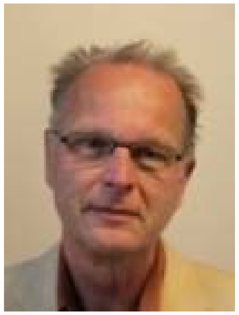
# Theories in Empirical Software Engineering

Roel Wieringa

Sidekicks:
Daniel Méndez
Lutz Prechelt

# Who are we?



**Roel Wieringa**

University of Twente, Germany

http://wwwhome.ewi.utwente.nl/~roelw/

**Lutz Prechelt**

FU Berlin

http://www.mi.fu-berlin.de/w/Main/LutzPrechelt

**Daniel Méndez**

TU München

http://www4.in.tum.de/~mendezfe/

# Who are you?

Quick round

- Who are you?
- What are your experiences in conducting empirical studies?
- What are your expectations?

# What do you think?

**Why do we need scientific theories in software engineering?**

4. Methodology (the study of research methods)
   a. Notion of conceptual framework; statements about them
   b. Notion of generalization; statements about them

Looking at research from the sky

3. Theory (statement about many research results)
   a. Conceptual framework
   b. Generalization

General knowledge is the gold we are after

2. Research questions (what, how, when where, …., why) aimed at generalizable knowledge, research method, and research result

Hard work to grow knowledge

1. Practice domain: SW, methods, tools, processes (as is / to be)

Grass roots

- Everything on the slides in this talk , except the examples, is at level 4.
  - *The examples on these slides contain explicit level indications.*
- The separate example slides report about research that contains 2 and 3.
- The reported research studies some aspect of 1.

# Agenda

| Time | Topic |
|---|---|
| 09:00 – 10:30 | Opening and Introduction |
| 10:30 – 11:00 | Coffee break |
| 11:00 – 12:30 | Inferring Theories from Data |
| 12:30 – 13:30 | Lunch |
| 13:30 – 15:00 | Designing Research based on Theories |
| 15:00 – 15:30 | Coffee break |
| 15:30 – 16:30 | Hands-on Working Session and Q&A |
| 16:30 – 17:00 | Wrap up (all) |

# What is a Scientific Theory

# Scientific theories

- A **theory** is a belief that there is a pattern in phenomena
- A **scientific** theory is a theory that
  - Has survived tests against experience
    - Observation, measurement
    - Possibly experiment, simulation, trials
  - Has survived criticism by critical peers
    - Anonymous peer review
    - Publication
    - Replication

***Examples (level 3)***

- *Theory of cognitive dissonance*
- *Theory of electromagnetism*
- *The Balance theorem in social networks*
- *Theories X, Y, Z, and W of (project) management*
- *Technology Acceptance Model*

- Hannay et al. A Systematic "Review of Theory Use in Software Engineering Experiments". IEEE TOSEM 33(2), February 2007
- Lim et al. "Theories Used in Information Systems Research: Identifying Theory Networks in Leading IS Journals"./ ICIC 2009, paper 91.
- Non-examples
  - *Speculations based on imagination rather than fact: Conspiracy theories about who killed John Kennedy*
  - *Opinions that cannot be refuted: The Dutch lost the World Championship because they play like prima donnas*

# Design theories

- A design theory is a scientific theory about an artifact in a context


- *Vriezekolk: What is a theory*
- *Méndez: What is a theory*

# The Structure of Theories

# The structure of scientific theories

1. **Conceptual framework**
   - Constructs used to express beliefs about patterns in phenomena
   - *E.g. The concepts of beamforming, of multi-agent planning, of data location compliance. (level 3)*
2. **Generalizations**
   - stated in terms of these concepts, that express beliefs about patterns in phenomena.
   - *E.g. relation between angle of incidence and phase difference,*
   - *Statement about delay reduction on airports. (level 3)*

- Generalizations have a scope, a.k.a. target of generalization

# The structure of **design** theories

1. **Conceptual framework**
2. **Generalizations**
   – Artifact specification X Context assumptions → Effects
   – Effects satisfy a requirement to some extent

# Two kinds of conceptual structures

1. **Architectural structures**: Class of systems, components with capabilities, interactions
   - *E.g. entities, (de)composition, taxonomies, cardinality, events, processes, procedures, constraints, … (level 4)*
   - Useful for case-based research (observational case studies, case experiments, simulations, technical action research)
   - **Typically qualitative**

