Introduction to Design Science Methodology

Prof. Dr. Roel Wieringa
University of Twente
The Netherlands

Outline

- Design science
- Theories
- Methods
  - Empirical research setup
  - Patterns of reasoning

- R.J. Wieringa. *Design Science Methodology for Information Systems and Software Engineering*. Springer, 2014.
- More information at http://wwwhome.ewi.utwente.nl/~roelw/

- Design science is the design and investigation of artifacts in context

- Design science is the design and investigation of **artifacts in context**

Subjects of design science

Problem context:

SW components & systems,
HW components & systems,
Organizations,
Business processes,
Services,
Methods, Techniques,
Conceptual structures,
People,
Values, Desires, Fears,
Goals, Norms, Budgets,
...

Interaction ↔

Artifact:

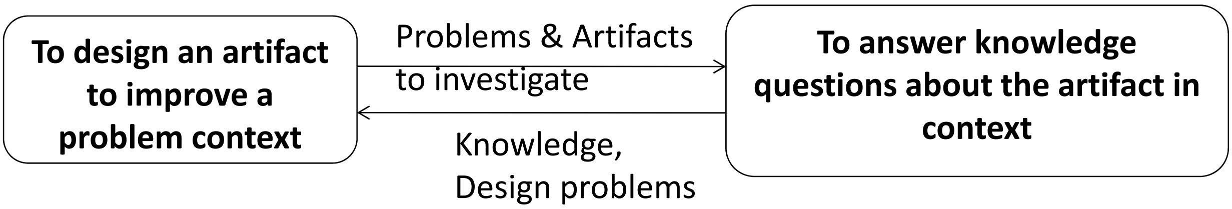
SW component/system,
HW component/system,
Organization,
Business process,
Service,
Method, Technique,
Conceptual structure,
...

Something to be influenced

Something to be designed

- Design science is the **design and investigation** of artifacts in context

Research problems in design science



- *"Design a DoA estimation system for satellite TV reception in a car."*
- *"Design a multi-agent aircraft taxi-route planning system for use on airports"*
- *"Design an assurance method for data location compliance for CSPs"*

- *"Is the DoA estimation accurate enough?"*
- *"Is this agent routing algorithm deadlock-free?"*
- *"Is the method usable and useful for cloud service providers?*

The design researcher iterates over these two activities

Discussion

- What is the research problem that you are working on?

Design problems

Template for design problems

- Improve <problem context>
- by <treating it with a (re)designed artifact>
- such that <artifact requirements>
- in order to <stakeholder goals>

- *Improve my body / mind health*
- *by taking a medicine*
- *such that relieves my headache*
- *in order for me to get back to work*

Template for design problems

- Improve <problem context>
- by <treating it with a (re)designed artifact>
- such that <artifact requirements>
- in order to <stakeholder goals>

- *Improve my body / mind health*
- *by taking a medicine*
- *such that relieves my headache*
- *in order for me to get back to work*

Problem context and stakeholder goals

Template for design problems

- Improve <problem context>
- by <treating it with a (re)designed artifact>
- such that <artifact requirements>
- in order to <stakeholder goals>

- *Improve my body / mind health*
- *by taking a medicine*
- *such that relieves my headache*
- *in order for me to get back to work*

Artifact and its desired interactions

- Design problems are usually not considered to be research problems
- They are stated in the form of questions
  - *How to plan aircraft taxi routes dynamically?*
  - *Is it possible to plan aircraft routes dynamically?*
  - *Etc.*
- And they are called “technical research questions”.
- This way, stakeholders, goals, and requirements stay out of the picture!

Discussion

- What is your top-level design problem?

The engineering cycle

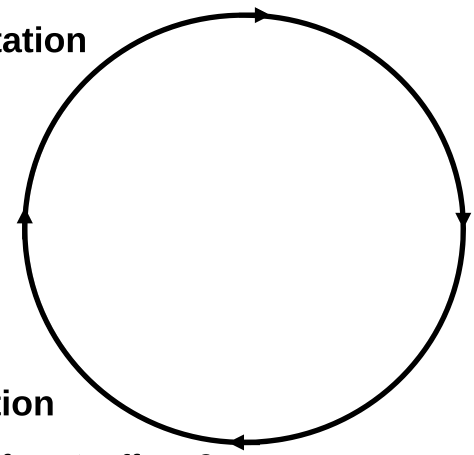
(Checklist for solving design problems)

Engineering cycle

Legend:
Knowledge questions?
Actions!

Design implementation

Implementation evaluation = Problem investigation



- Stakeholders? Goals?
- Conceptual problem framework?
- Phenomena? Causes? Effects?
- Effects contribute to Goals?

Design validation

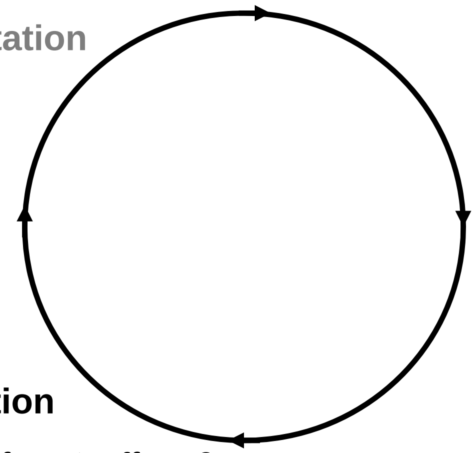
- Context & Artifact → Effects?
- Effects satisfy Requirements?
- Trade-offs for different artifacts?
- Sensitivity for different Contexts?

Treatment design

- Specify requirements!
- Requirements contribute to goals?
- Available treatments?
- Design new ones!

Implementation (transfer to problem context) is not part of research

Design implementation



Implementation evaluation = Problem investigation

- Stakeholders? Goals?
- Conceptual problem framework?
- Phenomena? Causes? Effects?
- Effects contribute to Goals?

Treatment design

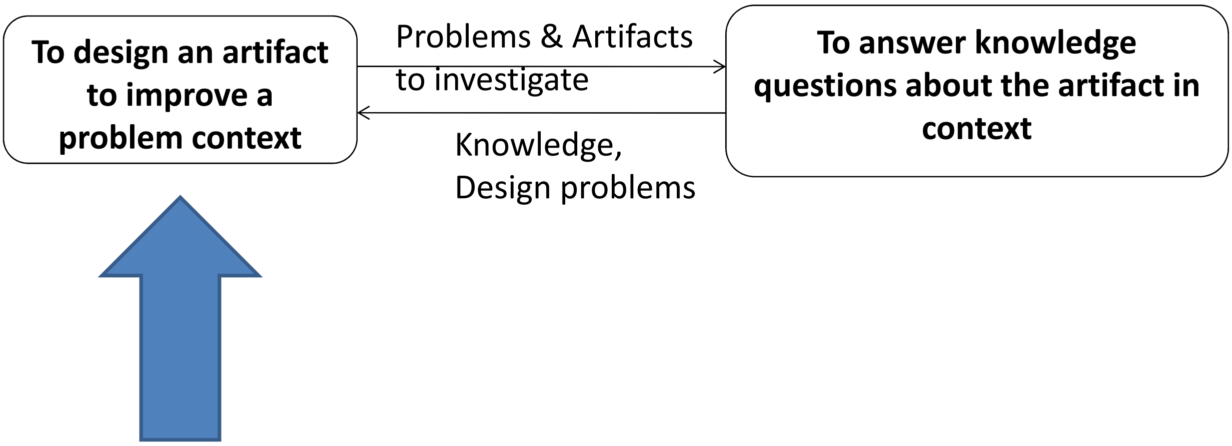
- Specify requirements!
- Requirements contribute to goals?
- Available treatments?
- Design new ones!

Design validation

- Context & Artifact → Effects?
- Effects satisfy Requirements?
- Trade-offs for different artifacts?
- Sensitivity for different Contexts?

