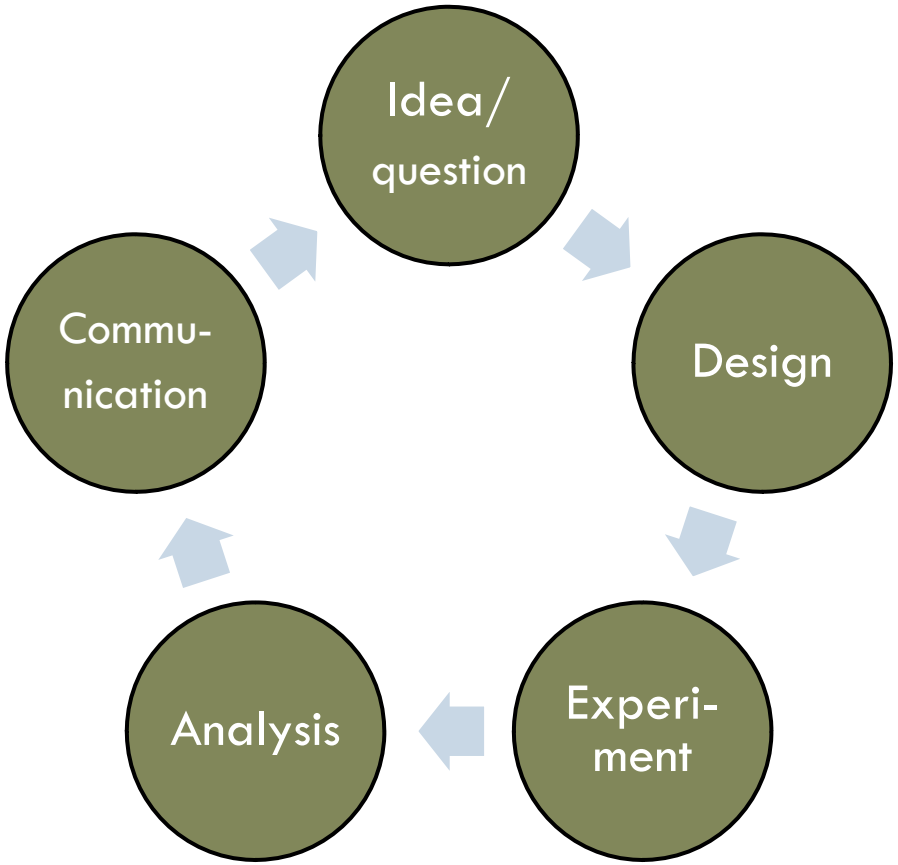
EXPERIMENTAL DESIGN IN HUMAN RESEARCH

Matthias Treder

Tutorial

Machine Learning Lab, TU Berlin 2009

Empirical cycle



Overview

1. From Idea to Design
2. Research design
    1. Types of research
    2. Design of an experiment
3. Measures (Messinstrumente)
    1. Kind of measures
    2. Operationalization
    3. Quality of a measure
4. ~~Confounding variables~~
    1. ~~Situational variables~~
    2. ~~Participant variables~~
    3. ~~Experimenter variables~~

Two running examples

1. What is the influence of environmental noise (e.g., traffic) on BCI performance?

2. How does mood (e.g., happyness, sadness) affect BCI performance?

From Idea to design: Relevance

What makes a *good* research question?

Practical relevance

↓

Benefit for society

Theoretical relevance

↓

Increase scientific understanding

From Idea to design: Relevance

„Influence of environmental noise on BCI performance“

Practical relevance

- Patients/users live in noisy environments → assess the robustness of your BCI system in a (more) realistic setting

Theoretical relevance

- Auditory noise can draw attention → learn more about the attentional processes involved in BCI

From Idea to design: Relevance

„Influence of mood on BCI performance“

Practical relevance

- Patients often suffer from depression, healthy subjects (usually) don‘t

Theoretical relevance

- Investigate the cognitive link between emotional states such as happyness/sadness and ERD-ERS/P300/etc.

