Vorlesung Fortgeschrittene Softwaretechnik

Wintersemester 2024/25

Prof. Dr. Stephan Diehl

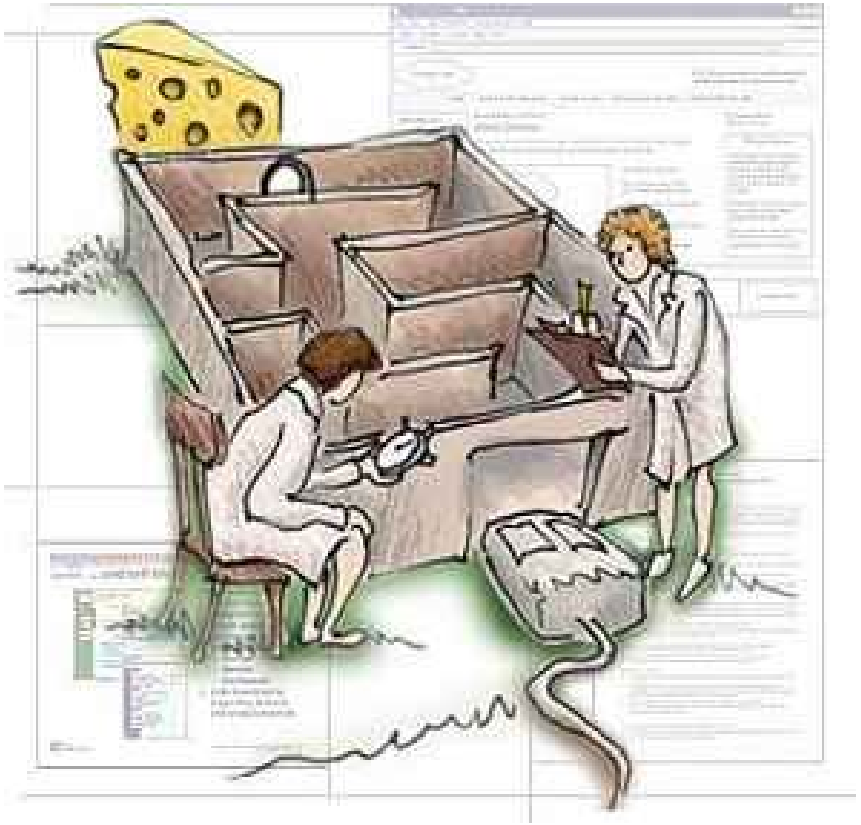
Informatik

Universität Trier

PLACEBO CHRISTMAS

Heute

- Research Design
- Quantitative Analyse



Goals of Research

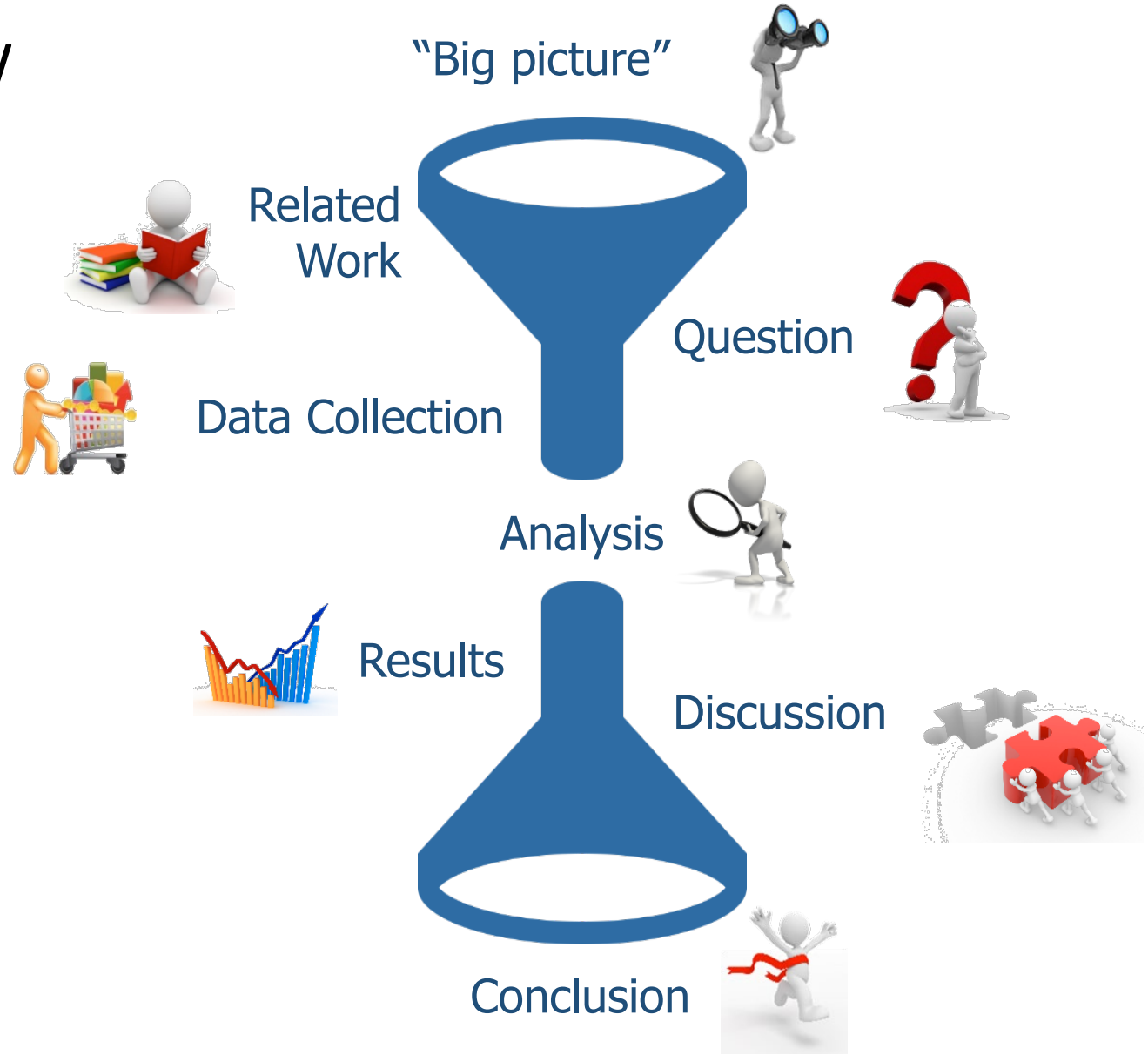
- Observe
- Describe
- Predict
- Explain
- Determine causes of