- Research projects may focus on
  - Implementation evaluation
  - Problem investigation
  - Treatment design and validation

Research problems in design science



10 July 2014 eBISS Summer School 20

Knowledge questions

- Descriptive questions:
  - What happened?
  - When?
  - Where?
  - What components were involved?
  - Who was involved?
  - etc.
- Explanatory questions:
  - Why?
    - What has caused the phenomena?
    - Which mechanisms produced the phenomena?
    - For what reasons did people do this?

Journalistic questions (descriptive questions)

Research questions (explanatory questions)

Effect questions

- Central effect question
  - **Effect question:** Context X Artifact → Effects?
- Generalizations
  - **Trade-off question:** Context X ***Alternative artifact*** → Effects?
  - **Sensitivity question:** ***Other context*** X artifact → Effects?
- Descriptive or explanatory questions

Contribution questions

- Central contribution question:
  - **Contribution question:** Do Effects contribute to Stakeholder goals?
- Preliminary questions:
  - **Stakeholder question:** Who are the stakeholders?
  - **Goal question:** What are their goals?

- In academic research projects, the answers to these questions may be speculative
  - From utility-driven to curiosity-driven projects

Example knowledge questions

- ***Effect:***
  - *What is the execution time of the DoA algorithm?*
  - *What is its accuracy?*
- ***Trade-off:***
  - *Comparison between algorithms on these two variables*
  - *Comparison between versions of one algorithm*
- ***Sensitivity:***
  - *Assumptions about car speed?*
  - *Assumptions about processor?*
- ***Stakeholders:***
  - *Who are affected by the DoA algorithm?*
- ***Goals:***
  - *What are their goals?*
- ***Contribution evaluation (after DOA algorithm is in use)***
  - *How well does the DoA algorithm contribute to these goals?*

Discussion

- Which knowledge questions do you have?
  - Effect questions
  - Trade-off
  - Sensitivity
  - Satisfaction of requirements
  - Contribution to stakeholder goals

Outline

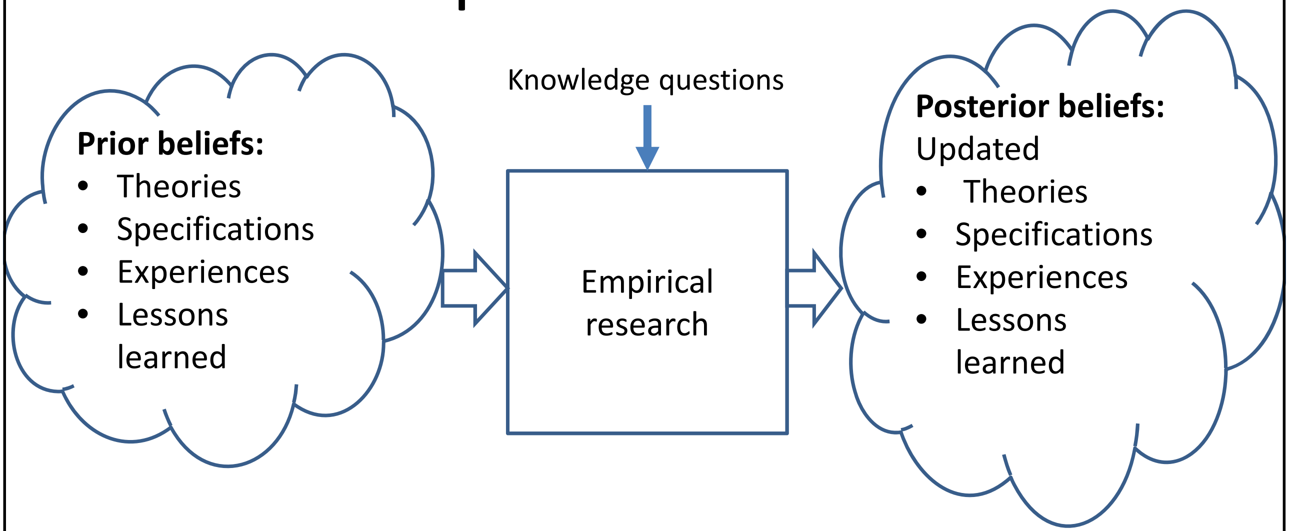
- Design science
  - – Design problems
  - – Engineering cycle
  - – Knowledge questions
- Theories
- Methods
  - – Empirical research setup
  - – Patterns of reasoning

Outline

- Design science
  - Design problems
  - Engineering cycle
  - Knowledge questions
- **Theories**
- Methods
  - Empirical research setup
  - Patterns of reasoning

- To answer a knowledge question, you may have to
  - Read the scientific literature
  - Read the professional literature
  - Ask experts
  - Do original research ⇐ **If scientific research, this is very expensive**

Empirical research



- The goal of empirical research is to develop, test or refine **theories**

- A **theory** is a belief that there is a pattern in phenomena
  - Speculations
  - Opinions
  - Ideologies
  - …
- A **scientific** theory is a theory that
  - Has survived tests against experience
  - Has survived criticism by critical peers
- All theories about the real world are fallible

The structure of scientific theories

1. **Conceptual framework**
   - *E.g. The concepts of beamforming, of multi-agent planning, of data location compliance*
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*
3. **Scope** of the generalizations. Population, or similarity relation
   - *Assumptions about the phenomena to which relation is applicable: plabne waves, narrow bandwidth, etc.*

The structure of **design** theories

1. Conceptual framework to **specify artifact** and **describe context**
2. Generalizations of the form **Artifact X Context → Effects**
3. The scope:
   - **constraints** on artifact design,
   - **assumptions** about the context

Variables

- Conceptual frameworks may define variables.
- Variables have data types and scales
- Generalizations are stated in terms of variables
- Examples (variables in **bold**):
  - *DOA performance graphs relating **noise**, **angle of incidence**, and **accuracy of estimation***
  - *DOA analytical generalization: change in **angle of incidence** causes change in **phase difference***
  - *Software engineering empirical generalization: Introduction of **agile development** causes **customer satisfaction** to increase*
  - *Software engineering laboratory generalization: **Programmer productivity** correlates well with **conscientiousness***

Architectures

- Conceptual frameworks may define an architecture for phenomena in terms of components and relationships
- Components have capabilities
- Generalizations can be stated in terms of capabilities of components and of interactions of components (mechanisms)
- Examples (components in **bold**):
  - *DOA mechanistic theory: e.g. input-output relation is explained by* ***components*** *and structure of the algorithm*
  - *A mechanism in observed in agile development: In agile development for SME, the* ***SME*** *does not put* ***customer*** *on-site. SME resources are limited and focus is on business.*
  - *A mechanism observed in requirements engineering: Introduction of* ***change control board*** *reduces requirements creep.*

Functions of scientific theories

The functions of scientific theories

- To **analyze** a conceptual structure
- To **describe** phenomena (descriptive statistics, interpretation)
- To **explain** phenomena
- To **predict** phenomena (important for design)
- To **design** an artifact by which to treat a problem

The functions of scientific theories

- To **analyze** a conceptual structure
- To **describe** phenomena (descriptive statistics, interpretation)
- To **explain** phenomena
- To **predict** phenomena **(important for design)**
- To **design** an artifact by which to treat a problem

The functions of scientific theories

- To **analyze** a conceptual structure
- To **describe** phenomena (descriptive statistics, interpretation)
- To **explain** phenomena (the classical function of theories)
- To **predict** phenomena (important for design)
- To **design** an artifact by which to treat a problem

Causal explanations (cause-effect relation between variables)

- If Y has been **caused** by X, then Y changed because X changed earlier in a particular way
- Examples
  - *Light is on because switch was turned*
  - *Cost increased because the organization had to perform additional tasks*
- Causation may be nondeterministic
  - Forward nondeterminism: X sometimes causes Y
  - Backward: Y is sometimes caused by X
- In the field, the causal influence of X on Y may be swamped by many other causal influences.
  - Lab research versus field research

