Design Science Research Methods

Slides available at https://wwwhome.ewi.utwente.nl/~roelw/DSM-SIKS-nov-2017.pdf.

Prof. Dr. Roel Wieringa
University of Twente,
The Netherlands
www.cs.utwente.nl/~roelw

Research methodology accross the disciplines

- Do these disciplines have the same methodology?
  - Technical science: Build cool stuff; test it; iterate
  - Social science: Observe people, interpret what they do or say; or select a sample, do a lot of statistics; iterate.
    - *For social scientists, engineers are slightly autistic tinkerers*
    - *For technical scientists, social scientists are chatterboxes*
  - Physical science: Build instruments, create phenomena, analyze data, create theories; iterate.
    - *For physicists, other sciences are like stamp collecting*
    - *For physicists, physics is the foundation of engineering*
  - Mathematics: Read, think, write, think; iterate.
    - *Mathematicians think that they provide the foundations of civilization*

Our approach

- All research in all disciplines is **problem-solving**
- Problems solved in rational problem solving cycle
  - Critical investigation of alternatives
  - Confrontation with facts

- Wieringa, R.J. (2014) *Design science methodology for information systems and software engineering.* Springer Verlag

Why are we doing this?

- For senior researchers: how to compete with other disciplines for funds?
- For students: How to structure my thesis?

- How to **justify** your research goals and research results?

Outline

1. Design problems versus knowledge questions
2. The design cycle
3. Design theories
   – Scientific theories
   – Scientific inference: from data to theories
4. The empirical cycle

What is design science?

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

Design problems versus knowledge questions

To design an artifact to improve a problem context

Problems & Artifacts to investigate →

← Knowledge, Design problems

To answer knowledge questions about the artifact in context

Design software to estimate Direction of Arrival of plane waves, to be used in satelite TV receivers in cars	• *Is the DoA estimation accurate enough in this context?* • *Is it fast enough?*
Design a Multi-Agent Route Planning system to be used for aircraft taxi route planning	• *Is this routing algorithm deadlock-free on airports?* • *How much delay does it produce?*
Design a data location regulation auditing method	• *Is the method usable and useful for consultants?*

Is the artifact **useful** in this context?

Is the answer about the artifact in context **true**?

Template for design problems

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

- *Reduce my headache*
- *by taking a medicine*
- *that reduces pain fast and is safe*
- *in order for me to get back to work*

- *BPMN Plus : a modelling language for unstructured business processes.*

Artifact

Context

- The objective of this study is
  - To investigate the way through which unstructured business processes can be modelled and managed without limiting their run-time flexibility.
- Research questions
  - Q1 What are the differences between structured and unstructured business processes?
  - Q2 What are the differences between Business Process Management and Case Management in dealing with unstructured business processes?
  - Q3 What are the capabilities of existing modelling notations to deal with unstructured business processes?
  - Q4 How to model an unstructured business process while providing run-time flexibility?

- **Improve <problem context in which unstructured business process is to be modelled>**
- **by <introducing a modeling language for unstructured business processes>**
- **such that <requirements such as run-time flexibility, and ... learnability etc?>**
- **in order to <stakeholder goals, e.g. provide better process improvement advice to clients>**

Empirical knowledge questions

- **Descriptive** knowledge questions:
  - What happened?
  - How much? How often?
  - When? Where?
  - What components were involved?
  - Who was involved?
  - Etc. etc.
- **Explanatory** knowledge questions:
  - Why?
    1. What has **caused** the phenomena?
    2. Which **mechanisms** produced the phenomena?
    3. For what **reasons** did people do this?

Journalistic questions. Yield facts.

Beyond the facts. Yields theories.

Two kinds of knowledge questions

1. **Design research problems** (a.k.a. *technical research questions*)
   – To improve some kind of artifact in some kind of context.
2. **Empirical knowledge questions**
   – To ask questions about the real world.
3. **Analytical knowledge questions**
   – To ask questions about the logical consequences of definitions

- [*BPMN Plus : a modelling language for unstructured business processes.*]
- The objective of this study is
  - To investigate the way through which the unstructured business processes can be modelled and managed without limiting their run-time flexibility.
- Research questions
  - Q1 What are the differences between structured and unstructured business processes?
  - Q2 What are the differences between Business Process Management and Case Management in dealing with unstructured business processes?
  - Q3 What are the capabilities of existing modelling notations to deal with unstructured business processes?
  - Q4 How to model an unstructured business process while providing run-time flexibility?

- Explanatory questions?
- Analytical questions?

