MSc Applied Information and Data Science

Data Analytics for Data Scientists

Design of Experiments (DoE)

Lecture 04: Properties of DoE

2025

Prof. Dr. Jürg Schwarz

Program: 16:15 until 17:55

16:15	**Start of the lesson**
	Lecture: Jürg Schwarz ◦ What is DoE? ◦ Research Process ◦ Quality aspects of DoE ◦ Preview of Lecture 05
	Exercises: Students / Jürg Schwarz / Assistants ◦ Working on the exercises ◦ Support by Jürg Schwarz / Assistants
17:55	**End of the lesson**

Detailed table of contents → Slide 21

What is DoE?

An example

<u>Research question</u>

Which factors influence the fuel consumption of a car?

DoE type: Trial and error

<u>Experimental design</u>

Procedure

- Change the fuel brand (Agrola / Avia / BP / ...)
- Change the octane number (Super 95 / Super 98)
- Change the speed (80 km/h / 100 km/h)
- Adjust the engine
- Change the tire pressure (2.5 bar / 3.0 bar)
- Wash the car

Measurement

- Fuel consumption: Liters of gasoline / kilometers traveled

Result

- No idea what factors influence fuel consumption!

DoE type: One-factor-at-a-time (OFAT)

Experimental design

Procedure

- Vary the first factor and then measure fuel consumption.
  Keep the setting with the lowest consumption and then vary the next factor.

Speed	Tire pressure	Octane number	Consumption
80	2.5	95	7.5
100	2.5	95	8.8
80	2.5	95	7.5
80	3.0	95	7.2
80	3.0	95	**7.2**
80	3.0	98	7.4

Properties

- Easy to implement
- Interaction between factors are not recognized.
- Research question is answered neither systematically nor exhaustively.
- Number of runs to determine the **global minimal consumption** is unknown.

DoE type: Full factorial design – Two-levels

Experimental design

Procedure

- Two levels (+/-) are defined per factor.
  All possible combinations of factor levels are varied.

Speed		Tire pressure		Octane number		Consumption
80	-	2.5	-	95	-	7.5
100	+	2.5	-	95	-	8.8
80	-	3.0	+	95	-	**7.2**
100	+	3.0	+	95	-	8.4
80	-	2.5	-	98	+	7.7
100	+	2.5	-	98	+	8.6
80	-	3.0	+	98	+	7.4
100	+	3.0	+	98	+	7.8