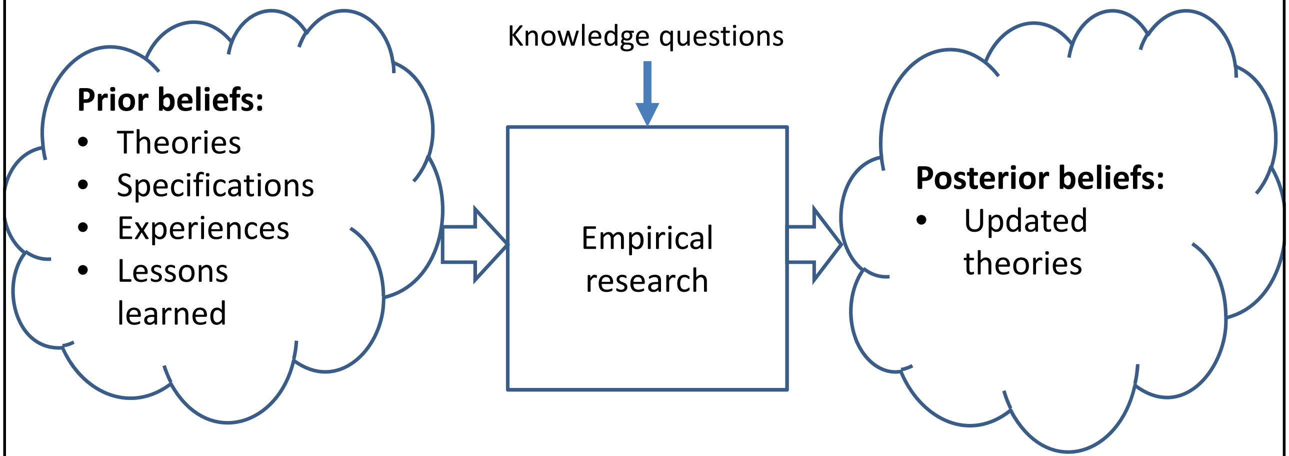
Architectural explanations (interactions among components)

- If system phenomenon E was produced by the interaction of system components C1, …, Cn, then C1, …, Cn is called a **mechanistic explanation** of E.
- Examples
  - *Light is on because it is connected by to electricity supply when switch was turned on*
  - *Cost increased because new people had to be hired to perform additional tasks*
- May be nondeterministic
- May be interfered with by other mechanisms in the field

Checklist for empirical research

(the empirical cycle)

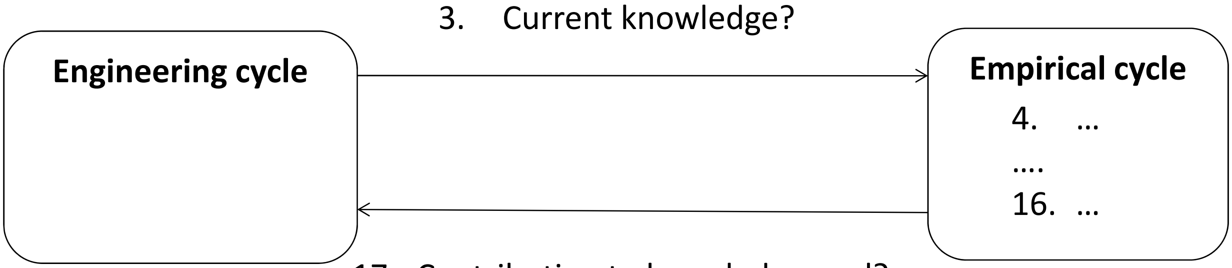
Empirical research



- The goal of empirical research is to develop, test or refine theories

Checklist questions about research context

1. Improvement goal?
2. Knowledge goal?
3. Current knowledge?



17. Contribution to knowledge goal?
18. Contribution to improvement goal?

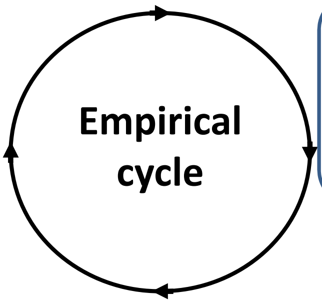
Data analysis

12. Data?
13. Observations?
14. Explanations?
15. Generalizations?
16. Answers?

New research problem

Research problem analysis

4. Conceptual framework?
5. Research questions?
6. Population?



Research & inference design

7. Object of study?
8. Treatment specification?
9. Measurement specification?
10. Inference?

Design validation

7. Object of study validation?
8. Treatment specification validation?
9. Measurement specification validation?
10. Inference validagtion?

Research execution

11. What happened?

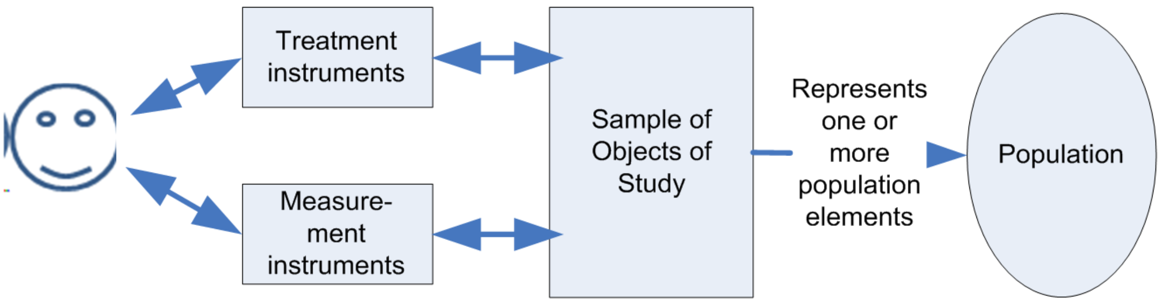
Outline

- Design science
- Theories
  - Structure: Conceptual framework, generalizations
  - Functions: explanation etc.
  - Empirical cycle
- Methods
  - Empirical research setup
  - Patterns of reasoning

Outline

- Design science
- Theories
  - Structure: Conceptual framework, generalizations
  - Functions: explanation etc.
  - Empirical cycle
- Methods
  - **Empirical research setup**
  - Patterns of reasoning

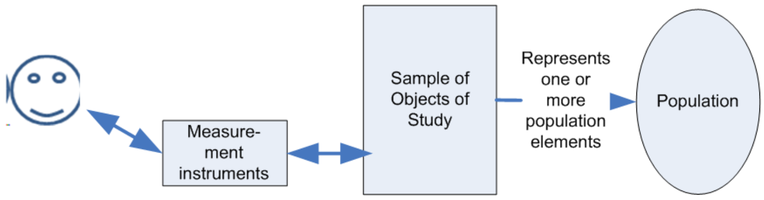
The empirical research setup



- The researcher wants to answer a question about a population.
- He or she selects a sample of objects of study (OoS) that represent population elements.
- In experimental research: S/he treats some/all OoS's in the sample.
- S/he measures phenomena in the OoS's.