} Phenomena $\Rightarrow$ **Expand knowledge**

Examples from Software Engineering research:

Does tool XY improve program comprehension?
How do software developers interact in a pair-programming setting?
Is software implemented in statically typed languages less error-prone than software implemented in dynamically typed languages?

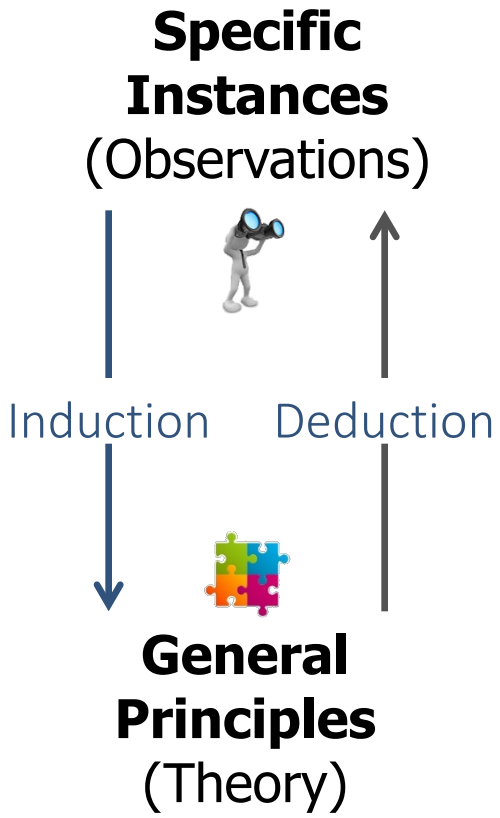
Scientific Workflow



https://explorable.com/research-basics

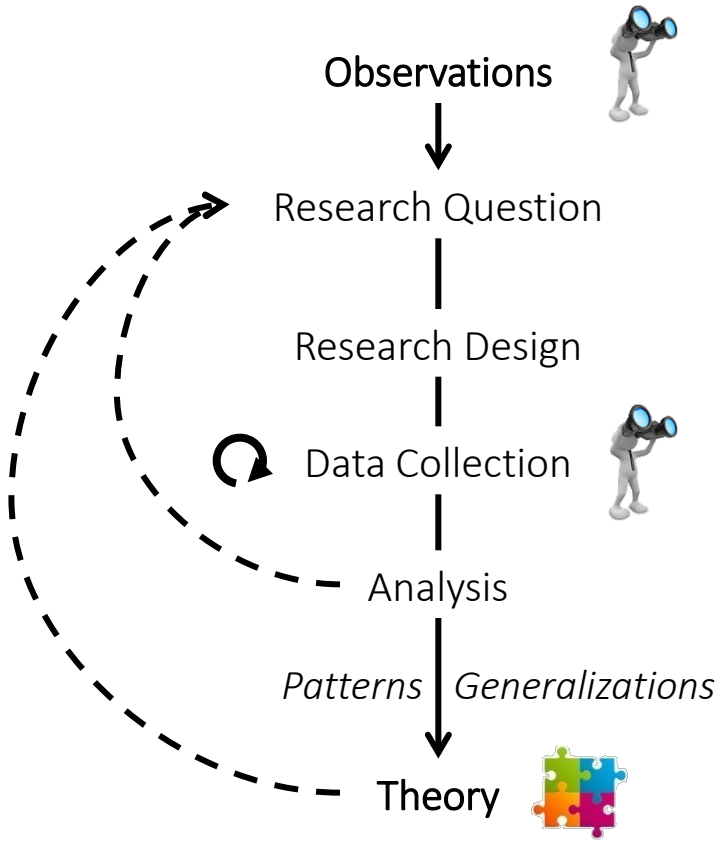
What is „Empirical“ Research?

