares, Greek-Latin squares, ...

→ Nesting of one IV under another IV → Hierarchical designs

	IV Independent variables	**DV** Dependent variable
Single factorial	**One IV**	**One DV**
	Two levels	Divided into two groups
	OR	
	Several levels	Several groups





IV
Independent variables

DV
Dependent variable

Multifactorial

Several IV — **One DV**

Two levels

1 A
2 B

A
B

Divided into two groups

ADDITIONAL

Two levels

1 C
2 D

	A	B
C	AC	BC
D	AD	BD

Four groups

ADDITIONAL

Several levels

1 E
2 F
3 G
:

	A	B
C	AC	BC
D	AD	BD

E F G

Several groups

etc.

Full factorial designs with control

Review of Lecture 03

Control of the secondary variance

- Elimination / Keeping constant
- Repetition / Randomization / Blocking
- Nuisance variable as covariate

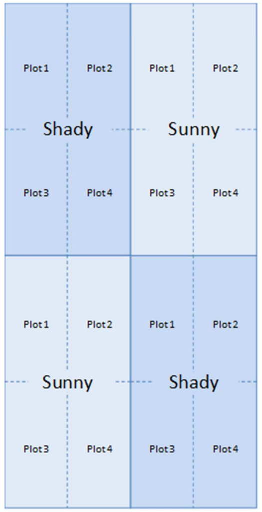
Methods for controlling secondary Variance (Individually or combined) …

- Keeping the experimental setup constant ↔ Particularly possible in laboratory experiments
- Repetition
  Several measurements are repeated on the same probands / trial objects.
- Randomization
  Trial objects are assigned randomly to *Treatment* and *Control* groups
  to eliminate systematic bias.
- Blocking
  Trial objects are grouped into homogeneous blocks based on one or more influential variables to reduce variability.
- Covariate adjustment
  Nuisance variables are included as covariates in the statistical model
  to account for their effects.

Example of blocking: Experiment to study the effect of fertilizer on wheat yield.

Effect of two fertilizers (+/–) on the yield of two types of wheat (+/–).

Problem: The areas of the field varies in terms of solar radiation → nuisance factor



→ **Block factor**
Shady = 1
Sunny = 2

Block	Plot	Fertilizer	Wheat
1	1	+	+
1	2	–	–
1	3	+	–
1	4	–	+
2	1	–	–
2	2		
2	3		
2	4		

In each
with all
indeper

Control by means of … ◦ Nuisance variable as covariate

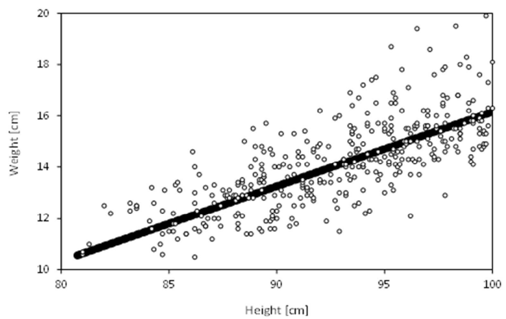
Conversion of a nuisance factor into an IV using the example of body growth in children

Primary variance is modeled using regression analysis.
The nuisance factor "sex" is included as an additional independent variable (control variable).

Model without sex as IV
Both sexes together

$y = \beta_0 + \beta_1 \cdot x$

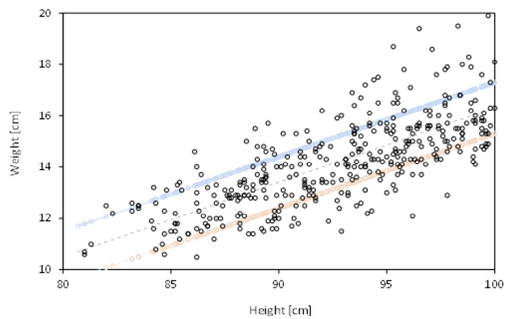
Model accuracy $R^2 = 0.557$



Model with sex as additional IV
Sexes separated

$y = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot \text{sex}$

Model accuracy $R^2 = 0.601$



The relationship between height and body weight is estimated more precisely
↔ Secondary variance is controlled through the inclusion of sex as an additional IV.

Example of average dwell time – with covariate

How do factors IV1 and IV2 affect the dwell time DV, measured in seconds?

DV	= dwell time [s]	→	output with metric scaling
IV1	= design variants [1,2,3]	→	factor with 3 levels
IV2	= navigation variants [1,2]	→	factor with 2 levels

The impact of an additional variable is taken into account:

IV3	= text amount [number of words]	→	metric variable

How can a metric variable be included in an ANOVA?