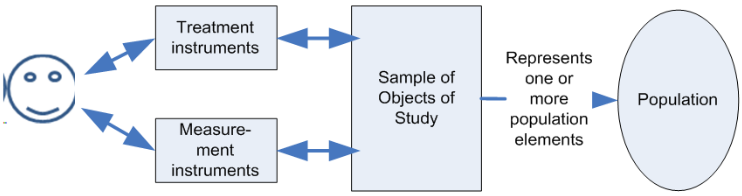
- Observational versus experimental setup
- Case-based versus sample-based research

Observational setup



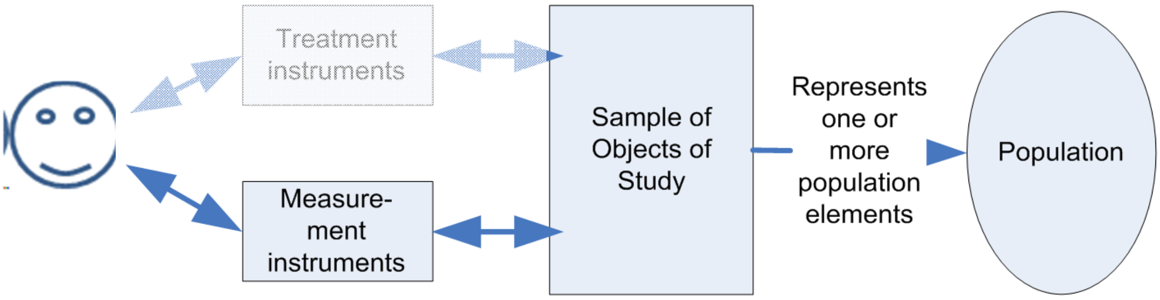
- The researcher wants to answer a question about a population.
  - *E.g. How is the UML used? About all SE projects that use UML.*
  - *What are the causes of project failure? About all IS development projects*
- He or she selects a sample of objects of study (OoS) that represent population elements.
  - *All projects in a company*
  - *Some projects in some companies*
  - *One project in some company*
- S/he measures phenomena in the OoS's.
  - *Modeling effort, model correctness,*
  - *Using as instruments primary documents, interviews, questionnaires, email logs, UML models, …*

Experimental setup



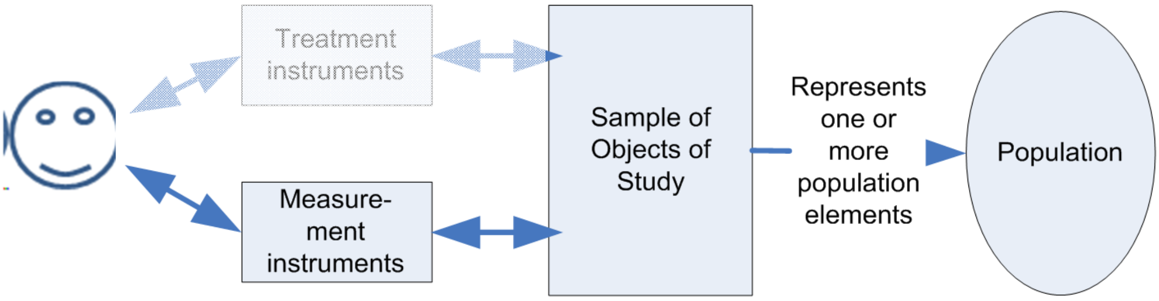
- The researcher wants to answer a question about a population.
  - *E.g. what is the effect of using UML? About all SE projects that use UML.*
- He or she selects a sample of objects of study (OoS) that represent population elements.
  - *Some projects in some companies*
  - *One project in some company*
- S/he treats some/all OoS's in the sample.
  - Ask some projects to use the UML
- S/he measures phenomena in the OoS's
  - *Modeling effort, model correctness, using similar instruments as before*

Case-based research



- Observational or experimental
- Study one OoS at a time:
- The sample is studied in series, with an analysis in between two case studies.
  - ~~*What is the effect of using the UML?*~~
  - *How is the UML used?*
  - *Which architecture does the case have? (e.g. actors, documents, artifacts)*
  - *Which mechanism take place? (interactions, communications, coordination)*

Sample-based reseach



- Observational or experimental
- Study samples of OoS's as a whole
- Sample statistics are used to derive estimations of statistical population parameters
  - *What is the effect of using UML?*
  - *What is the average modelling effort, compared to the modelling effort of other projects of similar size?*

	Case-based research	Sample-based research
Observational setup (No treatment)	**Observational case study:** Study the structure and mechanisms of a single case	**Survey:** Study a large population sample statistically
Experimental setup (treatment)	**Single-case mechanism experiment:** Testing a prototype, simulating a system, …. **Technical action research:** Experimental use of a novel artifact	**Statistical difference-making experiment:** Comparison of difference in statistical outcomes of treatments on two samples

Outline

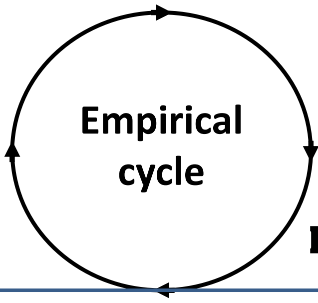
- Design science
- Theories
- Methods
  - Empirical research setup
    - Observational or experimental
    - Case-based or sample-based
  - **Patterns of reasoning**

Data analysis

12. Data?
13. Observations?
14. Explanations?
15. Generalizations?
16. Answers?

Research execution

11. What happened?



Research problem analysis

4. Conceptual framework?
5. Research questions?
6. Population?

Research setup and justification

Design validation

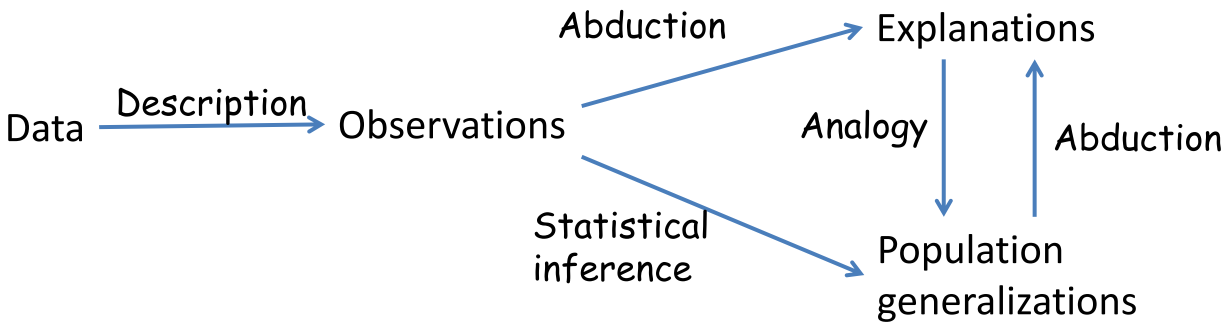
7. Object of study validation?
8. Treatment specification validation?
9. Measurement specification validation?
10. Inference validation?

Research & inference design

7. Object of study?
8. Treatment specification?
9. Measurement specification?
10. Inference?

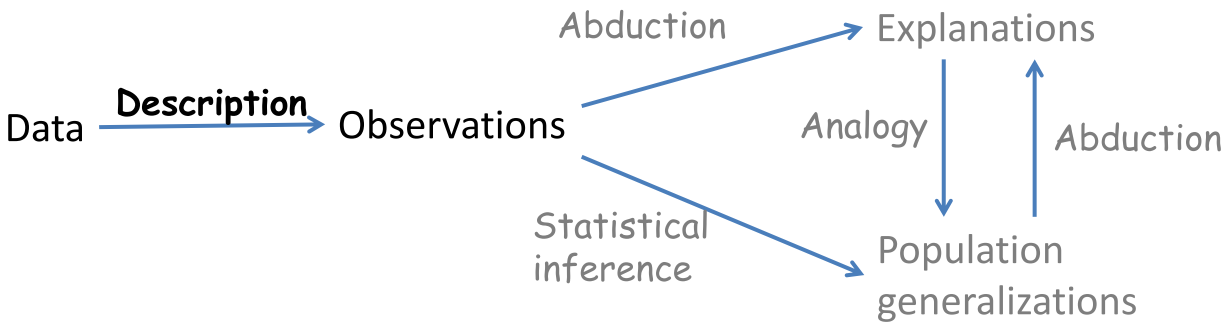
- Each of the choices in the design of a research setup has consequences for the kinds of inferences from data that we can do
  - Validity of research setup wrt planned inferences
  - Validity of inferences wrt research setup

Inferences from data



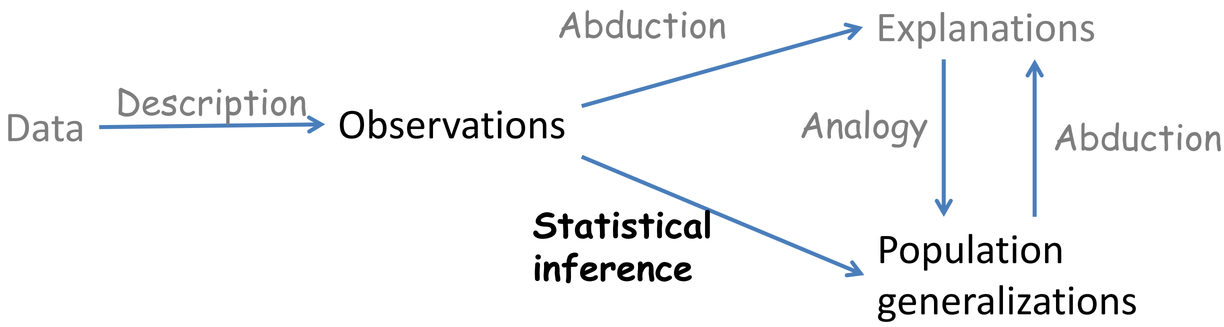
- **Conclusion validity**: How well is a statistical inference from a sample to a population supported?
- **Internal validity**: How well is an explanation supported by an abductive argument?
- **External validity**: How well is a generalization beyond a population supported by an analogy?