Properties

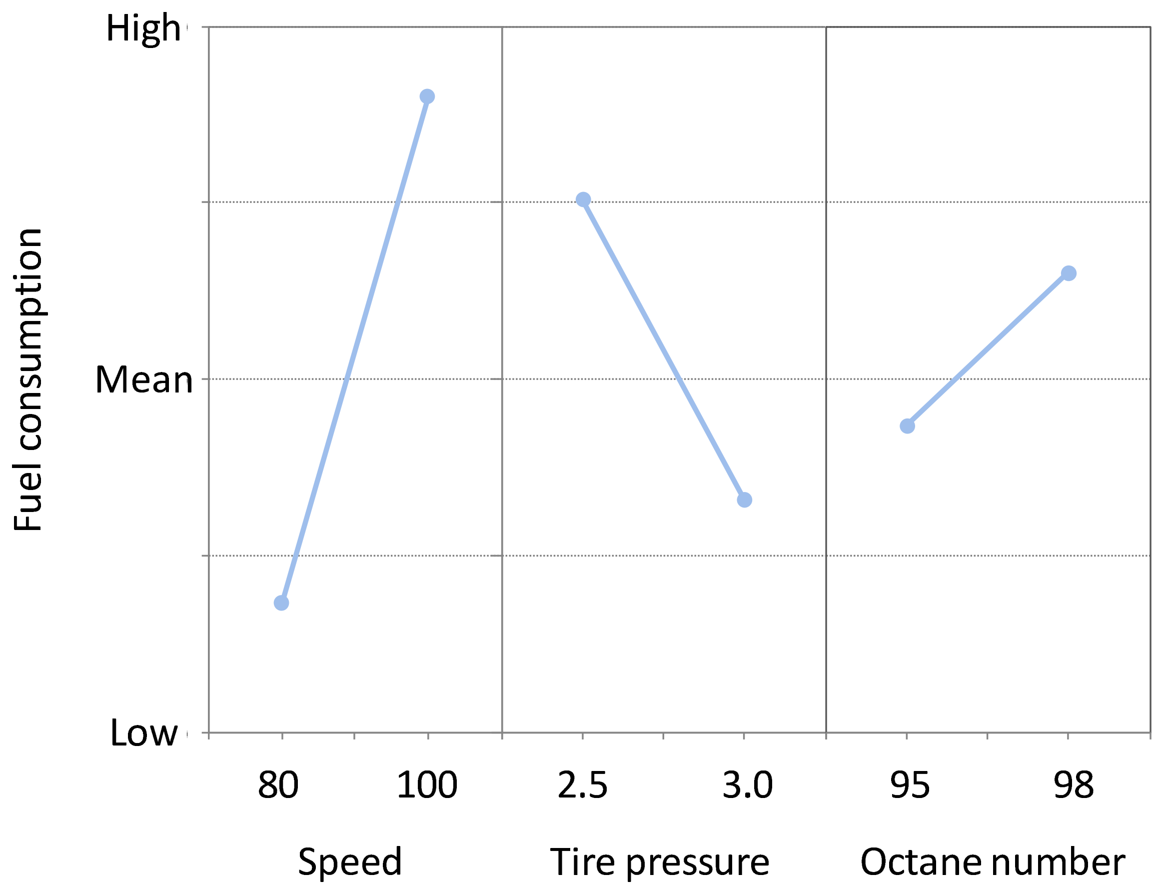
- All main effects and all interactions can be determined.
- Can be used as a *screening experiment* to identify potentially important variables.
- The effort involved increases rapidly as the number of factors increases:
  Each additional factor doubles the number of combinations.

Profile plots of the main effects*

Impact of the factors on the dependent variable fuel consumption

Model: fuel consumption = $\beta_0 + \beta_1 \cdot \text{speed} + \beta_2 \cdot \text{tire pressure} + \beta_3 \cdot \text{octane number}$

Main effects: speed, tire pressure, octane number



Example interpretation

The increase in speed from 80 km/h to 100 km/h causes the largest relative increase in fuel consumption.

Tire pressure and octane number are kept constant.

*More to come in Lecture 07

Legend

s → speed
t → tire pressure
o → octane number

Factorial design with interactions

Interactions can occur in experiments with two or more independent variables.
An interaction of two factors means that the two factors interact in a complex way.

If there is an interaction, the effect of one factor depends on the levels of the other factor.

Example speed and tire pressure:
The effect of speed on fuel consumption differs depending on tire pressure.

Interaction terms are written as multiplication: speed x tire pressure, etc.