Empirical: „originating in or based on **observation** or experience“ Merriam-Webster



- There is no “truth” in science, just evidence we trust
- Important property of theories: **falsifiability** (possibility to prove a theory to be false)
- Example: “*All swans are white*.” vs. “*There is a god.*”
- **Karl Popper:**
  - Theory only scientific if falsifiable
  - Theories cannot be verified, but can be accepted as long as attempts to falsify fail

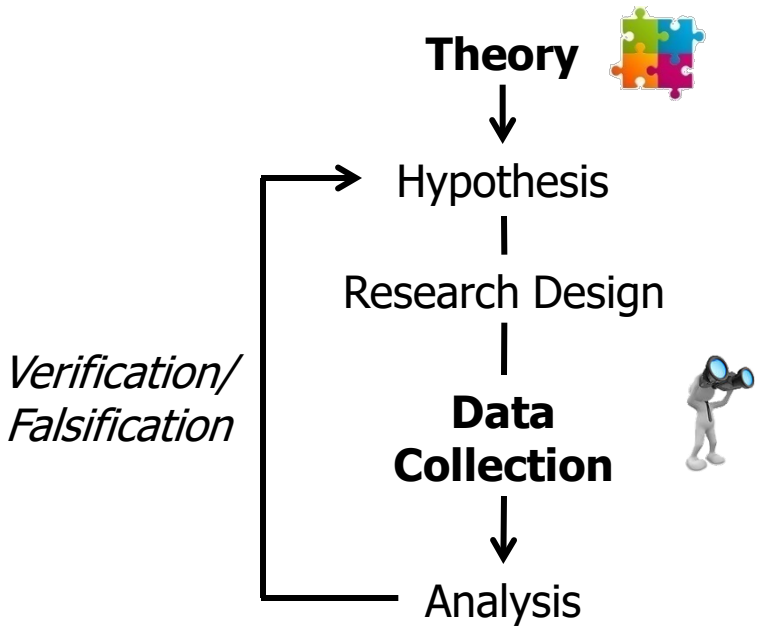
Inductive Research



- **Exploring** new phenomena
- Open-ended
- Process-oriented
- Focus on **qualitative data**
  (but quantitative data may also be used)
- Generating new theory from data
  (*grounded theory*)
- **Reiterating** until saturation
  (*constant comparison*)

Example: *How do software developers interact in a pair-programming setting?*

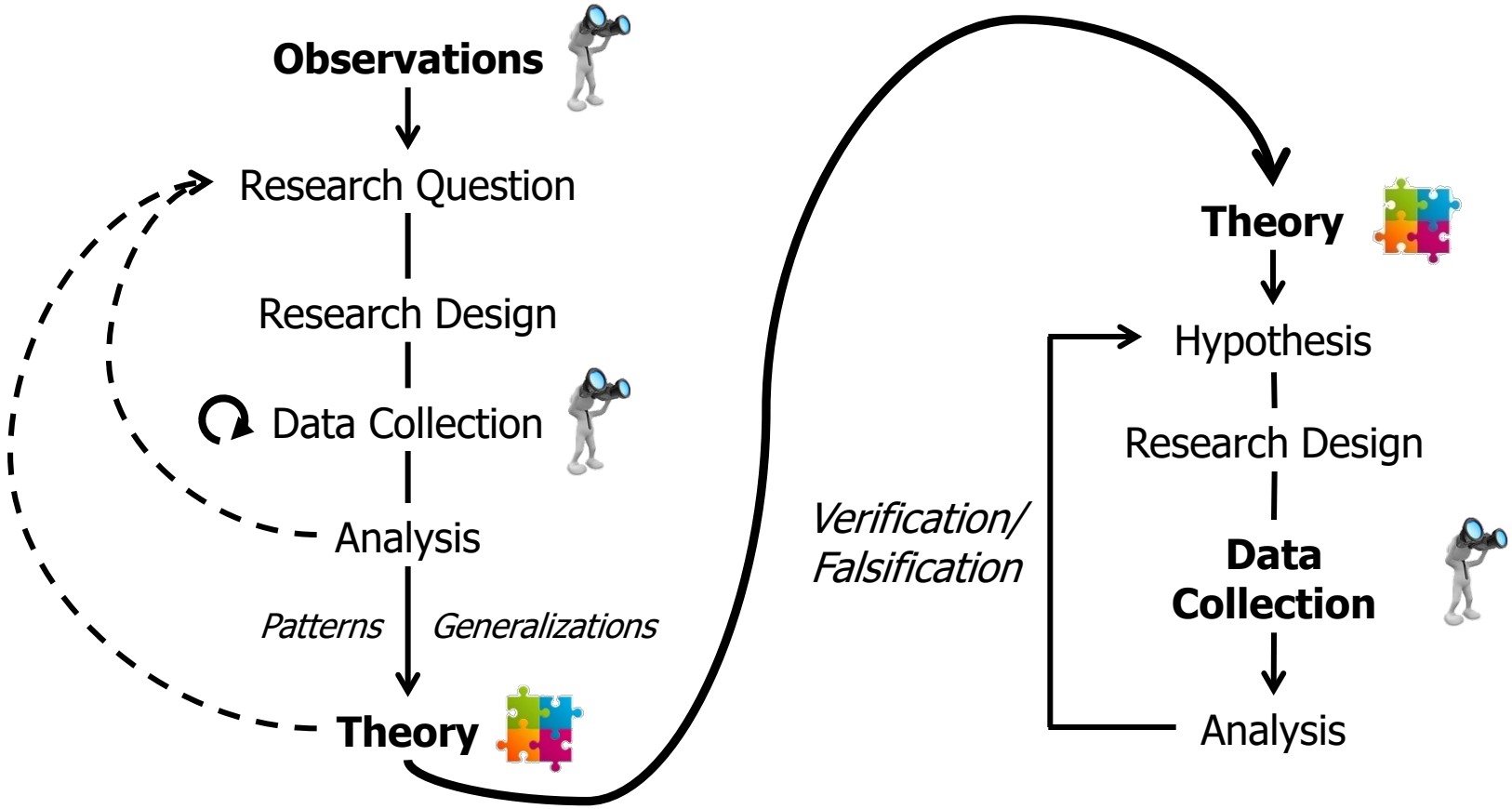
Deductive Research



- **Confirmatory** (testing hypotheses)
- Outcome-oriented (causality)
- Focus on **quantitative data** (often controlled experiments)
- Operationalization (“Messbarmachung”)
- **Statistical methods** (descriptive statistics or correlations)