Descriptive inference from raw data



- Removal of outliers, computation of statistics
- Visualization of data
- Interpretation of words and images
- Descriptive validity:
  – Descriptive inference should add no information to the data (= non-ampliative = non-defeasible)

Statistical inference



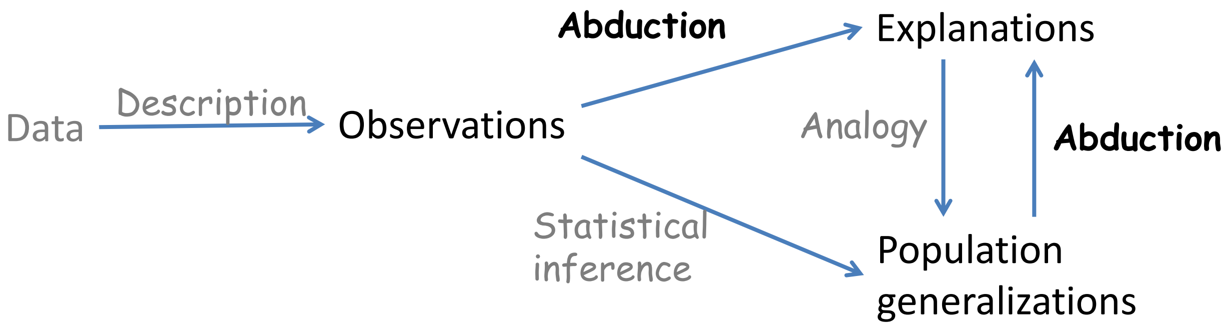
- Estimation of population parameters
- Computational explanation of observations by some statistical model
- Validty wrt research setup

Examples

- *Statistical inference is sample-based*
- *From an observational setup:*
  - *Classify the vulnerabilities found in a sample of 20 open source web applications*
  - *Find that in this sample, on the average 70% of the vulnerabilities in an OS WA are implementation vulnerabilities*
  - *Infer a confidence interval for the average proportion of implementation vulnerabilities in the population of web applications.*
  - *Validity: Assume that sample is random draw from a population, which has a constant probability of implementation vulnerabilities*

- From an experimental setup:
  - *Teach two programming techniques to two groups of students*
  - *Let them write programs using these techniques, and ask other students to perform maintenance tasks on these programs*
  - *Measure effort (= time to perform maintenance task)*
  - *Compute difference in average effort in the two groups.*
- Two kinds of statistical inference:
  - *(a) Estimate confidence interval for the average effort in the population; if 0 is not in this confidence interval, infer that there is a statistically discernable difference in average maintenance effort in the two populations …*
  - *(b) Compute probability of observing at least the measured difference if the population difference would be 0; if this probability is small, conclude that there is a difference in the population*
- *Validity*
  - *Random sampling & allocation, sufficient sample size, stable probability distribution, assumptions about the distribution (e.g. normality).*

Abductive inference



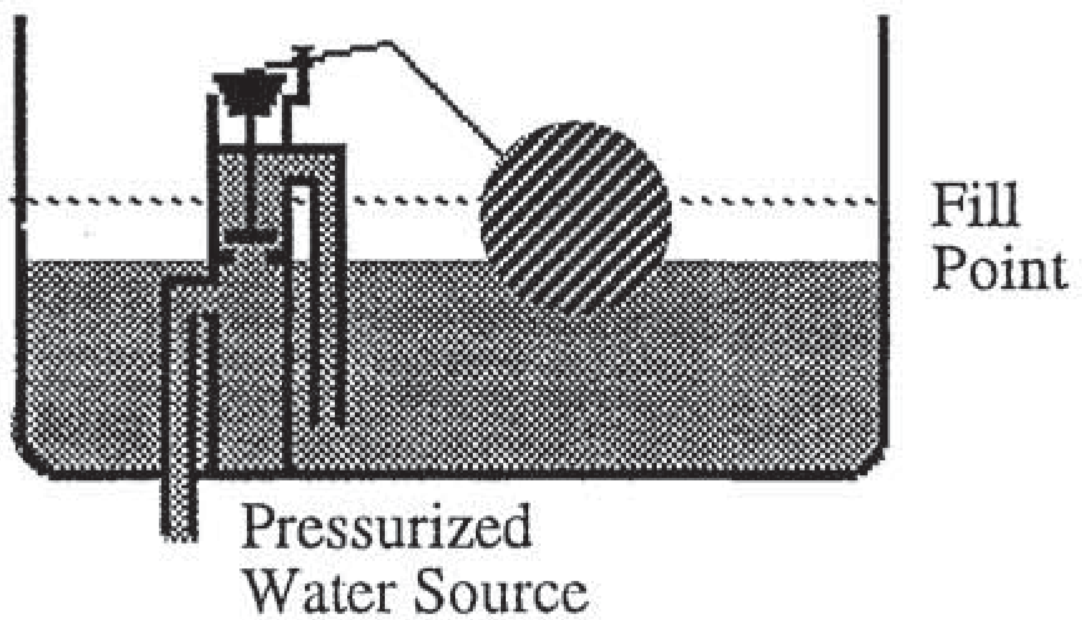
- Explanations of observations or of population-level generalizations
  - Causal explanations (one variable makes a difference to another)
  - Architectural explanations (components, capabilities, mechanisms)
  - Rational explanations (desires, goals, motivations)
- Validity wrt research setup

Causal explanations

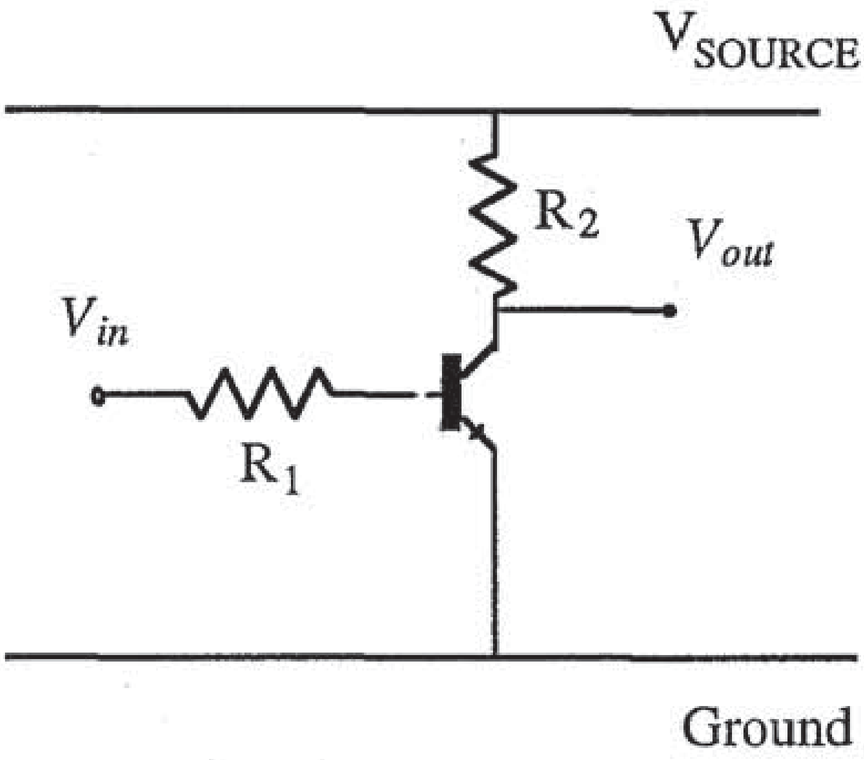
- Single-case causal experiment
  - Apply stimulus to object of study, withhold the stimulus, compare the effects.
  - Validity: Effect is transient, and all other conditions remain constant.
- Comparative case causal experiment
  - Apply stimulus to one OoSD, withhold from the other, compare effects.
  - Validity: OoS's are imilar, all other conditions constant.
- Randomized controlled trial
  - E.g. maintenance example given earlier.
  - In the long run, the only plausible cause of outcome difference is difference in treatments
- Quasi-experiment
  - Same, but with non-random sampling/allocation. Pre & posttest of relevant variables
  - Rank all possible causes on plausibility