2. **Statistical structures**: Population, variables with probability distributions, relations among variables
   - Useful for sample-based research (surveys, statistical difference-making experiments)
   - **Typically quantitative**


- *Prechelt: What is a theory, the structure of theories*
- *Vriezekolk: The structure of theories*
- *Méndez: The structure of theories*

# The Use of Theories

# Uses of a conceptual framework

- **Framing** a problem or artifact: choosing which concepts to use
  - *Using the theory of infectuous diseases to understand a patient's symptoms*
  - *Using concepts of force & energy to understand behavior of a machine*
  - *Using concept of a coordination gatekeeper to understand a distributed SE project (all three examples at level 1)*
- **Describe** a problem or **specify** an artifact: using the concepts
- **Generalize** about the problem or artifact
- **Analyze** a problem or artifact (i.e. analyze the framework)

# Functions of generalizations

- Functions of generalizations
  - **Explanation**: explain phenomena by identifying causes, mechanisms or reasons
  - **Prediction**: state what will happen in the future
    - Design: use generalizations to justify a design choice


- *Prechelt: the use of theories*
- *Vriezekolk: the use of theories*
- *Méndez: the use of theories*

# Usability of theories

- When is a design theory

  Context assumptions X Artifact design → Effects

  usable by a practitioner?

  1. He/she is capable to **recognize** Context assumptions
  2. and to **acquire/build** Artifact under constraints of practice,
  3. effects will **indeed** occur, and
  4. He/she can **observe** this, and
  5. They will **contribute** to stakeholder goals/satisfy requirements

- Practitioner has to asses the risk that each of these fails

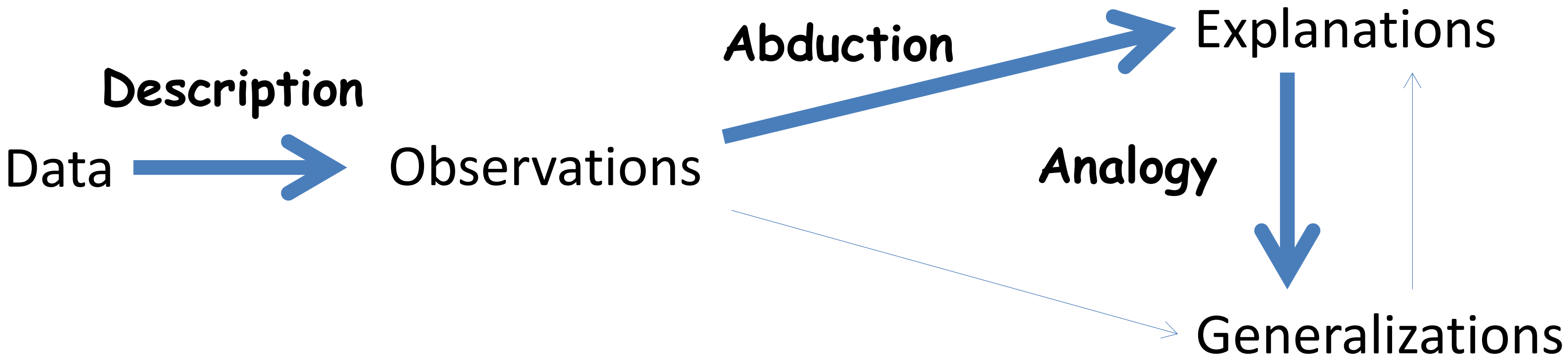

- *Prechelt: the usability of theories*
- *Vriezekolk: the usability of theories*
- *Méndez: the usability of theories*

# Agenda

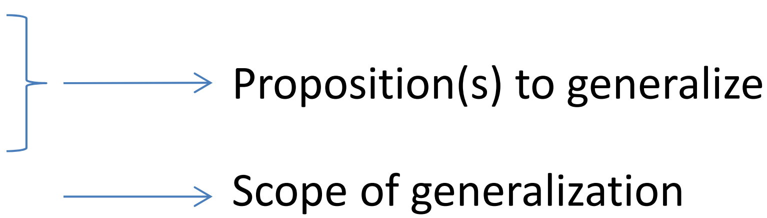
| Time | Topic |
|---|---|
| 09:00 – 10:30 | Opening and Introduction |
| 10:30 – 11:00 | Coffee break |
| 11:00 – 12:30 | Inferring Theories from Data |
| 12:30 – 13:30 | Lunch |
| 13:30 – 15:00 | Designing Research based on Theories |
| 15:00 – 15:30 | Coffee break |
| 15:30 – 16:30 | Hands-on Working Session and Q&A |
| 16:30 – 17:00 | Wrap up (all) |