Descriptive knowledge questions; (outcome of interviews)

Design problem

- **Curiosity/fun -driven science** starts with a knowledge question …
- … and continues with instrument design
- **Utility-driven science** starts with an improvement need of stakeholder …
- … and continues with artifact design or with a knowledge question

- Sponsors are always utility-driven
- Researchers are always curiosity and/or fun-driven

Reality check

- Each thesis has a top-level research problem
- For which theses is this a design problem and for which is it a knowledge qestion?
  - SIKS dissertations http://www.siks.nl/dissertations.php
  - Master theses in business informatics http://essay.utwente.nl/view/programme/60025.html
  - Master theses in computer science http://essay.utwente.nl/view/programme/60300.html
  - Master theses in human-media interaction http://essay.utwente.nl/view/programme/60030.html

Exercise

- What is your top-level research problem?
  - A design problem or a knowledge question?

Goal structure: example

Social context

To achieve stakeholder goals: *Reduce national health care cost*

Contributions

To improve a problem context: *To provide mobile home care for the elderly*

Contribution

Design research

To (re)design an artifact:
A remote health monitoring system

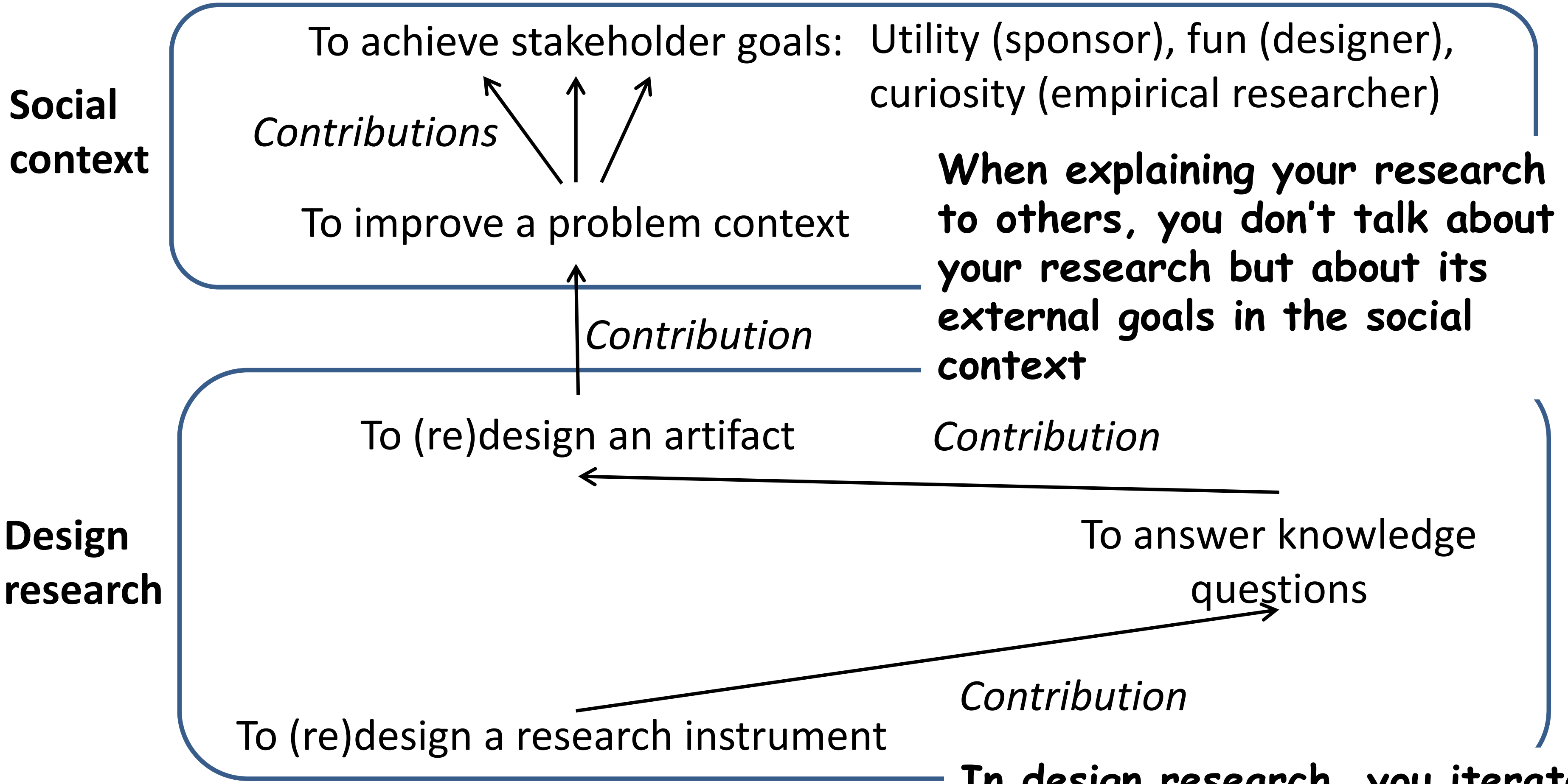
Contribution

To answer knowledge questions: *Is it usable? Does it save time? What quality of care is experienced?*