From Idea to design: Further points

- Ethical standards
  - Experimental design meets ethical standards of, e.g. Helsinki Declaration?
  - Approval of ethical standards committee?
- Ethics: rule of thumb
  - Participants volunteer (written consent)
  - Data is processed anonymously
  - You don‘t fool the participant / tell lies (➔ if you have to, detailed debriefing)
  - The experiment does not have negative consequences for the participant
- Feasibility
  - Time
  - Money
  - Availability of participants

Overview: 2. Research design

- Types of research
  - Descriptive research
    - ➔ Measure a dependent variable
  - Causal research /Experiment
    - ➔ Measure effect of independent variable on dependent variable
  - Relational research /Quasi-experiment
    - ➔ Measure correlation between two dependent variables
- Design of an experiment
  - Between-subjects design
  - Within-subjects design
  - Mixed design
- Validity
  - Internal validity
  - External validity

Research design: Variable

- Variable: „A factor or a particular situation that can change in observable and measurable ways“

Research design: Indepedent variable

- Manipulation (active)

  ➔ different states of the variable(s) are set by the experimenter

- Usually, the independent variable(s) is/are discrete and has/have a number of different levels (subconditions/Unterbedingungen)
  - Independent variable: Noise
  - Subcondition 1: No noise (NN)
  - Subcondition 2: Environmental noise @60 dB (EN)
  - Subcondition 3: Classical music @60 dB (CM)

Research design: Dependent variable

- Measurement (passive)
- Usually, the dependent variable(s) is/are continuous (can have many different values)
- Examples
  - Reaction time
  - Error rate
  - ERP (event related potential) amplitude

Research design: Causal research/Experiment

Independent variable

Dependent variable

- Subcondition 1 of independent variable → Value 1 of dependent variable
- Subcondition 2 of independent variable → Value 2 of dependent variable
- To establish a causal link between independent and dependent variable (for subconditions 1 and 2), Value 1 ≠ Value 2 (statistically significant)
- In other words, you reject the nullhypothesis that Value 1 = Value 2

Research design: Causal research/Experiment: Example

- Nullhypothesis (H0): „There is no effect of noise on BCI performance“

No noise → 80% correct BCI control

Environmental noise → 60% correct BCI control

Research design: Relational research/Quasi-experiment

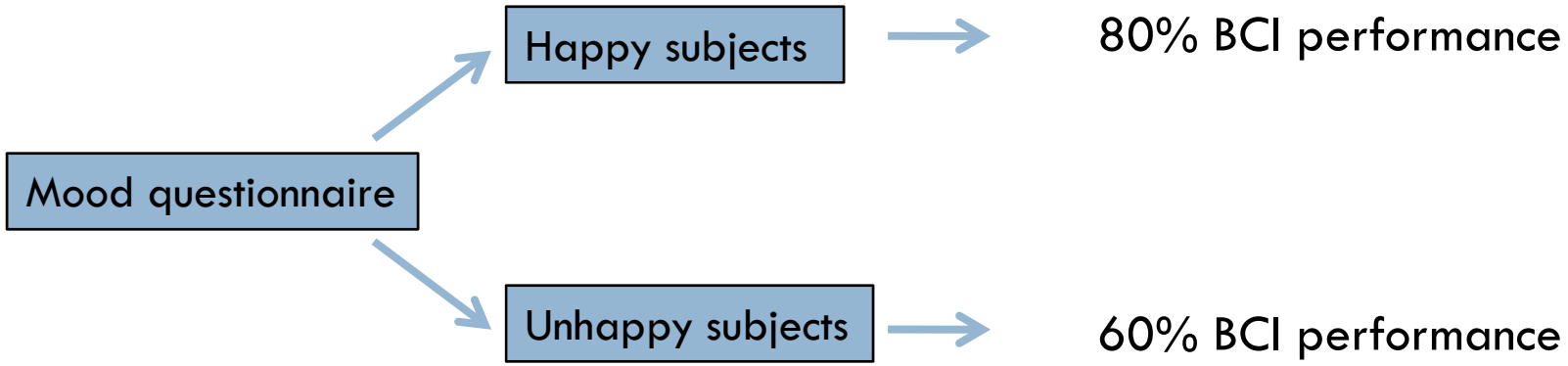
Dependent variable

Dependent variable

- Measure two (or more) dependent variables
- Establish whether there is a correlation between these variables
- You cannot investigate causal relationships!