# Scientific Inference

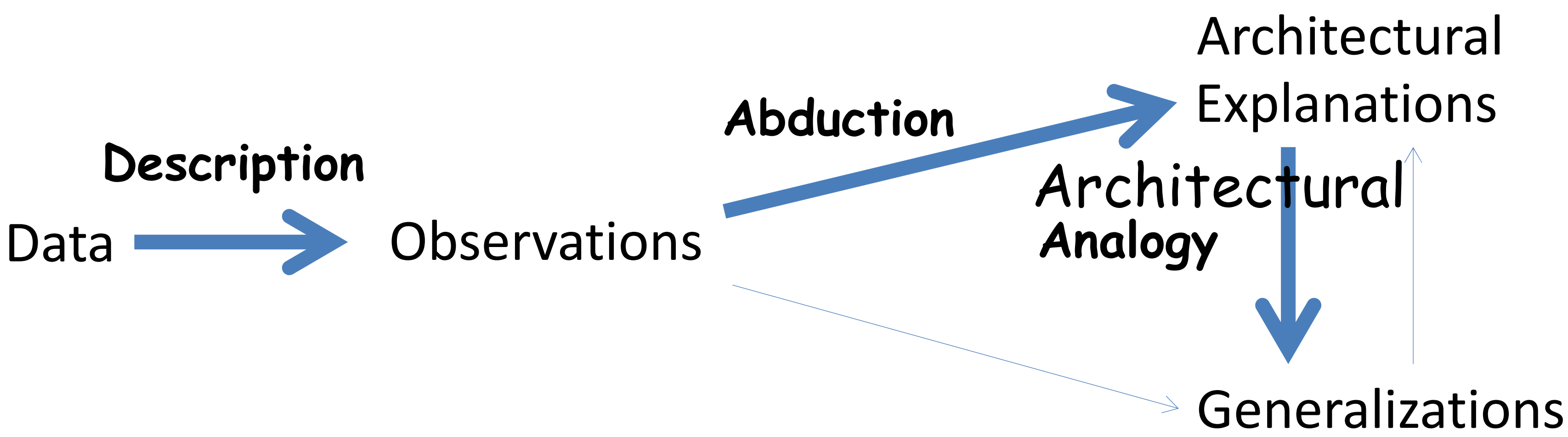
# Case-based inference



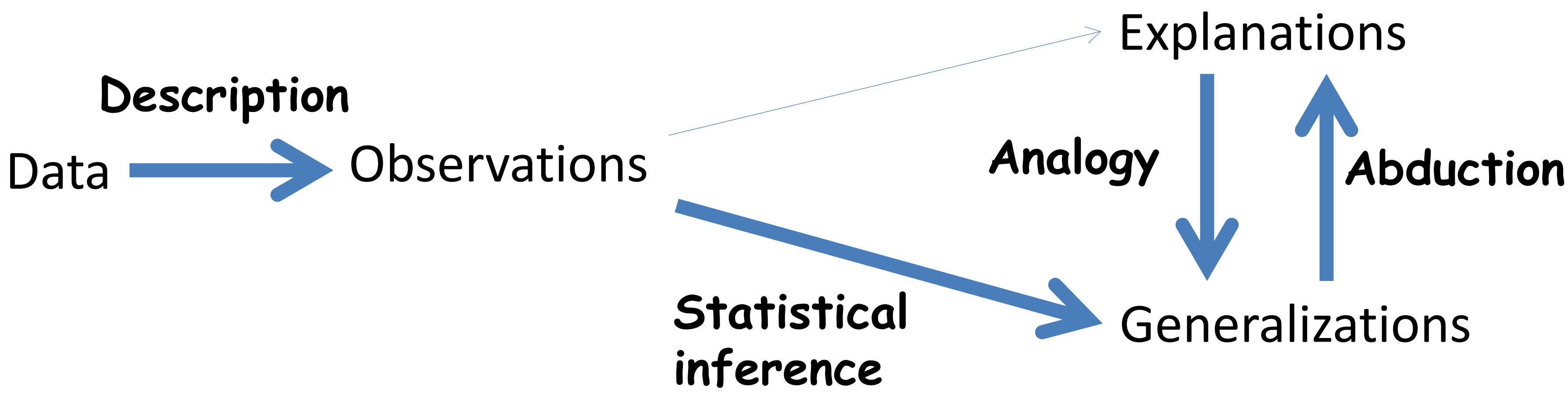

- Descriptive inference: Describing observations
- Abductive inference: Providing an explanation
- Analogic inference: Generalize to similar cases