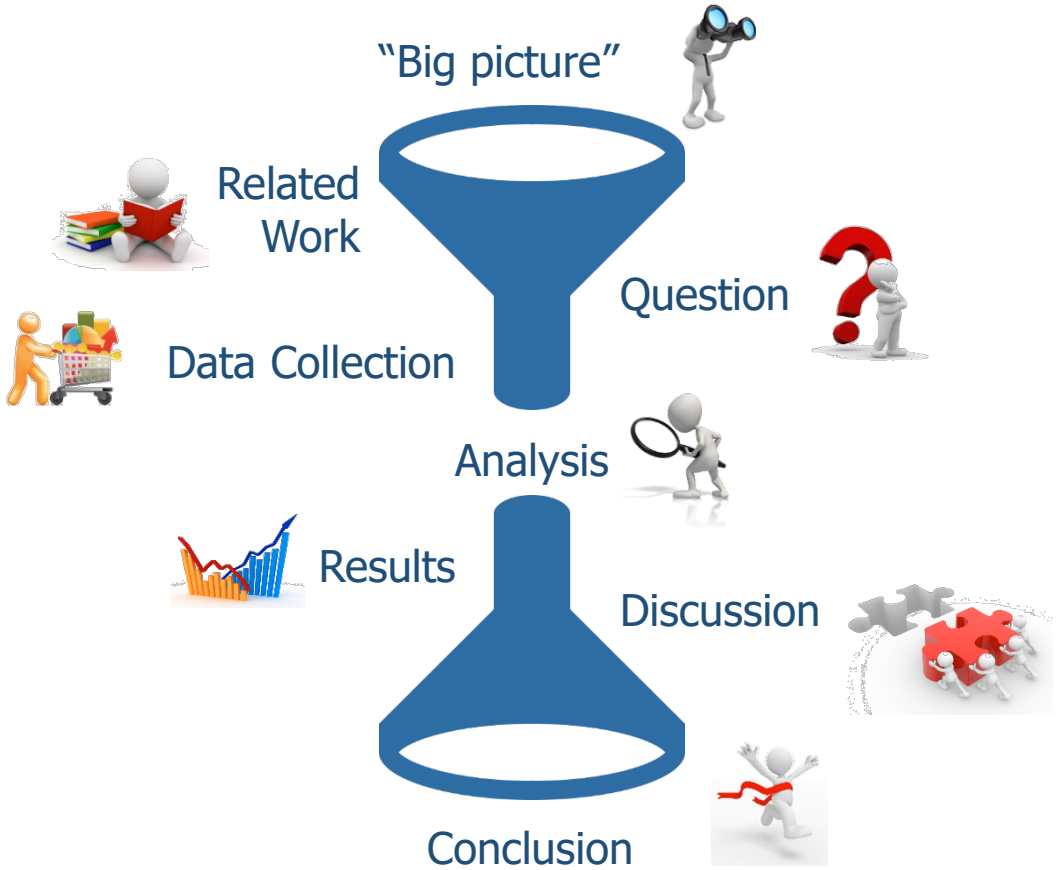
Example: (To what extent …) *Does tool XY improve program comprehension?*

„Mixed Methods“



Properties of „Good“ Research

- Reproducibility of Research



Provenance tracking

Publish results, raw data, and provenance information

Properties of „Good“ Research

Reliability: “the extent to which an experiment, test, or measuring procedure yields the same results on repeated trials“ [Merriam-Webster.com]

Validity: “the extent to which conclusions drawn from research provide an accurate description of what happened or a correct explanation of what happens and why”

[methods.sagepub.com]

Transparency: “honest and open, not secretive” [Merriam-Webster.com]

Reproducibility: “the ability of an entire experiment or study to be duplicated, either by the same researcher or by someone else working independently.”

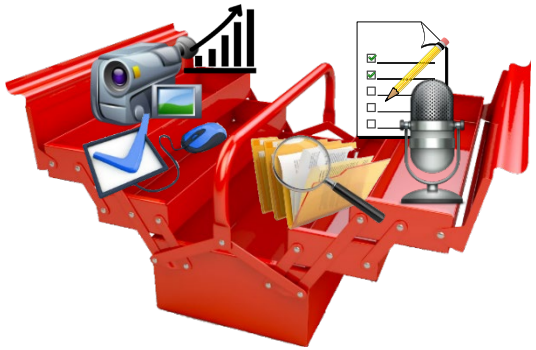
[en.wikipedia.org]

Ethical Issues: “following accepted rules of behavior”

[Merriam-Webster.com]