Contribution

To (re)design a research instruments:
a questionnaire, the setup of a field experiment

Goal structure



When explaining your research to others, you don't talk about your research but about its external goals in the social context

In design research, you iterate over design problems and knowledge questions

Outline

1. Design problems versus knowledge questions
2. The design cycle
3. Design theories
   – Scientific theories
   – Scientific inference: from data to theories
4. The empirical cycle

Engineering cycle

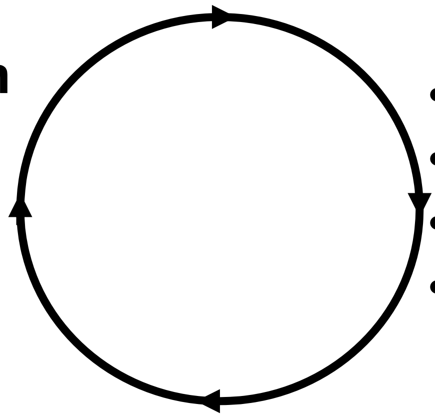
! = Action
? = Knowledge question

This is a checklist. See appendix
A in the book & on my web site

Implementation evaluation =
Problem investigation

- Stakeholders? Goals?
- Conceptual problem framework?
- Phenomena? Causes, mechanisms, reasons?
- Effects? Positive/negative goal contribution?

Design implementation



Treatment 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 = introducing the treatment in the intended problem context

- If the problem is to improve a **real-world** context…. implementation of a solution is **technology transfer to the real world**.
  - Not part of a research project

- If the problem is to learn about the performance of a design ... Implementation of a solution is the **construction of a prototype and test environment, and using it**.
  - Part of a research project

Nesting of cycles

Research project: **design cycle**

Real-world problem investigation	
Treatment design	
Treatment validation	Problem investigation (How to do the validation?)
	Design a prototype & test environment
	Validate a prototype & test environment
	Implement prototype & test environment (lab or field)
	Evaluation (analyze results)
Real-world implementation (tech transfer)	This is a very special engineering cycle, called the **empirical cycle**.
Real-world evaluation (in the field)	

- Do you recognize the structure of your thesis?

Exercise (design-driven thesis) your table of contents

- Make a poster with the outline of the table of contents of your thesis, following this pattern:
  1. Introduction: Societal improvement problem, stakeholders and their goals, current designs, gap with improvement needs.
  2. Research problem: top-level design problem; decomposition into subproblems; knowledge questions
  3. State of the art: existing designs
  4. Requirements for a new design; motivation in terms of stakeholder goals; evaluation of current designs against the requirements
  5. New design
  6. Validation of new design: prototypes, simulations, field experiments, etc.
  7. (More designs and validations)
  8. Conclusions, recommendations, and further work

Exercise (knowledge-driven thesis): your table of contents

- Make a poster with the outline of the table of contents of your thesis, following this pattern:
  1. Introduction: Societal improvement problem, stakeholders and their goals, current knowledge, gap with desired knowledge.
  2. Research problem: Top-level knowledge question; decomposition into sub-questions
  3. State of the knowledge: existing knowledge
  4. Research methods followed
  5. Study: observational study, experimental, case-based, sample-based, etc.
  6. (More studies)
  7. Conclusions, recommendations, and further work

Outline

1. Design problems versus knowledge questions
2. The design cycle
3. Design theories
   – Scientific theories
   – Scientific inference: from data to theories
4. The empirical cycle

What is a theory?

- A **theory** is a belief that there is a pattern in phenomena.
  - Idealizations: *“Merging two faculties reduces cost.” “This works in theory, but not in practice.”*
  - Speculations: *“Elvis lives.” “Jemenites are all terrorists.” “9/11 was executed by the CIA’”*
  - Opinions: *“The Dutch lost the soccer competition because the players are prima donna’s that do not play like a team.’’*
  - Wishful thinking: *``My technique works better than the others.’’*
  - Scientific theories*: Theory of electromagnetism*

What is a scientific theory?