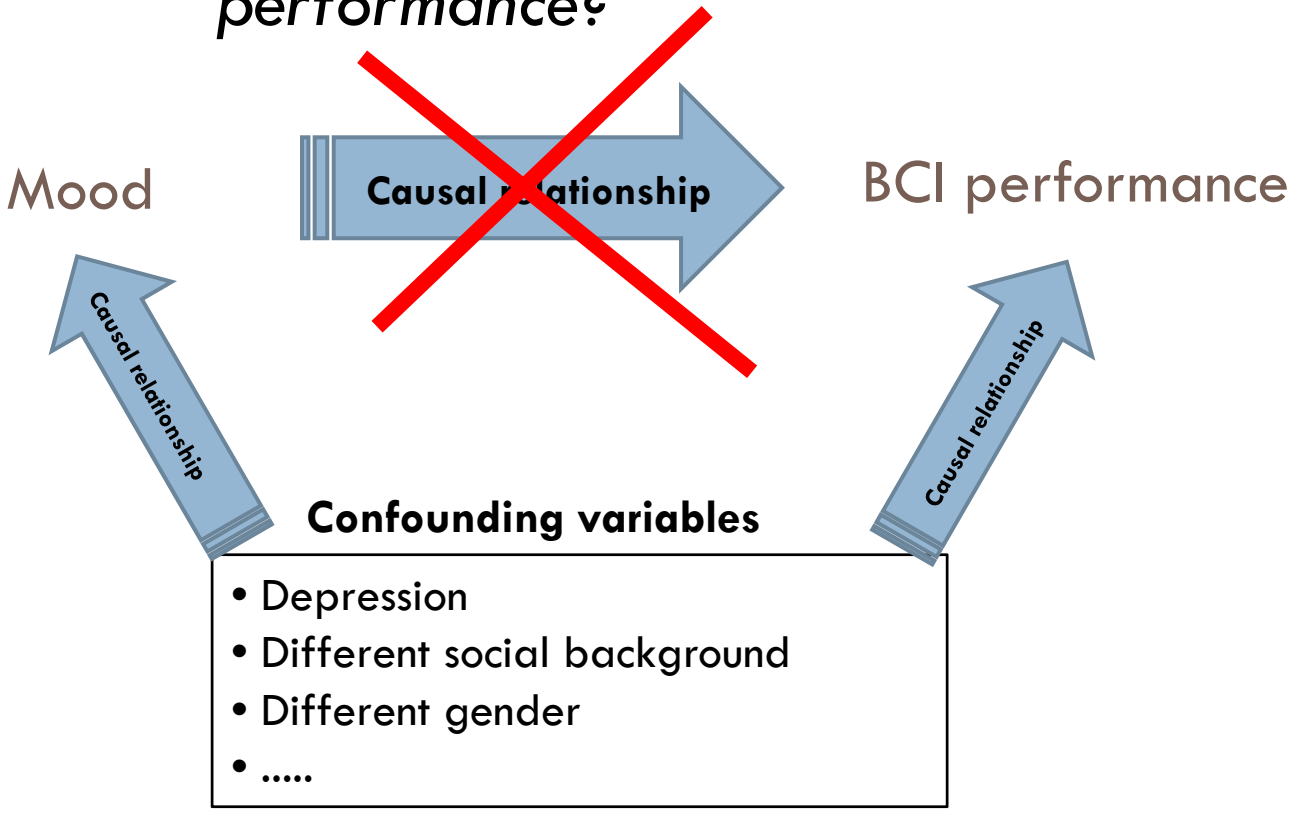
Research design: Relational research/Quasi-experiment: Example