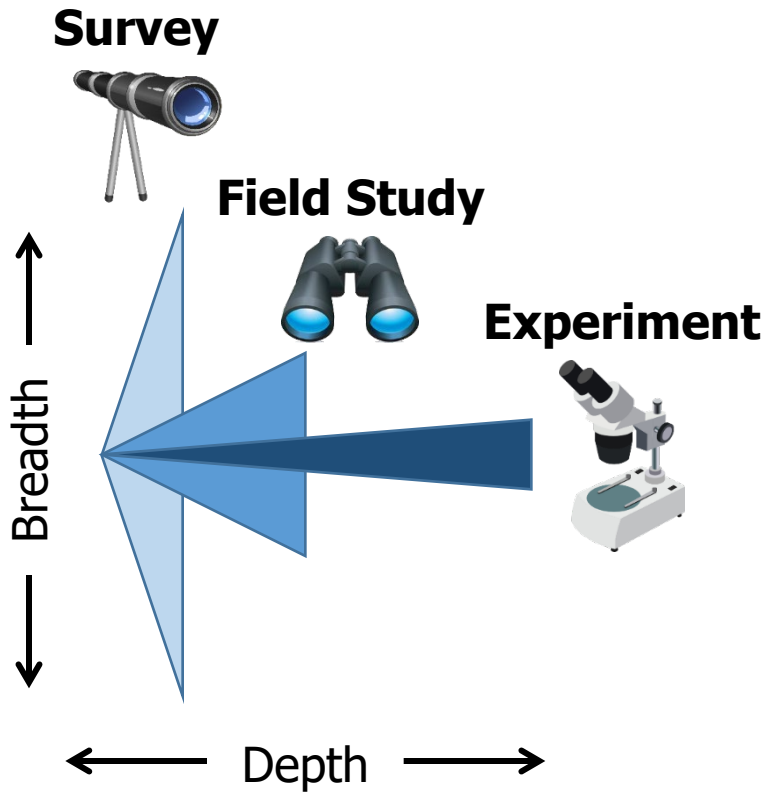
Good research should be well planned and documented!

Research Methods



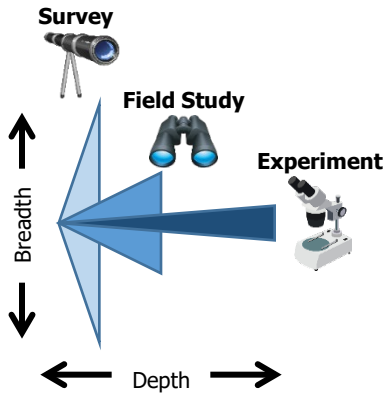
Basic Research Methods

→ vorherige Vorlesung



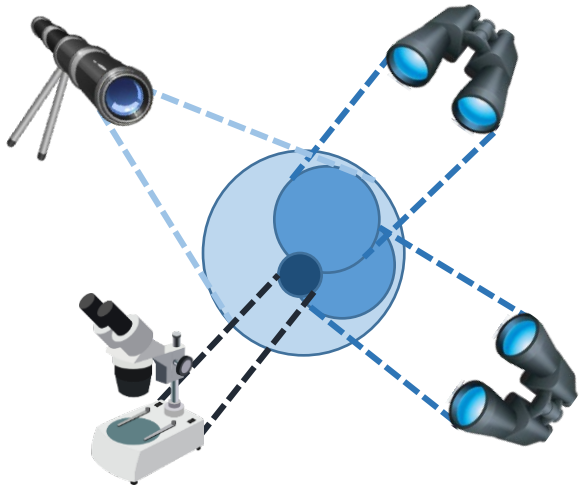
Mixed-Methods Research

- Remember goals of research:
  - **Exploration:** Understanding a problem in more detail, making it known to a broader audience, inform future research.
  - **Explanation:** Developing a theory based on previous explorations and exiting theories.
  - **Prediction:** Use a theory to build causal relationships to make predictions about phenomena.
  - **Demonstration:** Show that something can be done in a particular way (e.g. new tool).
- Each instrument can only help to achieve some of these goals.
- **Combine different research instruments** to gain both greater breadth and depth of insights.
- **Example:** Start with exploratory field study and validate results with large online survey.