- A **theory** is a belief that there is a pattern in phenomena.
- A **scientific** theory is a belief that there is a pattern in phenomena, that has survived
  - Tests against experience:
    - Observation, measurement
    - Possibly: experiment, simulation, trials
  - Criticism by critical peers:
    - Anonymous peer review
    - Publication
    - Replication

Examples

- *Theory of electromagnetism*
- *Technology acceptance model*

Non-examples

- *Religious beliefs*
- *Political ideology*
- *Marketing messages*
- *Most social network discussions*

What is a scientific design theory?

- A **theory** is a belief that there is a pattern in phenomena.
- A **scientific** theory is a belief that there is a pattern in phenomena, that has survived
  - Tests against experience,
  - Criticism by critical peers.
- A **scientific design theory** is a belief that there is a pattern in the interaction between an artifact and its context, that has survived tests against experience and criticism by critical peers.

Examples:

- *Theory of the UML in software engineering projects*
- *Theory about accuracy and speed of DOA algorithms in a context of plane waves and white noise*
- *Theory about delays in routes planned by MARP on airports*

The structure of scientific theories

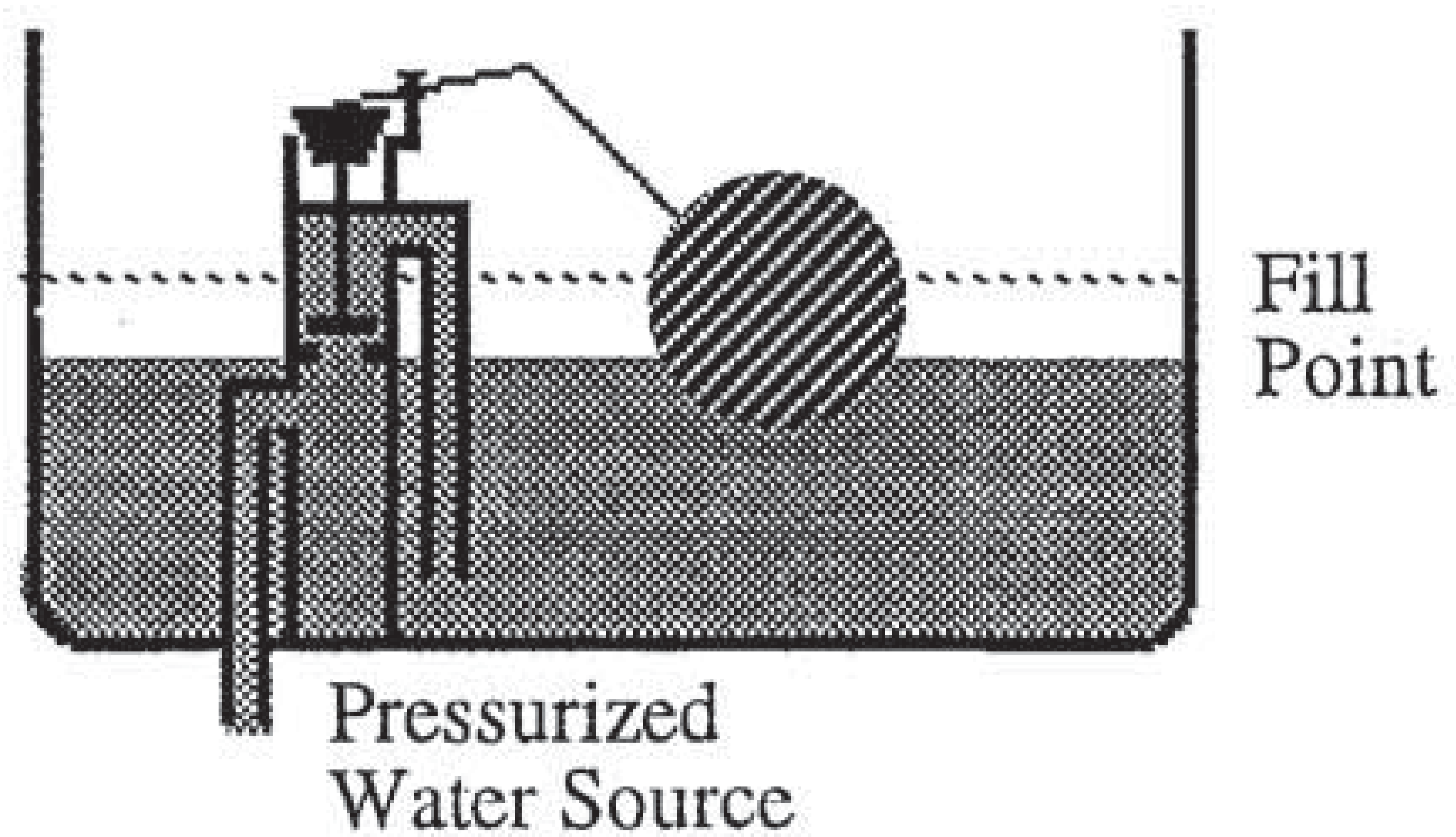
1. **Conceptual framework**
   – Definitions of concepts.
2. **Generalizations**
   – Express beliefs about patterns in phenomena.