„How does mood (e.g., happyness, sadness) affect BCI performance?“



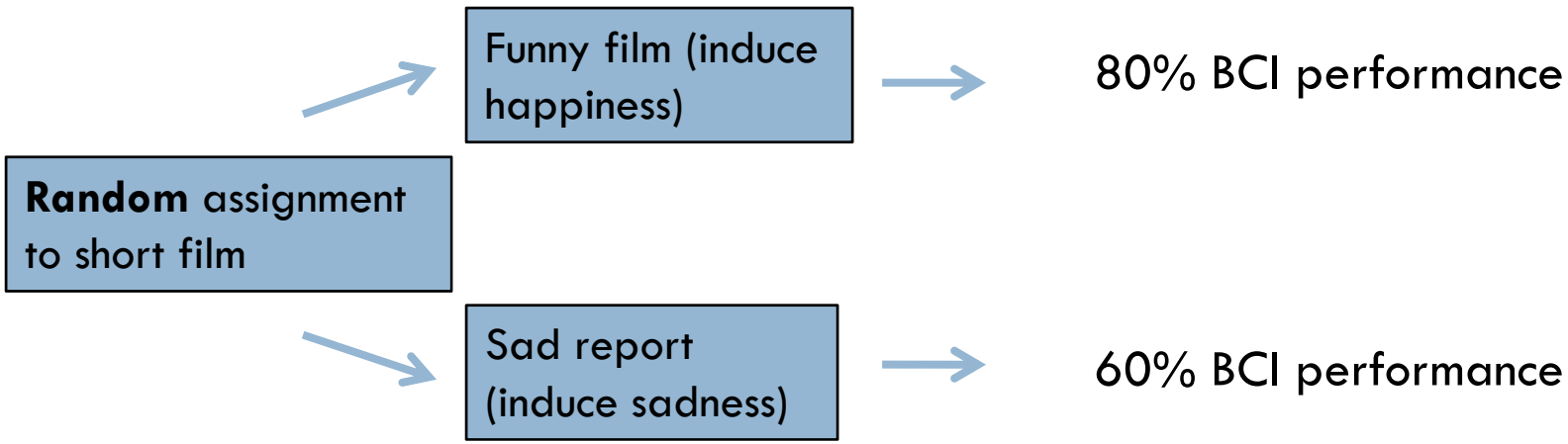
Research design: Relational research/Quasi-experiment: Example

„How does mood (e.g., happyness, sadness) affect BCI performance?“



Research design: From quasi-experiment to experiment: Example

„How does mood (e.g., happyness, sadness) affect BCI performance?“



Research design: Experiment vs quasi-experiment

(Hypothetical) research question: *What is the influence of meditation on BCI performance?*

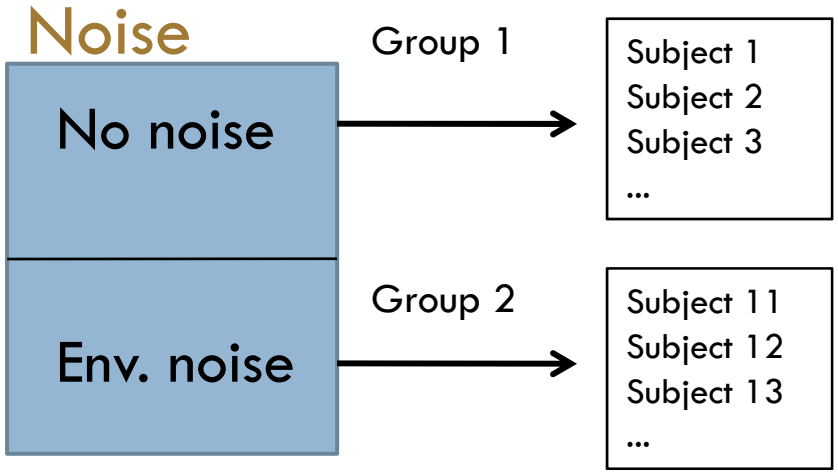
- Quasi-experiment
  - Group 1: People who meditate
  - Group 2: People who don‘t meditate
- ➔ passive observation of the factor „meditation“ (dependent)
- ➔ no causal relationship, only correlation

- Experiment
  - People who did not meditate yet, randomly assigned to:
  - Group 1: Meditate 15 min/day
  - Group 2: Do not meditate
- ➔ active manipulation of the factor „meditation“ (independent)
- ➔ causal relationship

Research design: Between/within subjects design

How to apply the independent variable?