- Architectural explanation must be the basis of the analogic generalization;
- Otherwise, we engage in wishful/magical thinking
  - *You have observed that some small companies did not put a customer representative on-site of an agile project;*
  - *you explain this as a result of tight resources (level 3);*
  - *you generalize by analogy that this will happen in (almost) all small companies (level 3).*

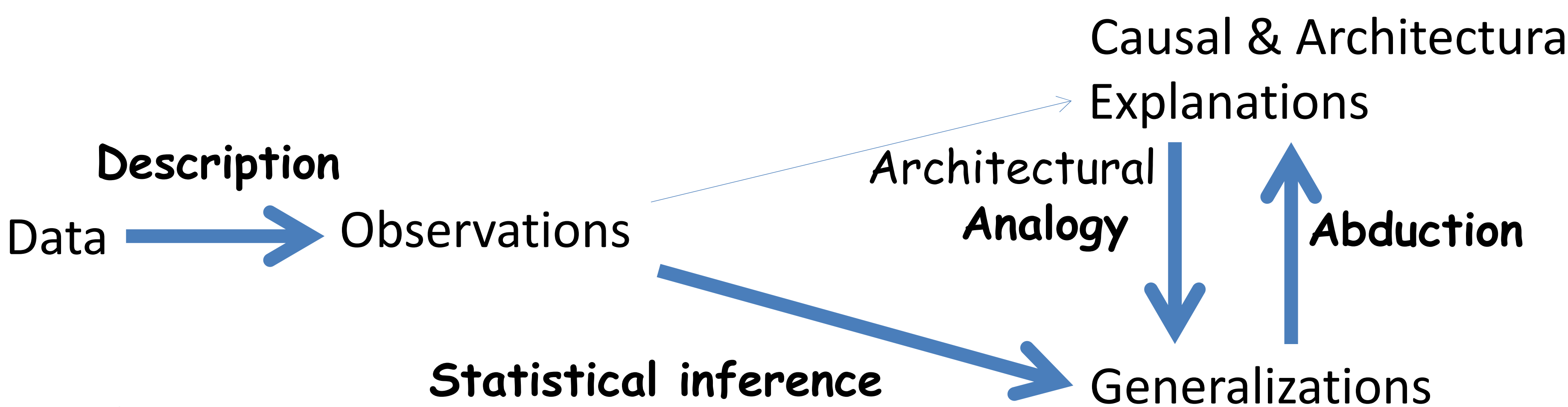


# Sample-based inference




- Descriptive inference: Describe sample statistics
- Statistical inference: Generalize to population parameters
- Abductive inference: Provide an explanation
- Analogic inference: Expand the scope of a theory based on similarity

- Causal explanations can be supported by sample-based designs (treatment group/control group)
- Generalization **from** a population, to similar populations must be based on architectural explanation
  - *In an experiment with a sample of students you observe a difference between treatment group and control group;*
  - *By randomness you generalize to population of students*
  - *Your explanation: this difference is caused by the treatment (level 3);*
  - *In turn explained by cognitive processes of students (level 3);*
  - *generalized by analogy to novice software engineers (level 3).*