Two way interactions **s** x **t** / **s** x **o** / **t** x **o**

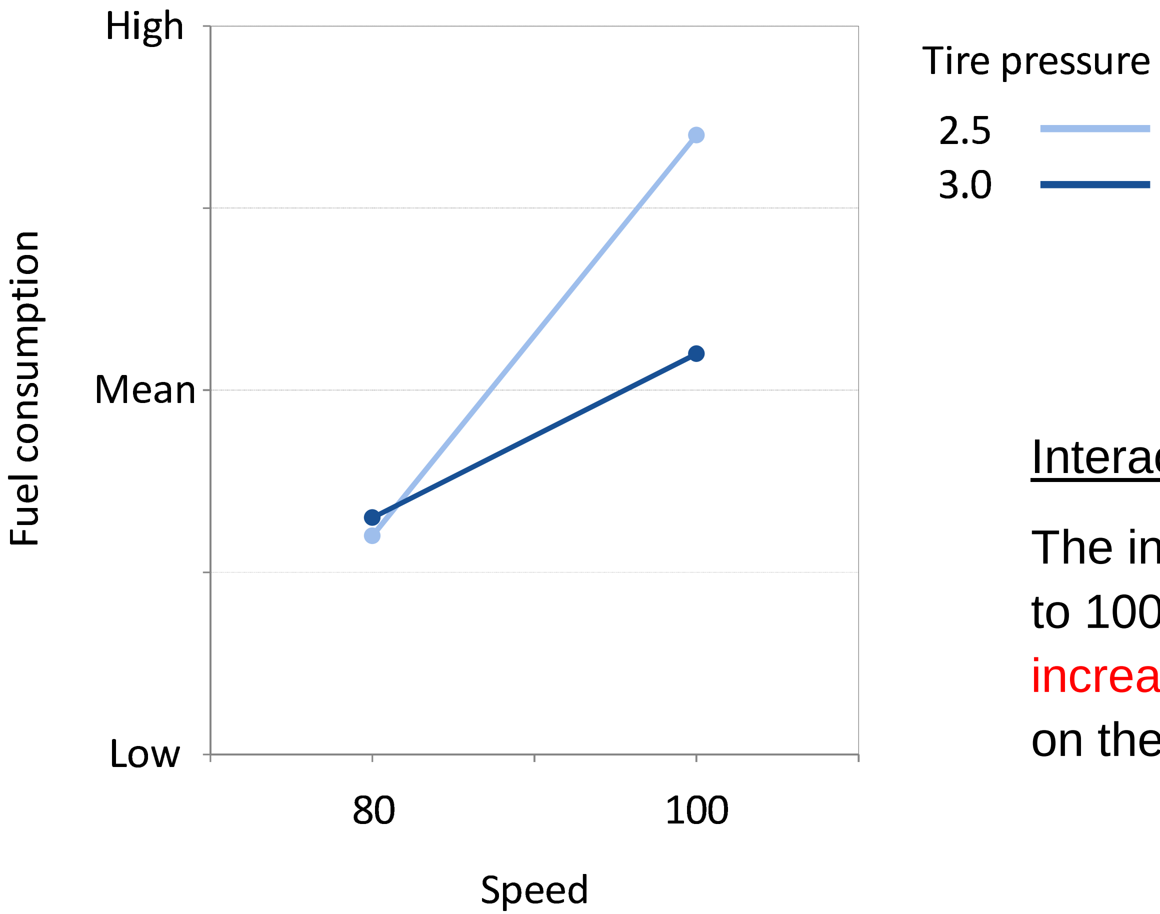
Three way interaction **s** x **t** x **o**

s	t	o	s x t	s x o	t x o	s x t x o	Consumption
+	-	-	+ -	+ -	- -	+ - -	…
-	+	-	- +	- -	+ -	- + -	…
-	-	+	- -	- +	- +	- - +	…
+	+	+	+ +	+ +	+ +	+ + +	…
+	+	-	+ +	+ -	+ -	+ + -	…
+	-	+	+ -	+ +	- +	+ - +	…
-	+	+	- +	- +	+ +	- + +	…
-	-	-	- -	- -	- -	- - -	…

Profile plot of the interaction using speed x tire pressure as an example

Interaction term: speed x tire pressure

Model: fuel consumption = $\beta_0 + \beta_1 \cdot$ speed ... + $\beta_3 \cdot$ octane number + $\beta_4 \cdot$ speed x tire pressure



<u>Interaction</u>

The increase in speed from 80 km/h to 100 km/h causes a different relative increase in fuel consumption depending on the level of the tire pressure.

The categories of the one factor are shown on the horizontal axis.
The legend shows the color-coded categories of the second factor.