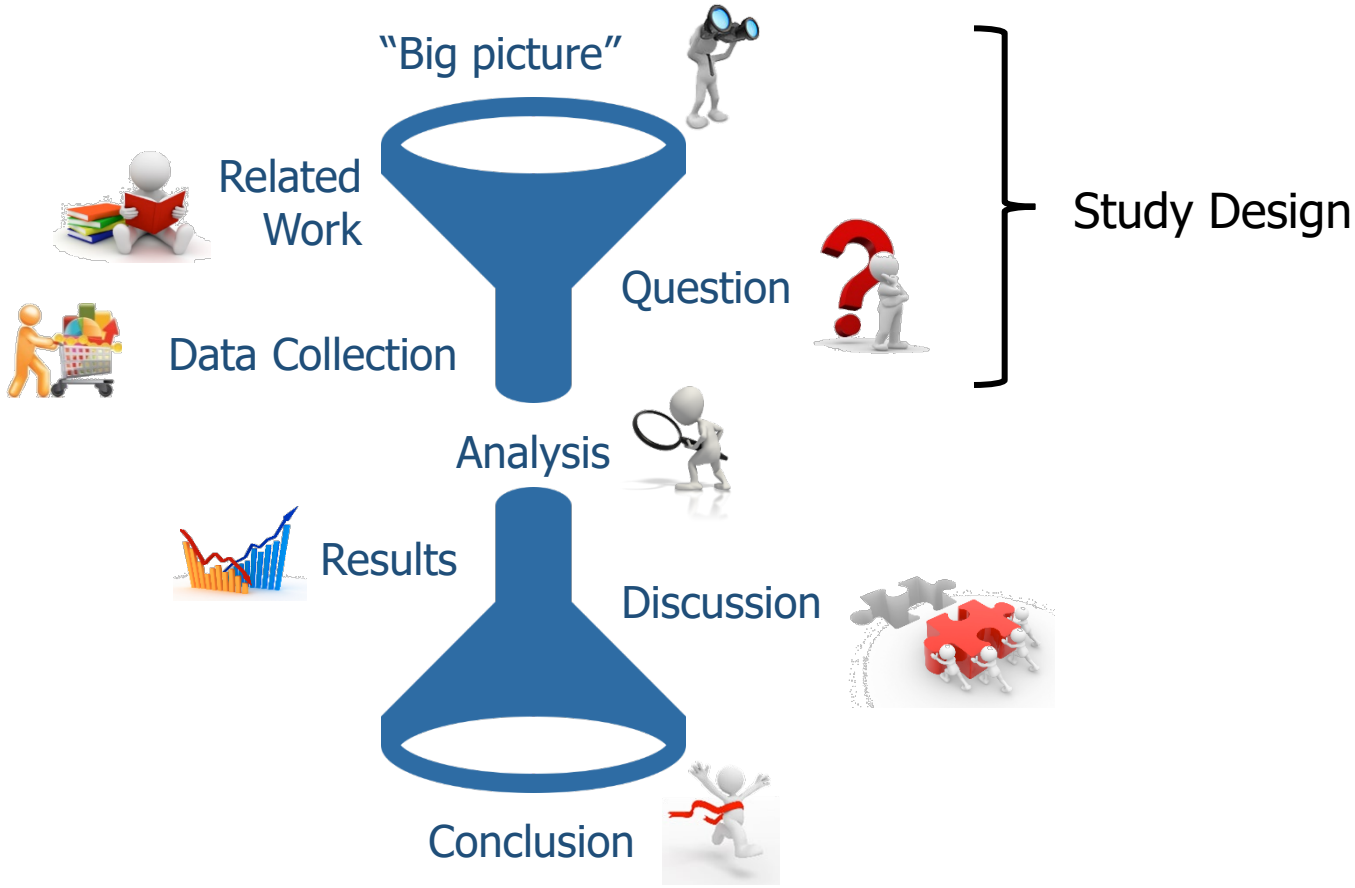
Triangulation

- Increase validity of research by **studying a phenomenon from several points of view**
- **Cross-validation** from two or more sources:
  - Different data sources
  - Different aspects of the same phenomenon
  - Different research instruments
  - Different researchers
- Term “triangulation” is an analogy to land surveying



Research Design

Scientific Workflow



Finding a Research Question

- A research question has to be **precise, operational**, small enough to be **feasible**, and should be **relevant**. (Alternative: SMART: specific, measurable, attainable, realistic, time bound)
- **Iterative** process
- Literature review, exploratory research, and discussion.
- Template for formulation of research goal*:*

We study **<property>** of **<object>** [in **<setting>**] [under **<treatment>**] [in order to **<purpose>**] by **<method>**.

In order to improve software development practices, we study effectiveness of pair-programming with students at university having single programmers and pairs of programmers in a laboratory experiment.

Finding a Research Question

- Alternative: **PICOC**
  - Population: Who?
  - Intervention: What or How?
  - Comparison: Compared to what? (possibly N/A)
  - Outcome: What are you trying to accomplish / improve?
  - Context: In what kind of organization / circumstances?

Criteria	Element
Population	Final-year undergraduate students in Software Systems Engineering with work experience, e.g., work placements or internships
Intervention	They will evaluate the proposed modelling language (i.e., DMML) and automated reasoning
Comparison	DMML will be compared with the i^* modelling framework
Outcome	It is expected that the use of DMML in designing digital motivation would be more effective, efficient, useful, and satisfactory in comparison to other goal-oriented modelling languages
Context	The experiment would be carried on in the context of a business information system

https://cebma.org/faq/what-is-a-picoc/

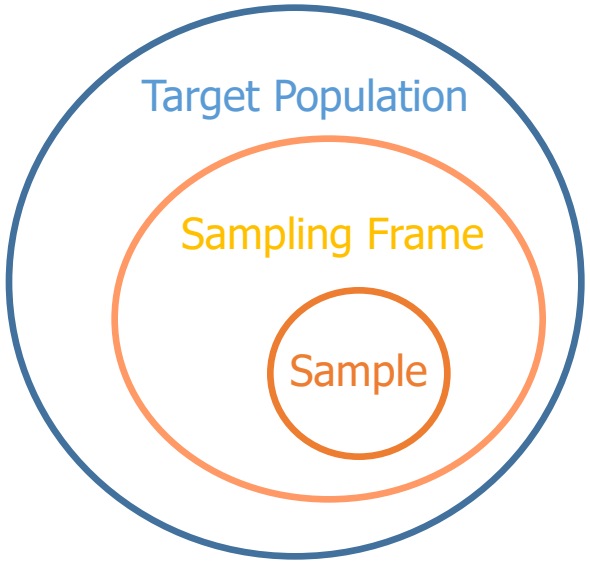
Population

What or who is being studied?