Between-subjects design



Within-subjects design

Subject 1	No noise	Env. noise
Subject 2	No noise	Env. noise
Subject 3	No noise	Env. noise
...	No noise	Env. noise

Disadvantages

- Groups might differ
- If there are N subjects, you have only N/2 subjects for each subcondition (with 2 subconditions)
- The order of the blocks can be important
- There might be interaction between the subconditions

Experimental design: Between-subjects

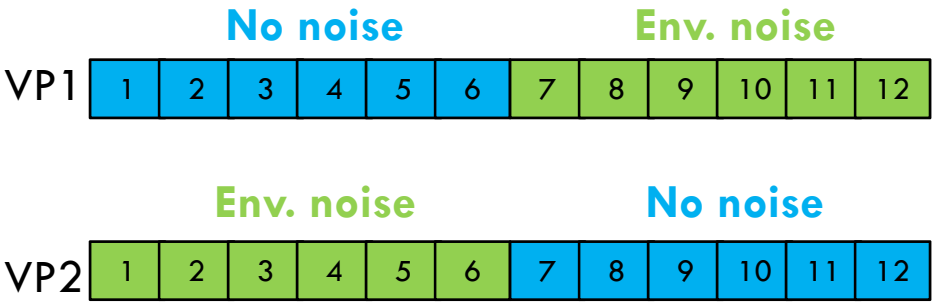
- (True) experiment → Random assignment of subjects to groups
- Patient studies: ALS patients vs. Healthy control group → relational studies (quasi-experiments)
- Use matching to reduce the number of confounding variables

Patient
Health status: ALS
Age group: 30-40y
Education: academic
etc

(Matched) control
Health status: healthy
Age group: 30-40y
Education: academic
etc

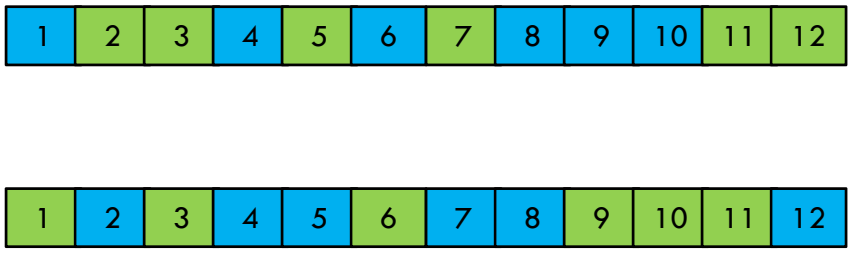
Experimental design: Within-subjects

Blocked trials



- Take into account order effects
  - Counterbalancing
  - Randomization
- ➔ Expectancy effects?