Generalization of two-level full factorial design – 2^k factorial designs

2 factors (A / B)

2 levels (+/-)

$2 \times 2 = 2^2 = 4$ combinations

A	B	C	D
-	-	-	-
+	-	-	-
-	+	-	-
+	+	-	-
-	-	+	-
+	-	+	-
-	+	+	-
+	+	+	-

3 factors (A / B / C)

2 levels (+/-)

$2 \times 2 \times 2 = 2^3 = 8$ combinations

-	-	+	+
+	-	+	+
-	+	+	+
+	+	+	+

etc.

4 combinations

		Factor A	
		-	+
Factor B	-	- -	- +
	+	+ -	+ +

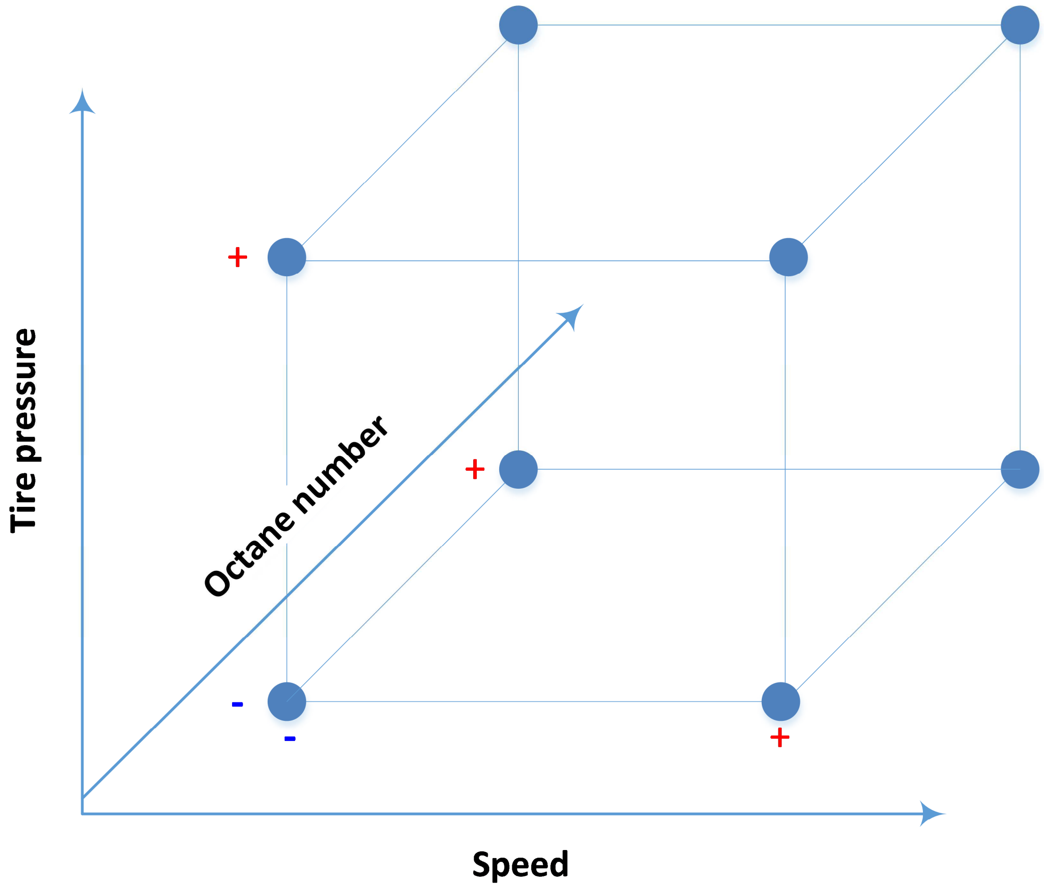
k factors (A / …)

2 levels (+/-)