Design theory of the UML

- *Concepts: UML concepts, definitions of software project, of software error, project effort, definition of concept of domain, understandability*
- *Descriptive generalization: (UML) X (SE project) → (Less errors, less effort than similar non-UML projects)*
- *Explanatory generalizations:*
    - *UML models resemble the domain more than other kinds of models;*
    - *They are easier to understand for software engineers;*
    - *So they they make less errors and there is less rework (implying less effort).*

Design theory of an algorithm to estimate direction of arrival of plane waves

- *Concepts:*
  - *Definitions of concepts to specify a direction-of-arrival recognition algorithm, and of concepts to describe antenna array, and of accuracy and excution time*
- *Descriptive generalization:*
  - *(Algorithm MUSIC) x (antenna array, plane waves, white noise) → (execution time less than 7.2 ms, accuracy 1 degree)*
- *Explanatory generalization*
  - *Algorithm structure explains functional correctness of output*
  - *(No explanation of exact performance numbers)*



- Conceptual model of an artifact architecture.
- Together with a narrative of the mechanism, <u>this diagram is a design theory</u> of an artifact.

Theory of electromagnetism

- *Conceptual framework (concepts defined to describe and explain the relevant phenomena):*
  - *Definitions of electric current, electric charge, potential difference, electric resistance, electric power, capacitance, electric field, magnetic field, magnetic flux density, inductance, …, … and their units.*
- *Generalizations*
  - *Electric charges attract or repel one another with a force inversely proportional to the square of their distance.*
  - *Magnetic poles attract or repel one another in a similar way and always come in North-South pairs.*
  - *An electric current inside a wire creates a corresponding circular magnetic field outside the wire.*
  - *A current is induced in a loop of wire when it is moved towards or away from a magnetic field*

All generalizations can be used to make predictions

- A **general problem theory** describes and explains a type of problem: Typical symptoms and diagnosis.
- **A general design theory** describes and possibly explains interaction between Artifact and Context in general.

- Both theories generalize, and so may be used to predic:
  - What will happen if the problem is untreated?
  - What will happen if the treatment is applied?

To design = to build a design theory

Update the design

- Create a design
- Create a theory about your design in context
- Test this theory
  - Analytically
  - Empirically (prototyping etc.)

Update the theory

Which is first?

1. Specify your desired descriptions & explanations
2. Design an artifact that, together with the intended context, satisfies those descriptions & explanations

1. Design an artifact for an the intended context that contributes to stakeholder goals
2. Describe & explain its behavior

Outline

1. Design problems versus knowledge questions
2. The design cycle
3. Design theories
   – Scientific theories
   – Scientific inference: from data to theories
4. The empirical cycle

Descriptions, generalizations, explanations

- Descriptive knowledge questions:
  - What happened?
  - How much? How often?
  - When? Where?
  - What components were involved?
  - Who was involved?
  - Etc. etc.
- Explanatory knowledge questions:
  - Why?
    - What caused this phenomenon?
    - What mechanisms produced it?
    - Why did people do this?

Yields descriptions of facts of the case(s)

May be generalized **beyond** the facts of the case(s)

Yields explanations **beyond** the facts of the case(s)

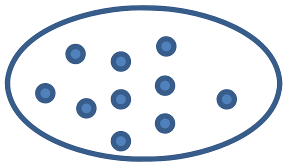
May be generalized **beyond** the explanations of the case(s)

Each bullet is an artifact in a context

From facts to theories

Descriptions of facts of case(s)

Observed sample of cases

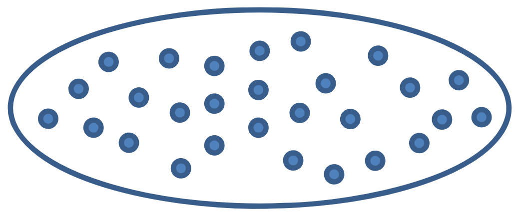


- What happens in these cases?
- What average, variance in this sample?