Architectural explanations

- Explain a phenomenon by interaction among components that produced the phenomenon
- Components have capabilities/limitations
- The architecture constrains possible interactions
- Mechanism = interaction triggered by stimulus

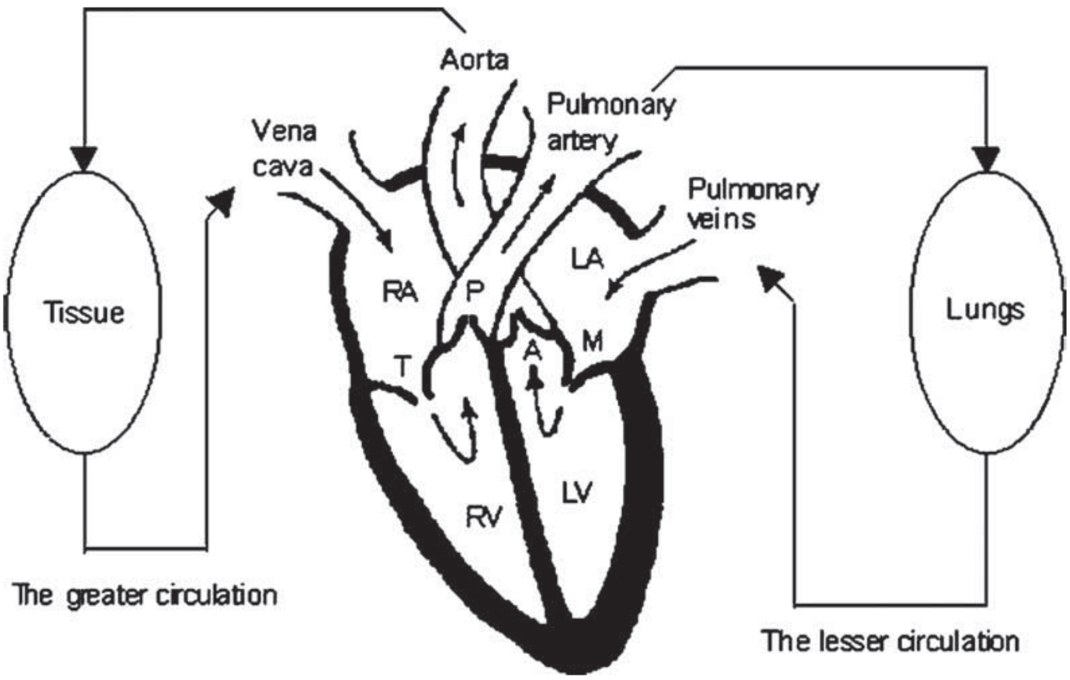


- Glennan - ``Mechanisms and the nature of causation''. 1996

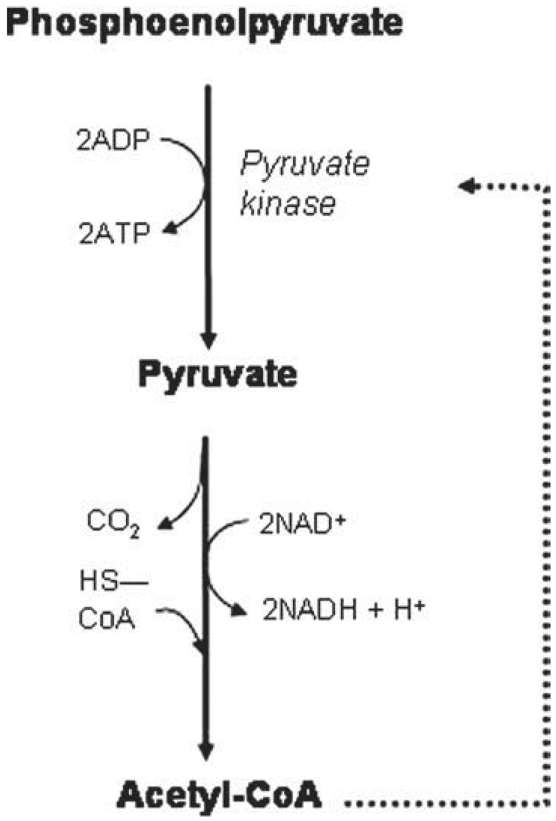


A voltage switch

- Glennan - ``Mechanisms and the nature of causation''. 1996



- Bechtel & Abrahamsen – ``Explanation; a mechanistic alternative.'' 2005



Bechtel & Abrahamsen – ``Explanation; a mechanistic alternative.'' 2005

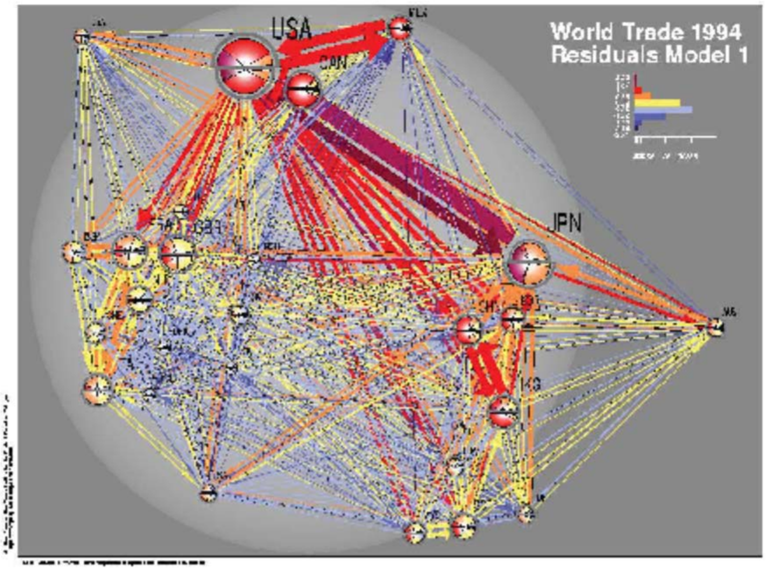


Figure 1.8: In a network representing international trade, one can look for countries that occupy powerful positions and derive economic benefits from these positions [262]. (Image from http://www.cmu.edu/joss/content/articles/volume4/KrempelPlumper.html)

resource or use it more efficiently. In fact, the interactions among people's behavior can lead