$2 \times 2 \times \ldots = \mathbf{2^k}$ combinations

Two-level full factorial design with 3 factors in 3D

The example of fuel consumption



3 factors (s / t / o)

2 levels (+/-)

$2 \times 2 \times 2 = 2^3 = 8$ combinations

Legend

s → speed
t → tire pressure
o → octane number

Fractional factorial design: 3 factors on 2 levels – 2^{3-1} factorial design

3 factors (s / t / o)

2 levels (+/-)

$\}$ 2 x 2 x 2 = 2^3 = 8 combinations

s	t	o	s x t	s x o	t x o	s x t x o	❶	Consumption
+	-	-	+ -	+ -	- -	+ - -	+1	...
-	+	-	- +	- -	+ -	- + -	+1	...
-	-	+	- -	- +	- +	- - +	+1	...
+	+	+	+ +	+ +	+ +	+ + +	+1	...
+	+	-	+ +	+ -	+ -	+ + -	-1	...
+	-	+	+ -	+ +	- +	+ - +	-1	...
-	+	+	- +	- +	+ +	- + +	-1	...
-	-	-	- -	- -	- -	- - -	-1	...

❶ In terms of mathematical operation

+		-		-	=	+
(+1)	x	(-1)	x	(-1)	=	+1

Half of them are selected: 4 combinations ↔ $2^{3-1} = 2^2 = 4$

But which half? → In general there are many possibilities of selection → Dean et al. (2017)

In many cases the selection is made based on the properties of the **main effects** (s / t / o) and the **interactions** (s x t / s x o / t x o and s x t x o).

In the example, the choice is that the triple interaction s x t x o takes on the same value (+1)

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

Applying the principles of DoE

DoE type: Full factorial designs

→ See above starting from Slide 5

DoE type: Fractional factorial designs

<u>Design</u>

Procedure

- The factor levels are determined before the experiment.

Factor combinations

- Only a part of the possible combinations of factors are selected.

Restrictions

- In fractional factorial designs, interactions can only be partially measured because not all possible combinations of factor levels are tested.
  Pilot studies and expert interviews can be useful for predicting the presence of interactions.

Advantages

- The effort involved is significantly lower compared to full factorial designs.

Statistical analysis

- As in the case of full factorial design, but without interactions.

Research Process

Integration of the research process in DoE

DoE comprises most of the elements of the research process (→ Lecture 02)

DoE …

- determines the study design
  (→ Lecture 02)
- defines and determines how data is collected and analyzed, in particular ...
  properties of the dependent variable (**DV**)
  information about the number and properties of the independent variable (**IV**)
  (→ Lecture 03)
- consists of a list of experimental conditions
  (→ Lecture 04)
- determines the type and implementation of the sampling
  (→ Lecture 05)
- determines sample size
  (→ Lecture 06)
- determines the statistical methods used to analyze the data
  (→ Lectures 08 & 09)

An example – Intervention experiment to raise the quality of teaching

- Research question: Can a specific coaching program improve how teachers interact with children in integration classes in primary schools?

- Study design: Longitudinal, randomized control-group experiment with a total of four measurement points (incl. follow-up) with a treatment group (n = 35 classes) and a control group (n = 35 classes) from integrative regular classes in the 4th to 6th grades.

- Dependent variable (**DV**) and independent variable (**IV**)
  Various variables are collected:
  - DV Interaction quality of teacher, based on video recordings of classroom teaching
  - IV Total 8, e.g. social behavior (physical and mental aggression, prosocial behavior ...)
    Performance of children in reading and writing

- Control variables: Children (age, first language, ...) / teachers (education, experience ...)

- Experimental conditions: Teaching observation and analysis, with feedback and coaching of teachers and experts in special education.