Generalize

Descriptive theory of the population

Unobserved population



- What happens in all cases?
- What average, variance in this population?

Explain

- Why?

Explanatory theory of the case/sample

Explain

- Why?

Explanatory theory of the population

- Theories may be general or particular
  - They may state that there is a pattern in the phenomena in a population
  - They may indicate that one case exhibits an instance of a pattern

Each bullet is an artifact in a context

From facts to theories

Descriptions of facts of case(s)

Observed sample of cases

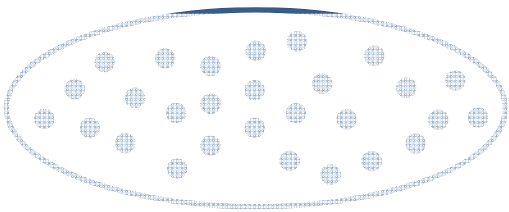
Physician
Detective
Lawyer
Engineer
Manager
Politician
Parent

- What happens in these cases?
- What average, variance in this sample?

Generalize

Descriptive theory of the population

Unobserved population



- What happens in all cases?
- What average, variance in this population?

Researcher

Explain

- Why?

Explanatory theory of the case/sample

Explain

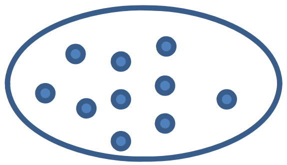
- Why?

Explanatory theory of the population

Three kinds of explanation

Descriptions of facts of case(s)

Observed sample of cases

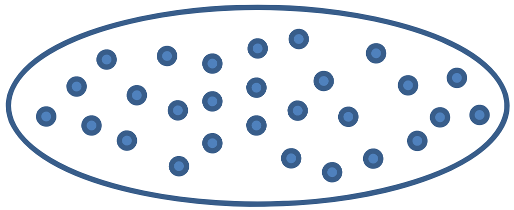


- What happens in these cases?
- What average, variance in this sample?

Generalize

Descriptive theory of the population

Unobserved population



- What happens in all cases?
- What average, variance in this population?

Explain by
- Causes
- Mechanisms
- Reasons

- Why?

Explanatory theory of the case/sample

Explain by
- Causes
- Mechanisms
- Reasons

- Why?

Explanatory theory of the population

Example explanations (1)

- *Descriptive question: Is the light on?*
  - *Based on observation: Yes.*
  - *When? Now.*
  - *Where? Here.*
- *Explanatory question: Why is it on?*
  1. ***Cause:** because someone turned the light switch, it is on (and not off).* Explains difference with off-state.
  2. *Why does this cause the light to switch on? **Mechanism:** because the switch and light bulbs are connected by wires to an electricity source, in this architecture …, and these components have these capabilities …..* Explains how on-state is produced.
  3. *By why did someone turn the light on? **Reasons:** Because we wanted sufficient light to be able to read, and it was too dark to read.* Explains which stakeholder goal is contributed to.