- **Population:** Complete set of items that share at least one property in common (e.g. all software developers in the world)
- **Subpopulation:** Subset of a population that shares one or more additional properties (e.g. all German software developers)
- **Target population:** Subset of a population that the researcher wants to draw conclusions about with his/her research (e.g. all software developers in the world or all German Java software developers)
- **Sample:** Subset of a population that does not require to share additional properties. Usually drawn from the population using a **sampling frame** (e.g. list of all software developers working on a project).
- **Unit of observation:** Entity described by the data that one analyzes (e.g. individual developers or commits)
- **Unit of analysis:** Entity being analyzed in the study (e.g. software development teams or the development activity of developers)

Sampling

- **Recruiting**: Getting people to participate in the study, selecting artifacts to study
- Probability vs. non-probability sampling
- Sampling influences **external validity** (generalization of results)
- *“Convenience sampling”* and *“snowball sampling”* common

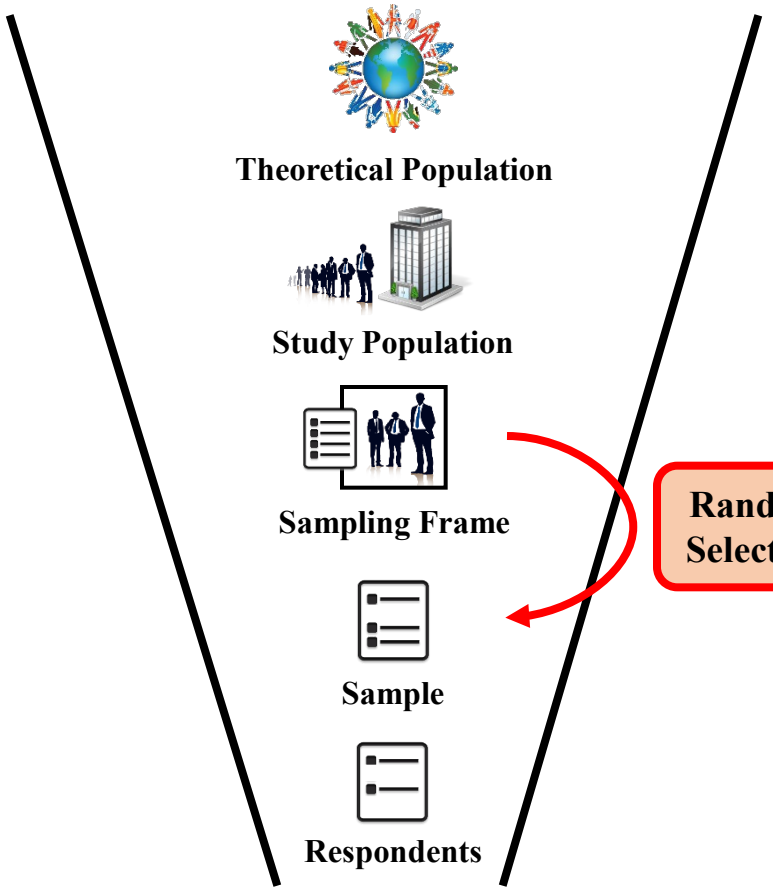


“Real-world” performance bugs

Performance bugs documented as issues in open source GitHub projects

Random Sampling

Sampling: Ideal Scenario



Who do you want to generalize to?

What populations can you get access to?

How can you get access to them?

Random Selection

Who is in your study?

Who participated?

Based on: http://www.socialresearchmethods.net/kb/sampterm.php

Sampling: Common Scenario

Main problem: Availability of suitable sampling frames, reachability of participants.

➔Reliance on available subjects:
convenience sampling, snowball sampling

➔Likely leads to **biased samples**:

- Human participants: Self-selection bias
- Researchers contact people from their own social and cultural group / select artifacts that they already know
- **Limited generalizability**

Strategies:

- (Try to) select broad cross-section of **the target population**
- **Clear description** of sampling approach and participants/selected artifacts
- Take care not to overgeneralize
- Alert readers to the **limitations**

Study Design

Research Question

Operationalization

Theoretical concept → observable or measurable variables/categories.

Select appropriate method(s)

Plan and conduct study(s)

Operationalization and method selection are of course connected.

Operationalization: Measurement

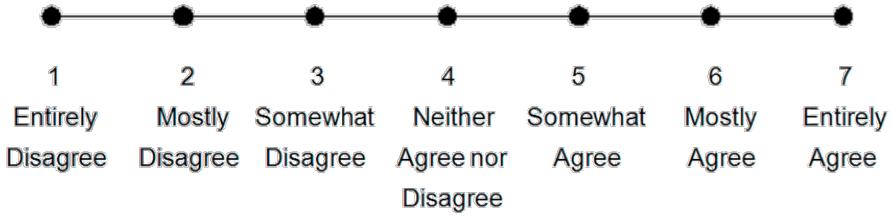
- *„Measure what is measurable, and make measurable what is not so.“* (attributed to Galileo Galilei)
- Measurement is central to the scientific method
- Mapping from studied object to a scale, e.g.

$$\text{height} : \text{people} \to \mathbb{R}^{\geq 0}$$
$$\text{sex} : \text{people} \to \{\text{male}, \text{female}, \text{intersex}\}$$

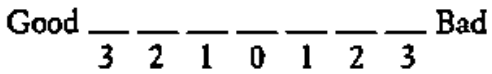
- **Scale types:** nominal (gender), ordinal (very unhappy, unhappy, …, very happy), interval (temp. °C), ratio scales (temp. K, height)
- **Examples from SE:** software metrics like *bugs per line of code* or *cyclomatic complexity*