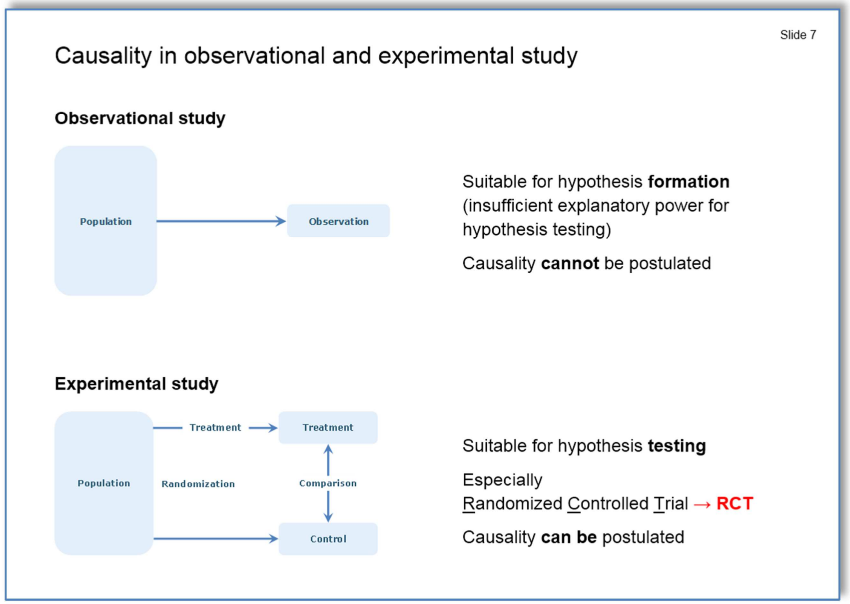
- Sampling: Assigning individual children to intervention and control groups makes little sense because the intervention is carried out at the class level. Therefore, entire school classes are assigned to one of the two groups by means of cluster randomization at the class level.

- Sample size: Based on power analyses → 35 classes (approx. 1,000 children)

- Statistical analysis methods: Multi-level analysis (hierarchical linear modelling)

Quality aspects of DoE

Looking back at Lecture 03: Introduction to DoE



Slide 7

Causality in observational and experimental study

Observational study

Population → Observation

Suitable for hypothesis **formation** (insufficient explanatory power for hypothesis testing)

Causality **cannot** be postulated

Experimental study

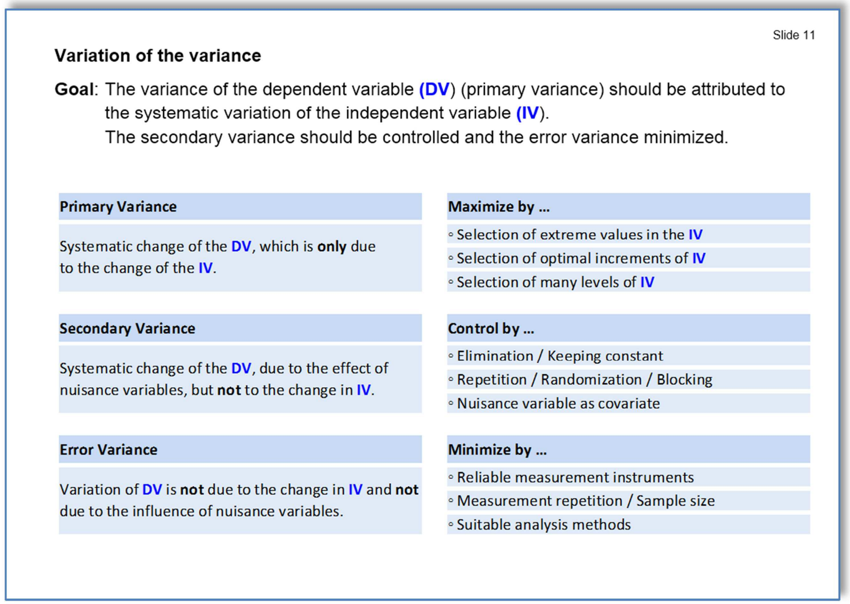
Population — Randomization — Treatment → Treatment; Comparison; → Control

Suitable for hypothesis **testing**

Especially
Randomized Controlled Trial → **RCT**

Causality **can be** postulated

Experimental studies
are best suited for obtaining scientific knowledge.
Among them, Randomized Controlled Trials (RCT) achieve the highest quality.