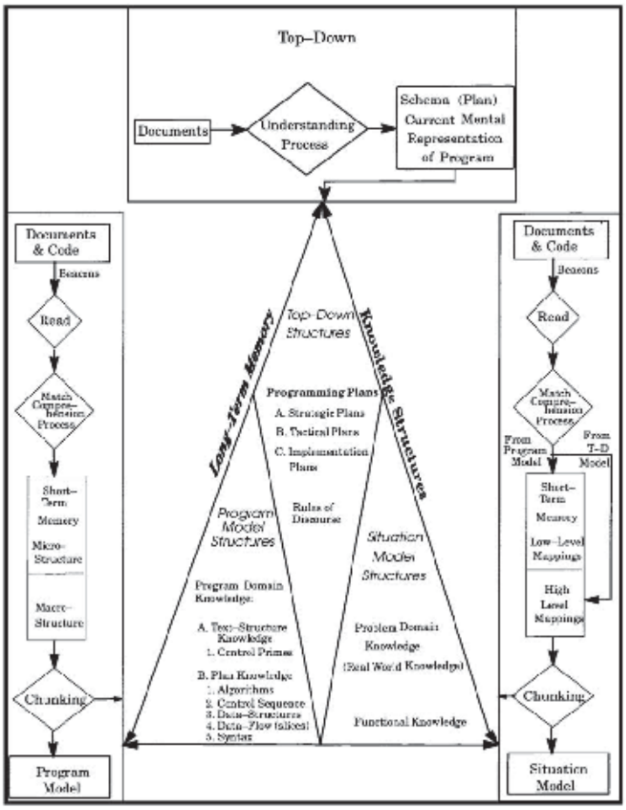


Figure 1. Diagrammatic representation of the Integrated Comprehension model from von Mayrhauser and Vans (1995a)

70

J. Softw. Maint., 9, 299–327 (1997)

Two kinds of effect questions

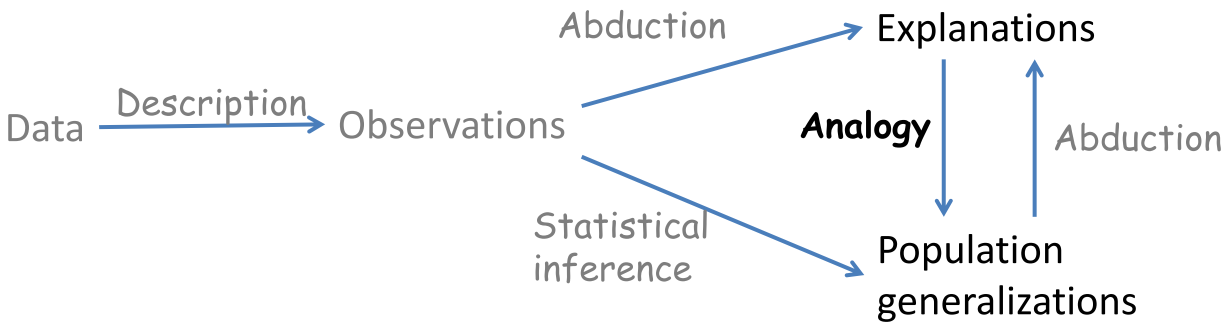
- *Test effect of programming technique on effort.*
- Effect of treatment
- Statistical observation of difference
- **Causal** explanation of outcome difference by difference in treatment
- This calls for a further architectural explanation

- *Test effect of personality on productivity.*
- Difference among capabilities
- Statistical observation of difference
- Explanation of difference in outcome by differences in **capability.**
- This too needs a further architectural explanation

Rational explanation

- Explain behavior of actors in terms of their goals
  - *Explain project failure by power struggles,*
  - *Deviation from business processes by personal goals of people, etc.*
- Validity: we know their goals, and they are motrivated by their goals
- NB rational explanations extend architectural explanations.

Analogic inference

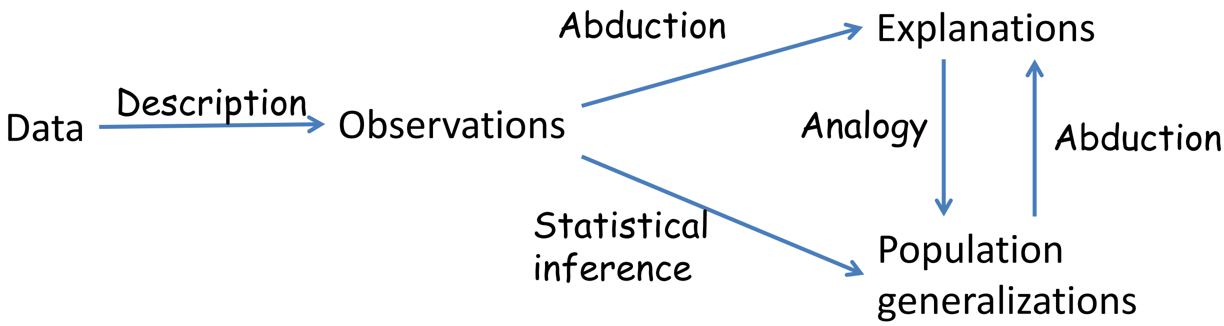


- Analogy
  - (Similar cases / similar populations) will exhibit similar observations produced by the same (causes / mechanisms / reasons)
  - Similarity in variables or similarity in architecture?
  - The **explanation** is generalized by analogy.
- Validity wrt research setup

Examples

- Case-based analogies:
  - *This agile project done for an SME is similar to that one, so probably the SME will not put a client on-site of the project here too.*
  - *This machine is of the same typer as that one, so it will probably contain the same mechanisms*
- Sample-based analogy:
  - *The elements of population are architecturally similar to that one, so the distribution of X is probably similar too.*
- Validity:
  - *Architectural similarity; no other mechanisms that interfere*

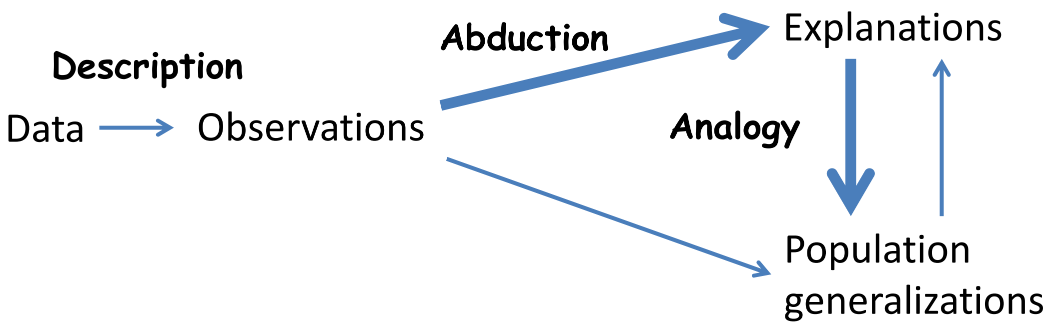
- Variable-based analogy is weak basis for analogic generalization
- Superficial analogy (similar values of variables) is the logic of sympathetic magic
- Inference is correct in rare cases
  - *Benjamin Franklin*
- We need similarity in architecture, so that we can assume similarity in mechanisms



Data
Description
Observations
Abduction
Explanations
Analogy
Abduction
Statistical inference
Population generalizations

- Case-based research and sample-based research have their own typical patterns of reasoning

Case-based inference

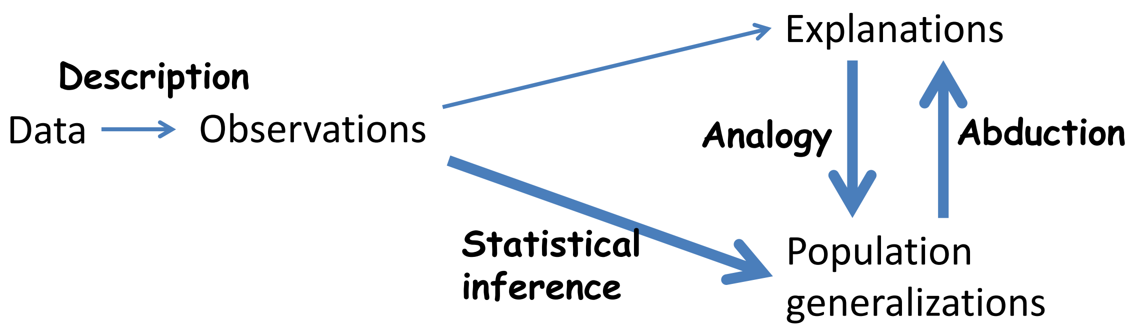


- In case-based inference,
  - we postulate a mechanism to explain behavior observed in a case and
  - reason by analogy that in architecturally similar cases, these mechanisms will produce similar effects