As a covariate

→ AN**C**OVA (analysis of **c**ovariance)

The number of factors and thus the number of groups remain the same →

2 Factors
6 Groups

ANOVA with data set *dwelltime* and R file *average_time*

With covariate IV3

Instead of entering the variable as a factor, enter it directly.

```
library(readxl)
dwelltime <- read_excel("<YOUR PATH>/dwelltime.xlsx")

options(contrasts = c("contr.sum", "contr.sum"))
fit  <-  aov(DV ~ factor(IV1) * factor(IV2) + IV3,  data  =  dwelltime)
summary(fit)
```

```
                          Df Sum Sq Mean Sq F value   Pr(>F)    
factor(IV1)                2  13256    6628  81.644  < 2e-16 ***
factor(IV2)                1   6829    6829  84.119  < 2e-16 ***
IV3                        1   1177    1177  14.496 0.000144 ***
factor(IV1):factor(IV2)    2   1228     614   7.562 0.000533 ***
Residuals               2297 186475      81                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

There is a main effect of IV1 (levels 1, 2, 3) on DV, $F(2, 2297) = 81.644$, $p = .000$.
There is a main effect of IV2 (levels 1, 2) on DV, $F(1, 2297) = 84.119$, $p = .000$.
There is a significant interaction, $F(2, 2297) = 7.562$, $p = .000$.
The covariate IV3 has a significant impact, $F(1, 2297) = 14.496$, $p = .000$.

Example of average dwell time – with block factor

How do factors IV1 and IV2 affect the dwell time DV, measured in seconds?

DV	= dwell time [s]	→	output with metric scaling
IV1	= design variants [1,2,3]	→	factor with 3 levels
IV2	= navigation variants [1,2]	→	factor with 2 levels

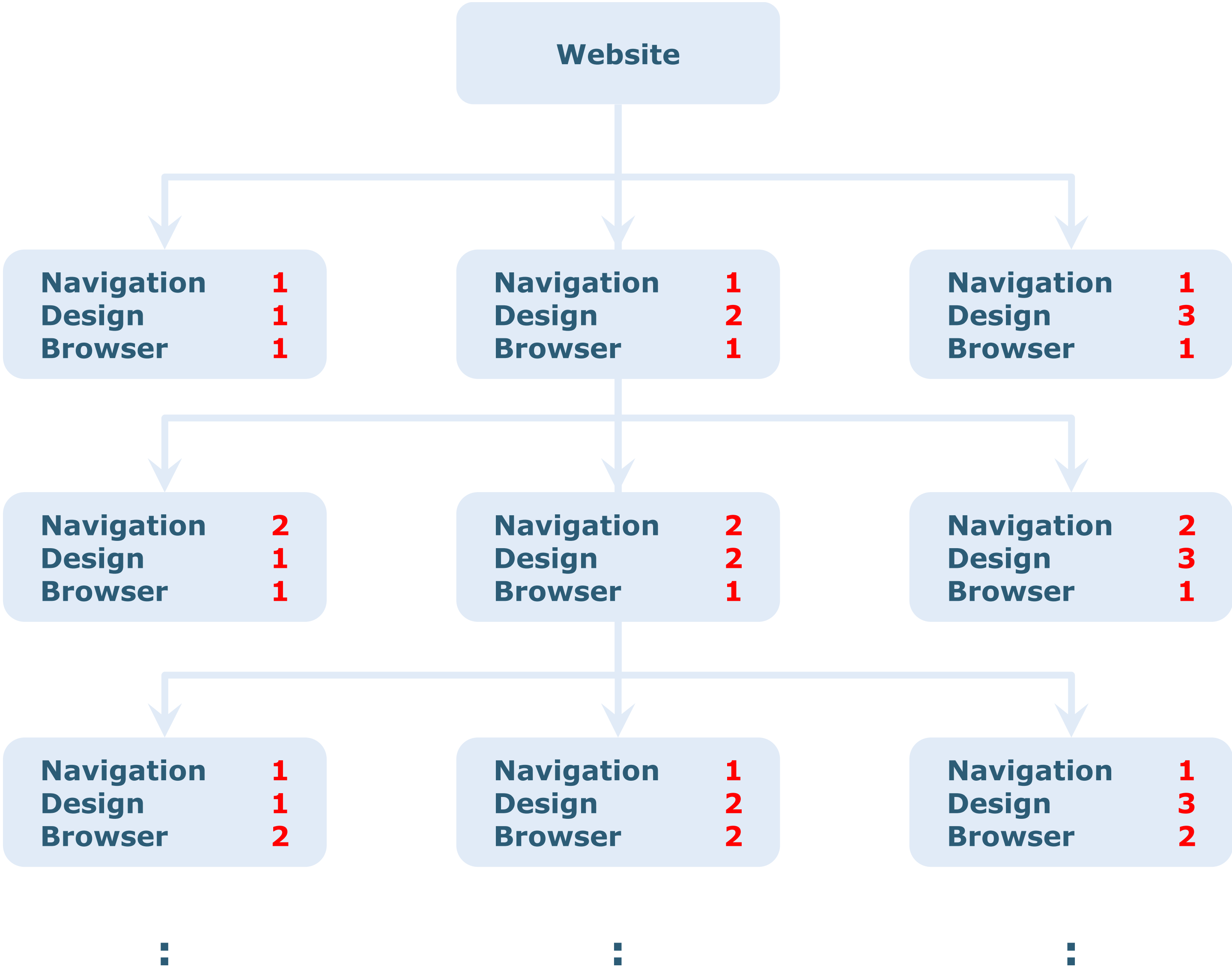
The impact of an additional variable is taken into account:

IV4	= Browser type [1,2,3]	→	factor with 3 levels

The impact of the additional variable IV4 is not part of the research question.

Therefore, although IV4 is introduced as a factor – as a **block factor** – it is not interpreted.

Note: ANOVA does not distinguish between factor and block factor.



3 Factors
18 Groups

Fractional factorial designs

Reduction of the number of groups

The more factors and the more characteristics, the larger the number of groups.
Question: How can the number of groups be reduced?

Example: A design for three factors IV1, IV2, IV3 with three levels each [1, 2, 3]

IV1	IV2	IV3
1	1	1
1	1	2
1	1	3
1	2	1
1	2	2
1	2	3
1	3	1
1	3	2
1	3	3
2	1	1
2	1	2
2	1	3

:

3 x 3 x 3 = 27 groups must be included

Provided that there are **no interactions**, the experiment can also be conducted successfully with a reduced number of groups.

By using **Latin squares** or related designs, the number of groups required can be significantly reduced compared to full factorial designs.

Fisher (1923) / Colors from Itten (1961)

Latin square

Introduction

A Latin squares design makes it possible to study the main effects of factors without having to observe all combinations of treatment levels.
Compared to the full factorial design, the Latin squares design requires smaller sample sizes.

A Latin squares design can only be used if it follows from theory or empirical evidence that the joint effect of the factors does not produce interactions.
That is, interactions between factors cannot be verified by a Latin squares design.
Accordingly, if interactions are actually present, the main effects are not clearly interpretable.

Structure of a Latin square – Examples with three factors

Three factors, which must have the same number of factor levels.

Example 1: Factors A, B, C, each with two factor levels (a_1, a_2), (b_1, b_2), (c_1, c_2) $\rightarrow$ $p = 2$

	a_1	a_2
b_1	c_1	c_2
b_2	c_2	c_1

The factor level combination a_1b_1 is combined with c_1, a_2b_1 with c_2, a_1b_2 with c_2 and a_2b_2 with c_1.

In an experiment, each of the four combinations is assigned a random sample of size n.

Example 2: Factors A, B, C, each with three factor levels (a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3) → $p = 3$

	a_1	a_2	a_3
b_1	c_1	c_2	c_3
b_2	c_2	c_3	c_1
b_3	c_3	c_1	c_2