Slide 11

Variation of the variance

Goal: The variance of the dependent variable **(DV)** (primary variance) should be attributed to the systematic variation of the independent variable **(IV)**.
The secondary variance should be controlled and the error variance minimized.

Primary Variance	**Maximize by ...**
Systematic change of the **DV**, which is **only** due to the change of the **IV**.	◦ Selection of extreme values in the **IV** ◦ Selection of optimal increments of **IV** ◦ Selection of many levels of **IV**
Secondary Variance	**Control by ...**
Systematic change of the **DV**, due to the effect of nuisance variables, but **not** to the change in **IV**.	◦ Elimination / Keeping constant ◦ Repetition / Randomization / Blocking ◦ Nuisance variable as covariate
Error Variance	**Minimize by ...**
Variation of **DV** is **not** due to the change in **IV** and **not** due to the influence of nuisance variables.	◦ Reliable measurement instruments ◦ Measurement repetition / Sample size ◦ Suitable analysis methods

Secondary variance and error variance reduce the quality of the explanatory power of experiments.

Control and minimization are therefore important elements for determining the quality of experiments.

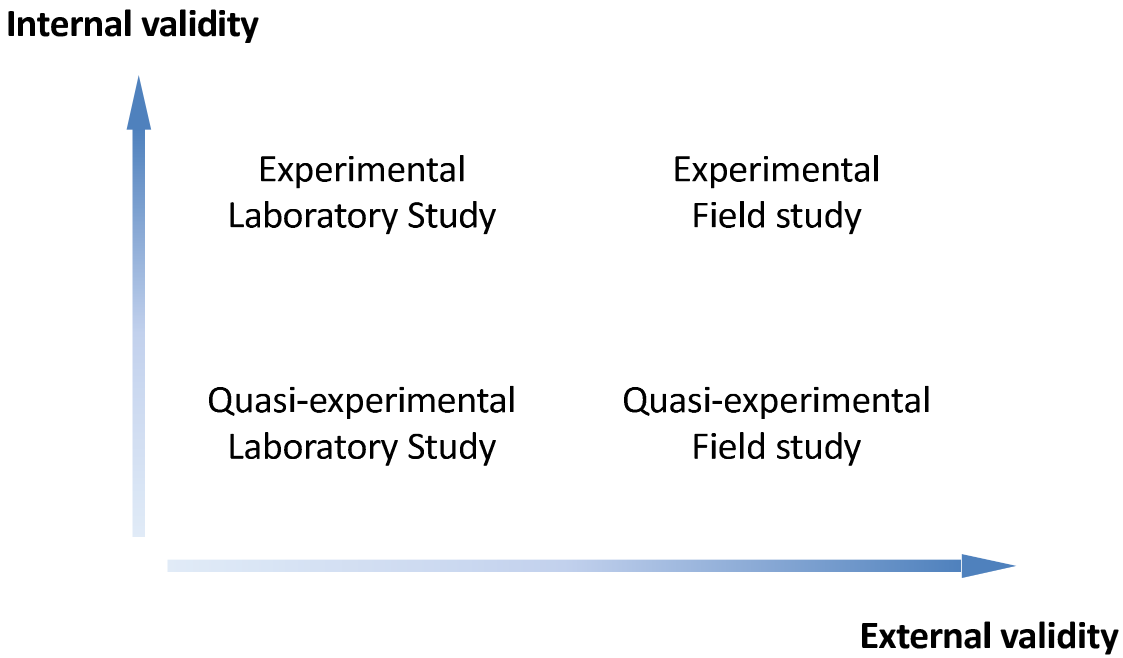
Quality criteria of experiments

Internal validity …

- exists when changes in dependent variables (**DV**) are attributed to independent variables (**IV**) (Quasi-experiment → Experiment / Maximize primary variance)
- increases with decreasing impact of nuisance variables (Quasi-experiment → Experiment / Control secondary variance / Minimize error variance)

External validity ...