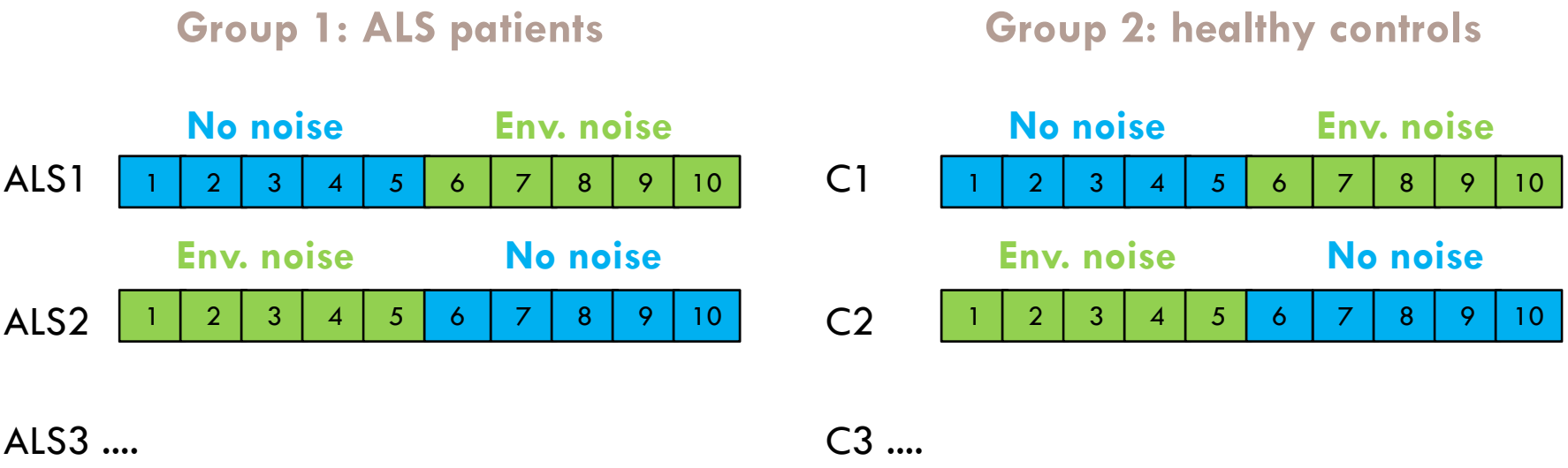
Randomized trials



- Reduces effects of expectancy
- But does not always make sense

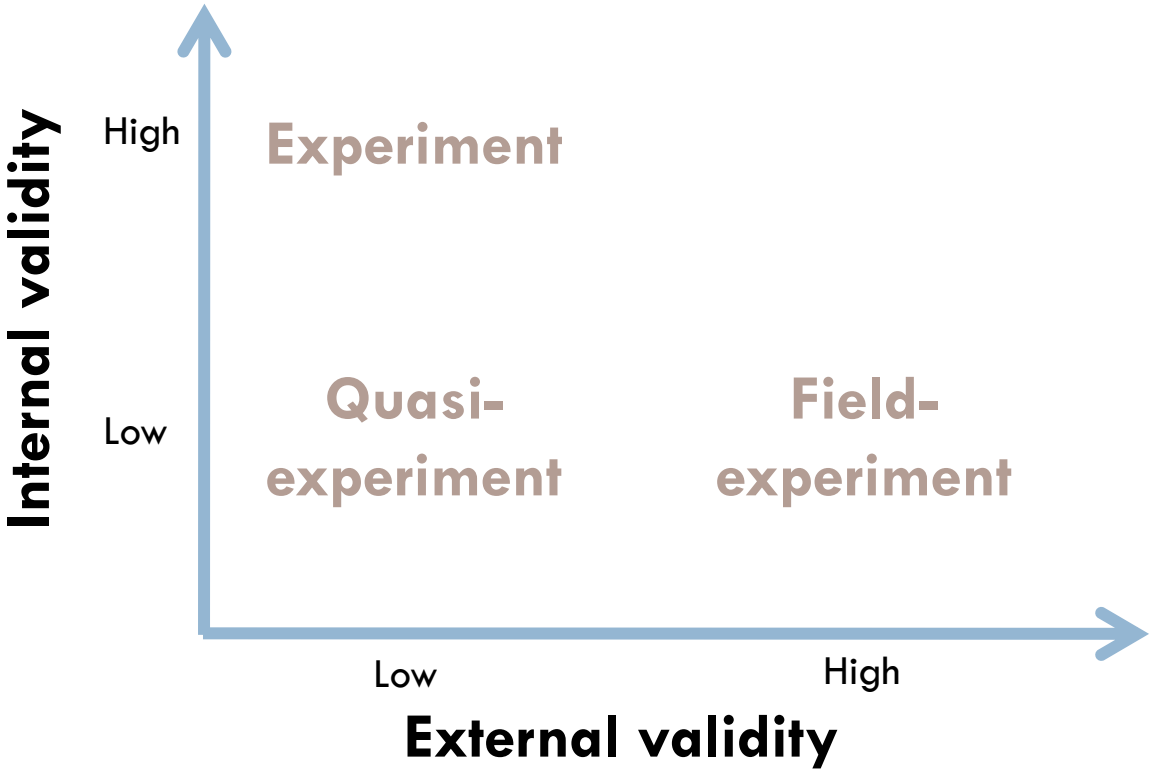
Experimental design: Mixed-design

- Both designs can be mixed: Example with a 2 factor design
- Factor 1: ALS (ALS patients, healthy subjects) ➔ between-subjects
- Factor 2: Noise (on,off) ➔ within-subjects