Case-based inference

- Examples
  - *Observational case study: Studying agile projects in the real world*
  - *Observational case study: studying coordination phenomena in a global software engineering project*
  - *Simulation: Testing a software prototype in a simulated context*
  - *Technical action research: Applying an experimental risk assessment technique for a client*

Sample-based inference



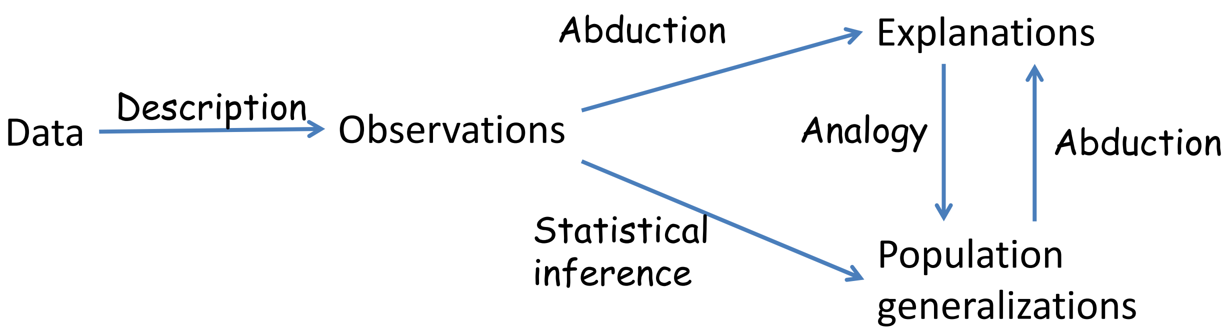
- In sample-based inference,
  - we statistically infer a property of a population from statistics of a sample,
  - Postulate one or more possible mechanism to explain this property,
  - Speculate about possible generalizations to other populations.

- Analogy plays a role twice in sample-based inference
  - Once in the definition of the population
  - Once in generalizing to similar populations (external validity)

Sample-based inference

- Examples
  - *Survey of sample of agile projects*
  - *Survey of coordination phenomena in global software engineering projects*
  - *Statistical difference-making experiment: Comparing two software engineering techniqaues in two samples of student projects*

Prediction



- We may stop after description, generalize by analogy, and then use this for prediction
  - Must assume a stable architecture in cases generalized about, even if we do not know it.
- We may stop after statistical inference, and use for prediction of statistics of future samples from same (or different!) population
  - Must assume stable architecture in population, even if we do not know it.

Patterns of reasoning

	Case-based research: case-based inference	**Sample-based research: sample-based inference**
Observational setup (No treatment)	Observational case study: **Architectural explanation, analogy.**	Survey **Statistical inference.**
Experimental setup (treatment)	Single-case mechanism experiment **Arch explanation, analogy; Causal reasoning too if similarity high enough** Technical action research **Architectural explanation, analogy**	Statistical difference-making experiment; **Statistical inference; causal inference.**

Take home

- **Design science**
  - Design problems
  - Engineering cycle
  - Knowledge questions
- **Theories**
  - Structure: Conceptual framework, generalizations
  - Functions: explanations etc.
  - Empirical cycle
- **Methods**
  - Empirical research setup
  - Empirical research setup
    - Observational or experimental
    - Case-based or sample-based
  - Patterns of reasoning
    - Description,
    - Statistical inference,
    - Abduction (causal. Architectural, rational)
    - Analogy

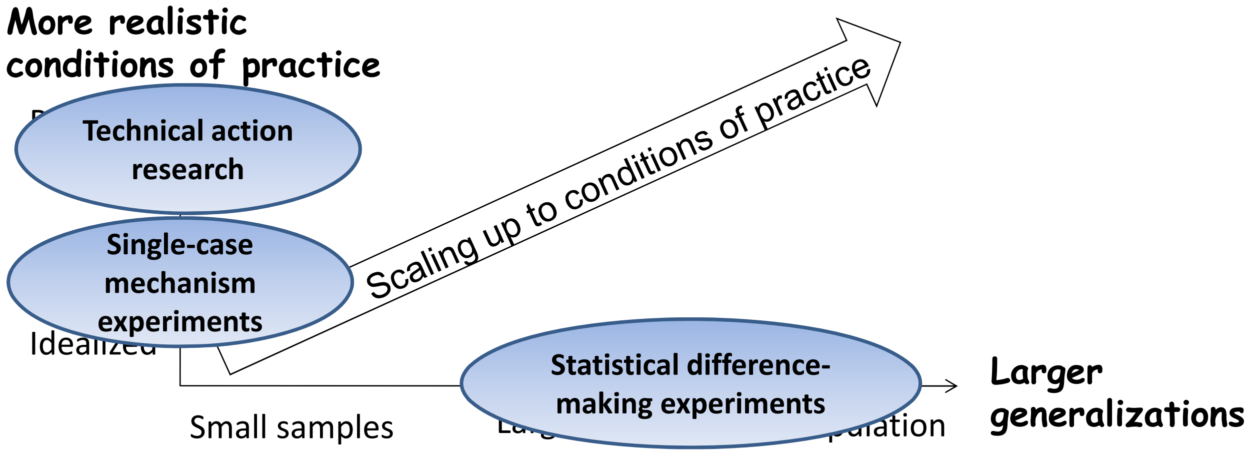
- Wieringa, R.J. *Design Science Methodology for Information Systems and Software Engineering*. Springer, 2014.
- Wieringa, R.J. Daneva, M. ``Six Strategies for Generalizing Software Engineering Theories''. Science of Computer Programming, to be published.
- Wieringa, R.J. (2014) "Empirical research methods for technology validation: Scaling up to practice." *Journal of systems and software*, on-line first.
- Wieringa, R.J. and Condori-Fernández, N. and Daneva, M. and Mutschler, B. and Pastor, O. (2012) *Lessons learned from evaluating a checklist for reporting experimental and observational research.* In: Proceedings, ESEM 2012, pp. 157-160. ACM.
- Wieringa, R.J. and Morali, A. (2012) *Technical Action Research as a Validation Method in Information Systems Design Science.* In: *Design Science Research in Information Systems. Advances in Theory and Practice 7th International Conference*, DESRIST 2012. pp. 220-238. LNCS 7286. Springer.
- Wieringa, R.J. (2010) *Relevance and problem choice in design science.* In: *Global Perspectives on Design Science Research (DESRIST)*. pp. 61-76. LNCS 6105. Springer.
- Wieringa, R.J. (2009) *Design Science as Nested Problem Solving.* In: *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, Philadelphia. pp. 1-12. ACM.

The big picture

Back to the design cycle

Summary of research designs and research goals

	Research goals		
Research designs	**Evaluation research / Problem research**	**Treatment survey**	**Validation research**
Survey	To survey problem owners / implementations	To survey possible treatments	
Observational case study	To study a problem / Implementation		
Single-case mechanism experiment; Expert opinion about an artifact	To diagnose a problem / Test an implementation in context	To test an artifact without context	To validate an artifact in context
Technical Action Research (TAR)			To validate usability and usefulness of an artifact in practice
Statistical difference-making experiment	To compare the effect of interventions on random samples		To compare the effect of treatments on random samples



- **Single-case mechanism experiments:** Investigate underlying mechanisms (interaction between components) in single cases
  - Test a single instance of the artifact in the lab/field
  - Technical action research: Use the artifact to solve real-world problem
- **Statistical difference-making experiments:** Investigate average effects (average difference between treating and not treating) in large samples