The factor level combination a_1b_1 is combined with c_1, a_2b_1 with c_2, a_3b_1 with c_3, a_1b_2 with c_2 etc.

In an experiment, each of the nine combinations is assigned a random sample of size n.

Construction rules

In each row and in each column, each C-level appears only once.
For $p = 3$, there are twelve arrangements that satisfy this condition.

Two further examples

	a_1	a_2	a_3
b_1	c_3	c_1	c_2
b_2	c_2	c_3	c_1
b_3	c_1	c_2	c_3

	a_1	a_2	a_3
b_1	c_2	c_1	c_3
b_2	c_1	c_3	c_2
b_3	c_3	c_2	c_1

However, in the case of p = 3, there is only one arrangement in which the C-levels in the first row and the first column appear in a natural sequence (c1, c2, c3) → **Standard form** (see above)

Balancing and decomposition

Balancing = Balance in the combination of factor levels
Decomposition = Breaking down a complete design into smaller, balanced subsets

Full* Factorial — Each level of C is combined only once with each level of A and with each level of B. The design is completely balanced in terms of the main effects.

Fractional** Factorial — Each level of C occurs only partially with other combinations of A and B. The Latin square is only partially balanced with respect to first-order interaction.

Relationship between a complete design C and a Latin square design L (balanced form)

	a_1									a_2									a_3								
	b_1			b_2			b_3			b_1			b_2			b_3			b_1			b_2			b_3		
C	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
	↑				↑				↑		↑				↑	↑					↑	↑				↑	
L	1				2				3		4				5	6					7	8				9	

Arrows point to the factor level combinations in the Latin square → Slide 23

C Full factorial → 27 combinations
L Latin square → 9 combinations

Ratio = $\frac{9}{27} = \frac{1}{3}$ General***: Ratio = $\frac{1}{p}$

Shifting the arrows one unit to the right (*modulo*) results in further factor level combinations, which in turn produce other Latin squares.

→ A complete 3×3×3 design can be divided into three balanced Latin squares.

*Also called "Complete design" / **A Latin square is an example / ***p = Factor levels

Other fractional factorial designs / other methods

Greco-Latin square

Extension of the Latin square to four factors.

Counterbalancing in case of repeated measures

Problem → influence of a treatment on the result in one of the following treatments.
For example learning and fatigue effects also called «carry-over» effect.
Solution «counterbalancing» → conducting as many test sequences as needed to ensure that each stage of the treatment occurs equally frequently at each stage of the study.

More → Dean et al. (2017)

More → Oehlert (2010)

For example

Preview of Lecture 10

What has happened so far

We looked at full factorial designs of experiments in more detail by using an example.

We also looked at examples of full factorial designs with control.

For full factorial designs, the complexity and sample size can increase very quickly, resulting in a need for fractional factorial designs.

A classic is the implementation with the Latin square.

What follows in Lecture 10

First, the question occurs «What are large data quantities?»

Specific properties of large data quantities become apparent, which have to be considered more closely in the context of design of experiments and statistics.

Possible solutions and procedures are discussed to deal with the problems associated with large data quantities.

Appendix

Details of a research project

Reaerch question

How do factors A, B and C affect the dependent variable (DV) dwell time, measured in seconds?

DV	= dwell time [s]	→ output with metric scaling
A	= design variants [a_1, a_2, a_3]	→ factor with 3 levels
B	= navigation variants [b_1, b_2, b_3]	→ factor with 3 levels
C	= browser type [c_1, c_2, c_3]	→ factor with 3 levels

Sample size

Based on power analysis, 1,200 persons are collected for each combination of factors.

Complete design	→ 27 combinations	27 x 1,200 = 32,400 persons
Latin square	→ 9 combinations	9 x 1,200 = 10,800 persons

Experimental design

Creating a Latin square with R script: ###### EXERCISE 09: Factorial Designs