Experimental design: Validity

- **Internal validity:** The validity of (causal) inferences in scientific studies
- **External (aka ecological) validity:** The validity of the generalization of inferences in scientific studies (to the whole population)



Experimental design: Increase validity

How to increase internal validity

- Between-subjects design: Random assignment or matched groups
- Within-subjects design: Counterbalancing or randomization
- Elimination or constantness of (potentially) confounding variables
  - Artificial environment (constant lighting, constant noise)
  - Well-defined or automatized experimental procedure
  - Good operationalization of your constructs (next section)

How to increase external validity

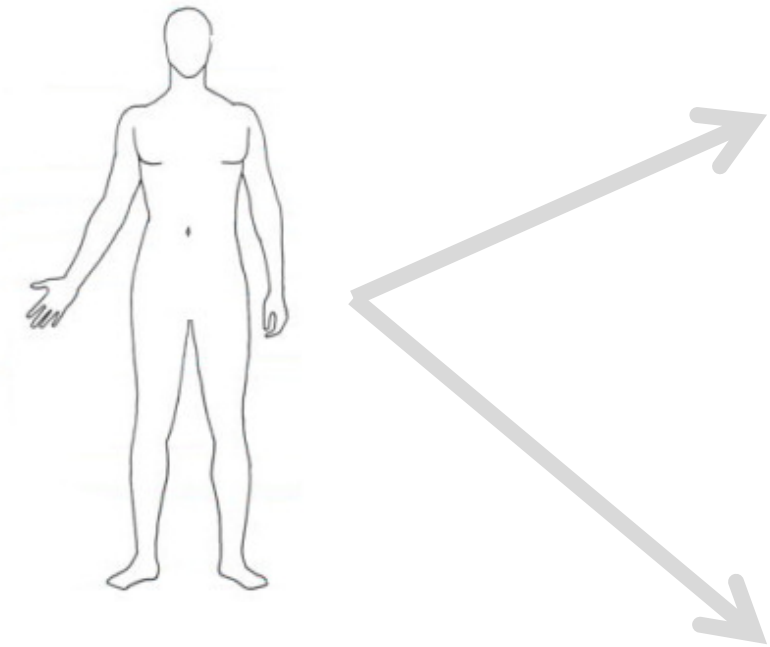
- Aselect group ➔ subjects of different age, different social background etc.
- Natural(istic) environment

Unfortunately, increases in internal validity usually cause decreases in external validity and vice versa

Overview: 3. Measures / Messinstrumente

- Measuring human function
- Operationalization
- Quality of a measure
  - Objectivity
  - Reliability
  - Validity