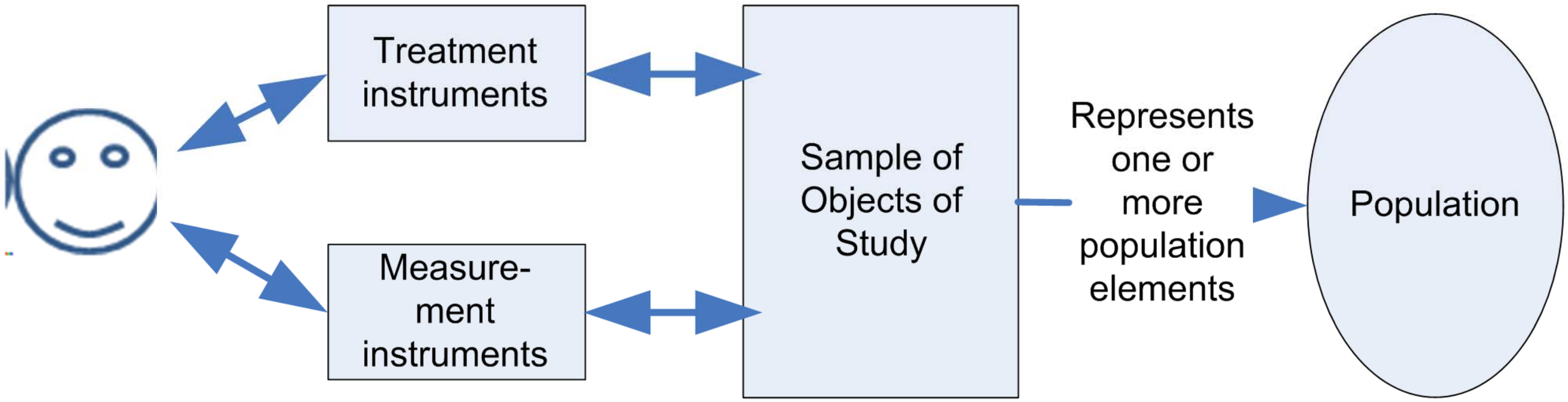

- *Vriezekolk: Inferring theories from data*
- *Méndez: inferring theories from data*
- *Prechelt: Applying/inferring theories to/from data*

# Agenda

| Time | Topic |
|---|---|
| 09:00 – 10:30 | Opening and Introduction |
| 10:30 – 11:00 | Coffee break |
| 11:00 – 12:30 | Inferring Theories from Data |
| 12:30 – 13:30 | Lunch |
| 13:30 – 15:00 | Designing Research based on Theories |
| 15:00 – 15:30 | Coffee break |
| 15:30 – 16:30 | Hands-on Working Session and Q&A |
| 16:30 – 17:00 | Wrap up (all) |

# Research Design

# The research setup




- In experiments we are interested in the effect of the treatment on the OoS
  - Requires capability to apply treatment and control
- In observational studies we are interested in the structure and dynamics of the OoS itself
  - Only weak support for causality

- Case-based designs
  - provide architectural explanations
  - generalize by architectural analogy
  - Nondeterminism across cases is not quantified
- Sample-based designs
  - Collect sample statistics
  - Infer properties of distribution over population
  - May be purely descriptive!
  - Possibly a causal explanation
  - To generalize further, need architectural explanation too
  - Nondeterminsim within the population is quantified, but not across analogous populations

# Field versus lab

- If a phenomenon cannot be (re)produced in the lab, it can only be investigated in the field
- Which of the following designs can be done in a lab?

| | Case-based inference | Sample-based inference |
|---|---|---|
| No treatment (observational study) | Observational case study | Survey |
| Treatment (experimental study) | Single-case mechanism experiment, Technical action research | Statistical difference-making experiment |

E.g. simulation, test of individual OoS

E.g. test with client, pilot project

Treatment group / control group designs


- *Vriezekolk The research setup*
- *Méndez: The research setup*
- *Prechelt: The research setup*

# Agenda

| Time | Topic |
| --- | --- |
| 09:00 – 10:30 | Opening and Introduction |
| 10:30 – 11:00 | Coffee break |
| 11:00 – 12:30 | Inferring Theories from Data |
| 12:30 – 13:30 | Lunch |
| 13:30 – 15:00 | Designing Research based on Theories |
| 15:00 – 15:30 | Coffee break |
| 15:30 – 16:30 | Hands-on Working Session and Q&A |
| 16:30 – 17:00 | Wrap up (all) |

# Hands-on Working Session

# Hands-on Working Session

1. What is your research question?
2. Describe a research setup to answer it
3. What inferences do you plan to base on this setup?

Groups of 3

- 15:30 Each person first drafts a flipchart with his/her answers for own research
- 15:45 Each group member comments on the two flipcharts of others in his/her group, in particular on:
  - Are the answers clear?
  - Are the answers defensible?
- 16:30 Each person finalizes (for now) his/her flipchart
- 16:31 Paste to the wall. See what you can learn from other designs.
- 16:45 Plenary wrap-up

# Q&A

**You probably can't ask Google anyway, so ask us!**