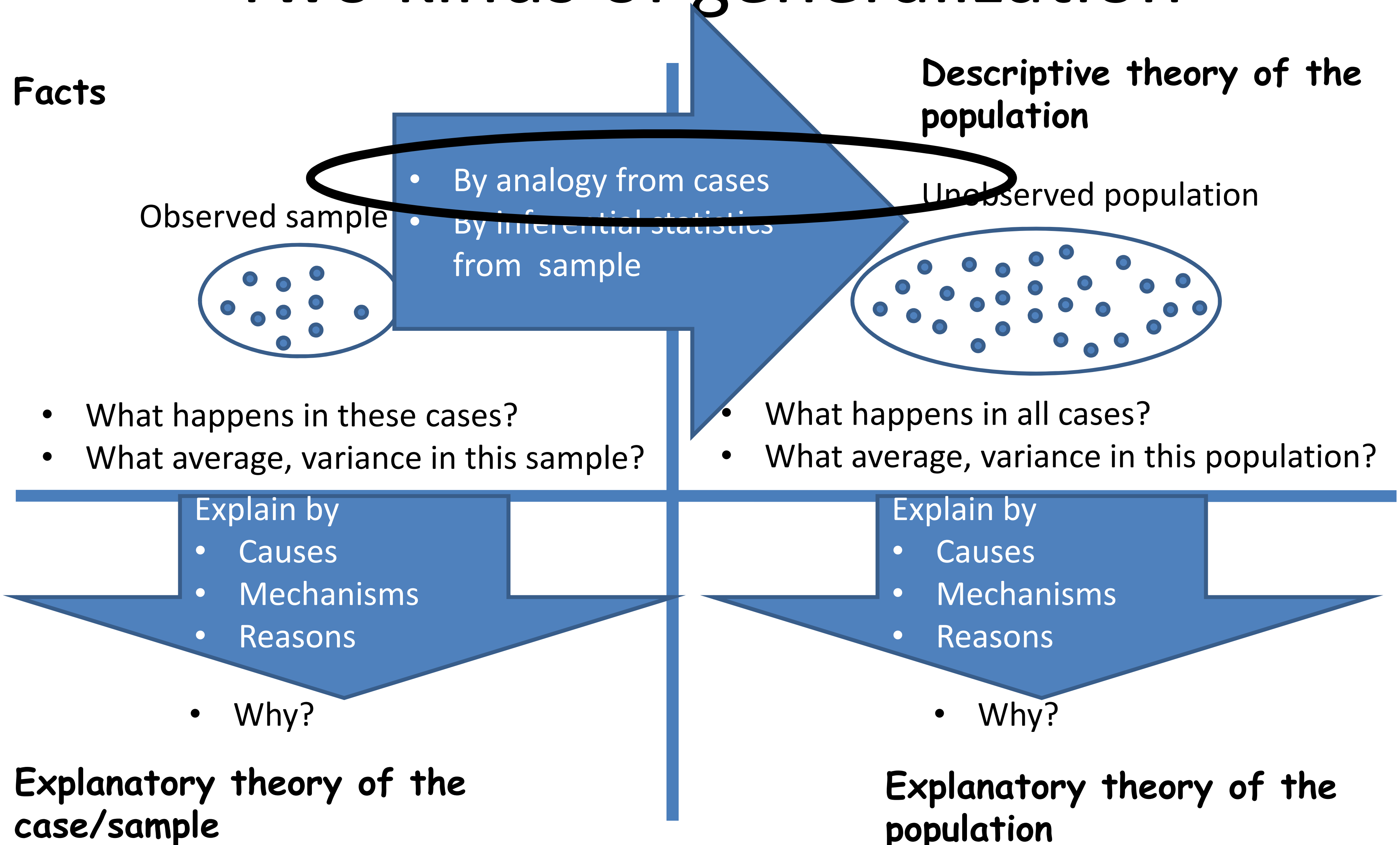
Example explanations (2)

- *Descriptive question: What is the performance of this program?*
  - *Execution time for different classes of inputs?*
  - *Memory usage?*
  - *Accuracy?*
  - *Etc. etc.*
- *Explanatory question: Why does this program have this performance (compared to others)?*
  1. ***Cause:*** *Variation in execution time is caused by variation in input; etc.*
  2. ***Mechanism:*** *Execution time varies this way because it has this architecture with these components*
  3. ***Reasons:*** *Observed execution time varies this way because users choose to drive on busy roads with a lot of signal interference*

Example explanations (3)

- *Descriptive question: What is the performance of this method for developing software?*
  - *Understandability for practioners*
  - *Learnability*
  - *Quality of the result*
  - *Perceived utility*
  - *Etc. etc.*
- *Explanatory question: Why does this method have this performance?*
  1. ***Cause:*** *Difference in project performance is attributed to difference between UML and non-UML methods.*
  2. ***Mechanism:*** *The difference in effects is by the match between UML and the structure of cognition.*
  3. ***Reasons:*** *Difference in performance may be explainable by difference in motivation of developers to use UML or something else.*

Two kinds of generalization



Case-based generalization (1)

- *Observation:*
  - *Artifact: A light switch*
  - *Context: next to the door in the wall of a room with ceiling lights*
  - *Effect: toggles the ceiling light on and off.*
- *Explanation:*
  - *The switch and context architectures produce this behavior*
- *Generalization by analogy:*
  - *All **similar** switches*
  - *Running in **similar** contexts*
  - *Will show **similar** effects*

Descriptive generalization. Implicit assumptions:

1. Similar artifacts and contexts have similar components with capabilities, and
2. There is no interference from other mechanisms.

Case-based generalization (2)

- *Observation:*
  - *Artifact: This prototype implementation of the MUSIC algorithm,*
  - *Context: when used to recognize direction of arrival of plane waves received by an antenna array, in the presence of only white noise, running on a Montium 2 processor,*
  - *Effect: has execution speed less than 7.2 ms and accuracy of at least 1 degree.*
- *Explanation:*
  - *Algorithm structure*
- *Generalization by analogy:*
  - *All* ***similar*** *implementations*
  - *Running in* ***simila****r contexts*
  - *Will show* ***similar*** *performance*

Descriptive generalization. Implicit assumptions:

1. Similar artifacts and contexts have similar components with capabilities, and
2. There is no interference from other mechanisms.