Measures: Human function



Overt/behavioral measures

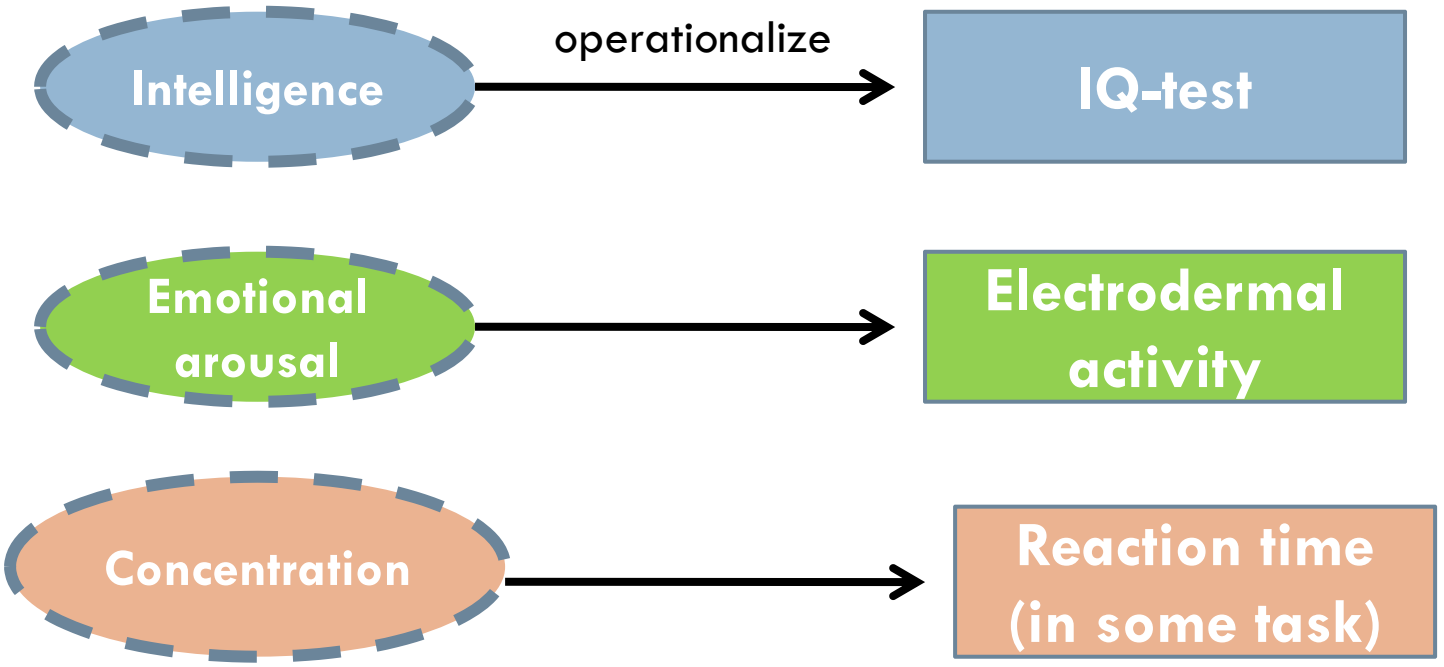
- Psychophysical measures
  - Reaction times (RT)
  - Accuracy
- Subjective report
  - Questionnaire
  - Interview
- Observation

Physiological measures

- Brain activity (EEG, MEG, ...)
- Skin conductance
- EMG
- Derived measures (e.g., classifier output)

Measures: Operationalization

- Process of translating (an aspect of) a psychological construct (latent variable) into a measurable variable



Measures: Quality-> objectivity

- **Objectivity**
  - The measurement should be independent of the person who is making the measurement (Versuchsleiter)
- Objectivity can be violated due to
  - Variability in the behavior of an individual experimenter/Versuchsleiter
  - Variability between multiple experimenters/Versuchsleiter

Measures: Increasing objectivity

- Strict experimental procedure (how-to's)
- The role of the experimenter should be minimized
- Constant order of submeasurements (e.g., Vormessungen and Hauptmessung)
- Written instruction!!
- ➔ standardized, automatized experiment

Measures: Quality-> reliability

- **Reliability**
  - The precision of a measure
  - The part of the variance which is explained by real differences and not by measurement errors

Measures: How to measure reliability

- **Test-retest reliability**
  - Repetition of the measurement should yield identical results (if the construct you measure is stable)
  - → sometimes violated due to memory effects
- **Parallel-test reliability**
  - A comparable measure should yield comparable results
- **Split-half reliability**
  - Split in two halves → same result?
- **Interrater reliability** (observation)

Measures: Quality -> validity

- **Validity** (aka construct validity)
  - The extent to which a test measures the construct that it should measure
  - ➔ the quality of the operationalization, crucial for making (causal) inferences
- Validity is not easily given by a single statistic, but by a body of research demonstrating that a measure is related to the psychological construct it should measure

Measures: How to measure validity

- **Content validity**
  - The different parts of the measurement (e.g., items in a questionnaire) cover the whole range of aspects of a construct → e.g. IQ-score and creativity
  - → Sampling analogy
  - → high correlation with other (accepted) tests of the construct
- **Divergent validity**
  - Two measures should not be correlated if their underlying constructs are unrelated → e.g., intelligence & emotional arousal