```
library(agricolae)

my_design_lsd <- design.lsd(trt = c("c1","c2","c3"))
my_design_lsd$sketch
```

trt = short for "Treatments"

```
     [,1] [,2] [,3]
[1,]
[2,]
[3,]
```

Output
This is part of the task in the exercise

Interpretation
This is part of the task in the exercise

Sampling

A sample with n = 1,200 is drawn for each of the combinations This is part of the task in the exercise

	1	2	3	4
1	AV	A	B	C
2	43.7	1	1	1
3	28.7	1	1	1
4	40.0	1	1	1
5	41.4	1	1	1
6	29.7	1	1	1
7	48.5	1	1	1
8	34.7	1	1	1

dwelltime_Latin.xlsx

CRAN.R-project.org/package=agricolae

Statistical data analysis

ANOVA with data set *dwelltime_Latin* and R script: ##### EXERCISE 09: Factorial Designs

```
library(readxl)
dwelltime_Latin <- read_excel("<YOUR PATH>/dwelltime_Latin.xlsx")
View(dwelltime_Latin)

fit  <-  aov(DV  ~  factor(A) + factor(B) + factor(C),  data = dwelltime_Latin)
summary(fit)
```

```
                Df Sum Sq Mean Sq F value Pr(>F)    
factor(A)        2  76057   38028 465.761 <2e-16 ***
factor(B)        2  27159   13580 166.320 <2e-16 ***
factor(C)        2    552     276   3.383  0.034 *  
Residuals    10338 844075      82                   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

<u>Interpretation</u>

This is part of the task in the exercise

Table of contents