Objective vs. Subjective Measures

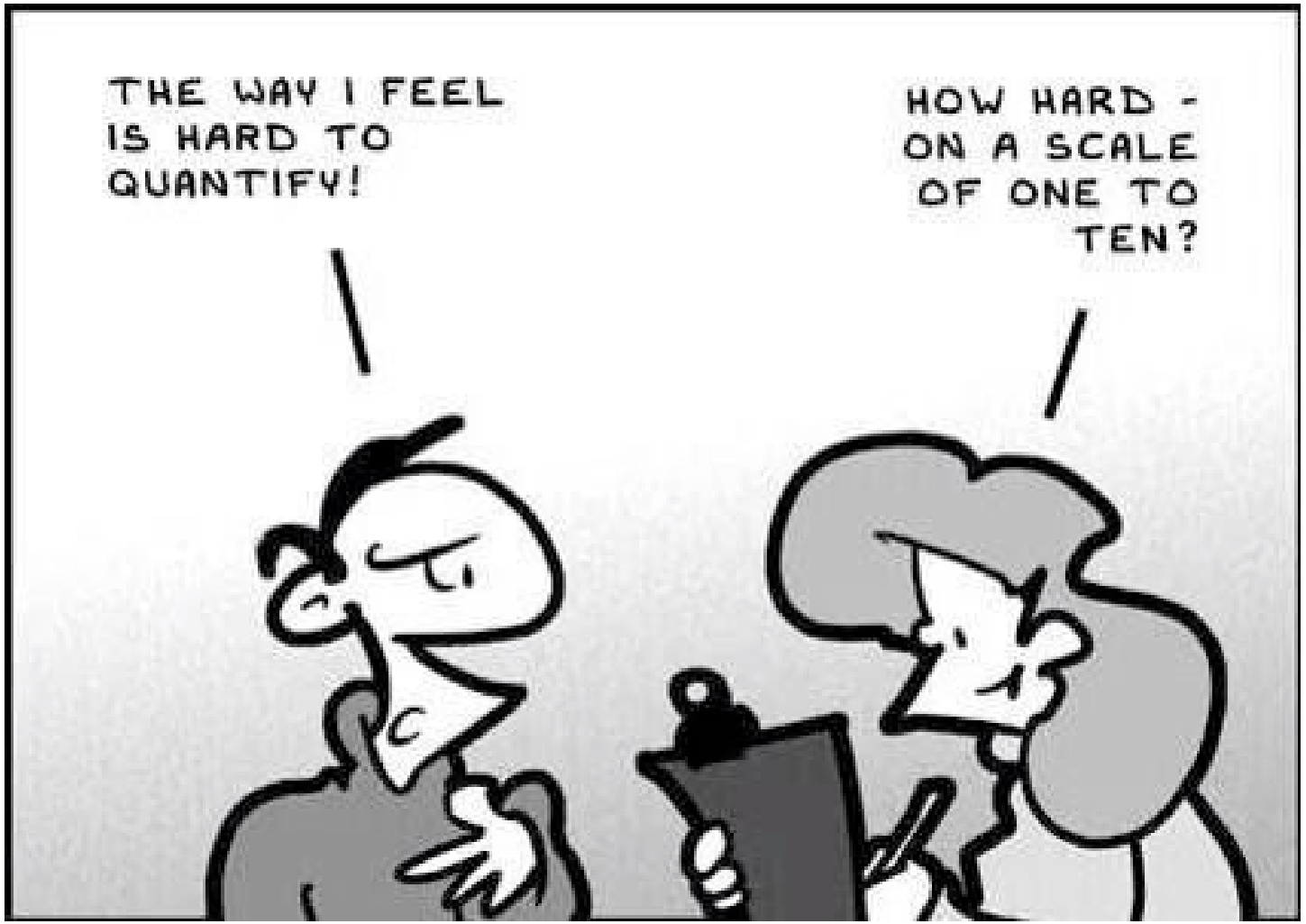
- Software artifacts: objective measures exist (e.g. software metrics)
- Human subjects:
  - Measuring more difficult
  - **Few objective measures** can be obtained easily (e.g. demographic data)
  - Many properties cannot be measured directly (e.g. cognitive load, program comprehension)
  - **Self-reports**: many cognitive biases involved, but strategies exist
  - **Likert scale**: participant is asked to rate his/her agreement with a statement



  - **Semantic differential: "**measures people's reactions to stimulus words and concepts in terms of ratings on bipolar scales defined with contrasting adjectives at each end" [Heise70]



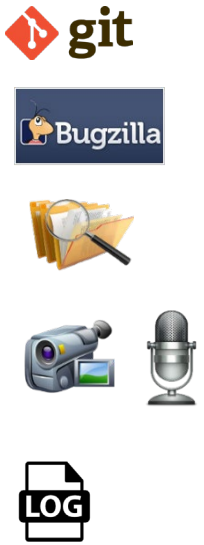
Objective vs. Subjective Measures



„Measuring“ Behavior

- **Many sources** for behavioral data can be utilized e.g. to analyze usability of a tool or search for usage patterns

- Sources include:
  - Source code repository (commit history)
  - Bug tracking system
  - Electronic traces of social interaction (e.g. mailing list, online discussions, tweets)
  - Video recordings and audio recordings of actual behavior (e.g. screen recordings of tool usage, pair programming sessions)
  - Log files (e.g. from instrumented tools)



- **Analysis can be difficult** and time-consuming (→ lecture “Qualitative Analysis”)

From Research Question to Study Design

Research Question

Operationalization

Select appropriate method(s)

What methods yield observations that can be analyzed to answer the RQ?

Plan and conduct study(s)

Scientific Workflow