- exists when experimental results from a sample can be generalized to the entire population.
- increases with increasing naturalness (Laboratory → Field).



Internal validity – More details in → Lecture 03

External validity – Two subtypes

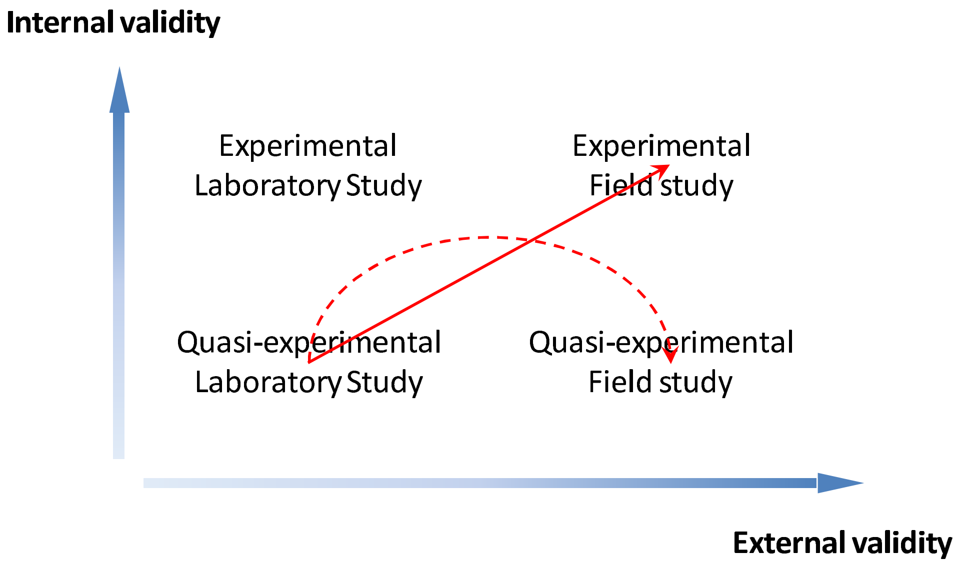
- Population validity → Degree to which the results of a study can be generalized from the sample to the whole population ↔ Lecture about sampling will address "representativeness"
- Situation validity → Degree to which the findings of a study can be applied to different situations.

Construct validity – Additional, distinct concept

- → Effectiveness of the measurement methods in precisely capturing the intended construct ↔ Lecture 02 → Slide 26: Operationalization

Relationship: Internal vs. external validity

- The lowest general level of validity is at the bottom left for the quasi-experimental laboratory study, and the highest is at the top right for the experimental field study.
- In principle, the goal is to arrive at the top right, but with experimental field studies generally being the most expensive (→ solid line →)
- A well-controlled lab experiment may maximize internal validity by eliminating confounding variables, maximizing internal validity. However, its artificial conditions may reduce external validity (→ dashed line --->)



Westermann (2000) / Several articles from Wikipedia

Preview of Lecture 05

What has happened so far

Variants of Design of Experiments (DoE) are known.

Properties of Design of Experiments (DoE) are known.

Quality aspects of Design of Experiments (DoE) are known.

What follows in Lecture 05

Quantitative studies

An important aspect of quantitative studies is finding the "right" sample.
But how do you get there?

Lecture 05 about "Sampling" will provide …

- some definitions ("Population")
- an overview over sampling procedures
- the first answer to the "famous" question "How large must the sample be?"
  (more insight will be given in "Lecture 06 Effect size & Power analysis"

Table of contents