Case-based generalization (3)

- *Observations:*
  - *Artifact: this version of the UML*
  - *Context: Used in this software project*
  - *Effect: Produces software with less errors and less effort than in similar projects without the UML,*
- *Explanation:*
  - *UML models are easier to understand for software engineers because they resemble the domain more than other kinds of models,*
  - *so the software engineers make less errors and there is less rework.*
- *Generalization*
  - *In similar projects, UML will have similar effects*
  - *Assumptions: Similar projects have software engineers and tools with similar capabilities, and*
  - *The effects will not be undone by other mechanisms*

Case-based generalization (generally)

- Case-based generalization = analogic generalization.
- We observe some mechanism in a particular case.
- Assumptions of analogic generalization:
  1. Similar artifacts and contexts have similar components with capabilities, and
  2. There is no interference from other mechanisms.

- Similarity conditions:
  - Similar cases have components with similar capabilities
  - Similar cases have similar mechanisms involving these components

- Analogy based in similarity of superficial features, without knowledge of underlying mechanisms, is too weak a basis for generalization:

  – *Wallnuts look like brains.*
  – *Brains can think.*
  – *Therefore .... Wallnuts can think*

- *There is no shared mechanism that produces thinking in brains and wallnuts!*

Fallibility

- Generalization based on architectural similarity gives uncertain conclusions
- Explicitly describe this uncertainty

- Reduce this uncertainty by replication!

Sample-based generalization

Facts

Observed sample

- By analogy from cases
- By inferential statistics from sample

Descriptive theory of the population

Unobserved population

- What happens in these cases?
- What average, variance in this sample?

- What happens in all cases?
- What average, variance in this population?

Explain by
- Causes
- Mechanisms
- Reasons

- Why?

Explain by
- Causes
- Mechanisms
- Reasons

- Why?

Explanatory theory of the case/sample

Explanatory theory of the population

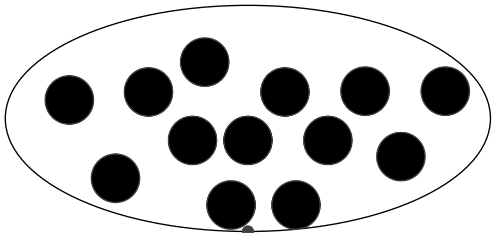
Statistical inference

Do you *agree* with these distinctions?

1. **By big data:** If the sample is almost the size of the population, then the population probably has similar statistics.
   – Only true if the sample is random. Law of large numbers.
2. **By statistical learning:** Use a sample of (X, Y) values to estimate Y as a function of X in the population.
   – E.g. regression. Different methods come with different assumptions.
3. **Bayesian inference.** Use a sample to update a hypothesized distribution of a variable over the population
   – Need to start with an initial hypothesized distribution.
4. **Frequentist statistical inference:** In repeated random sampling from the same population, the sample averages are approximately normally distributed around the population mean.
   – Central-limit theorem. Assumes random samples.

Methodology of statistical inference

Theoretical population

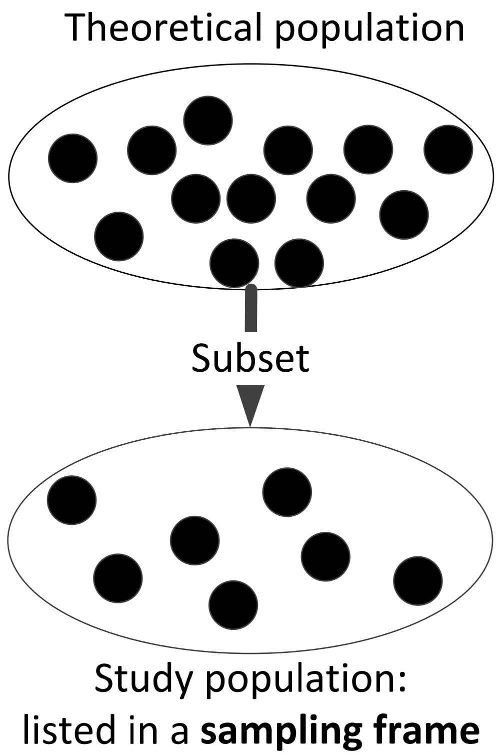


E.g.

- *The set of all instances of an algorithm running in a context;*
- *The set of all global SE projects;*
- *Etc.*

Our ultimate ***target of generalization***

Methodology of statistical inference



Research methodology:

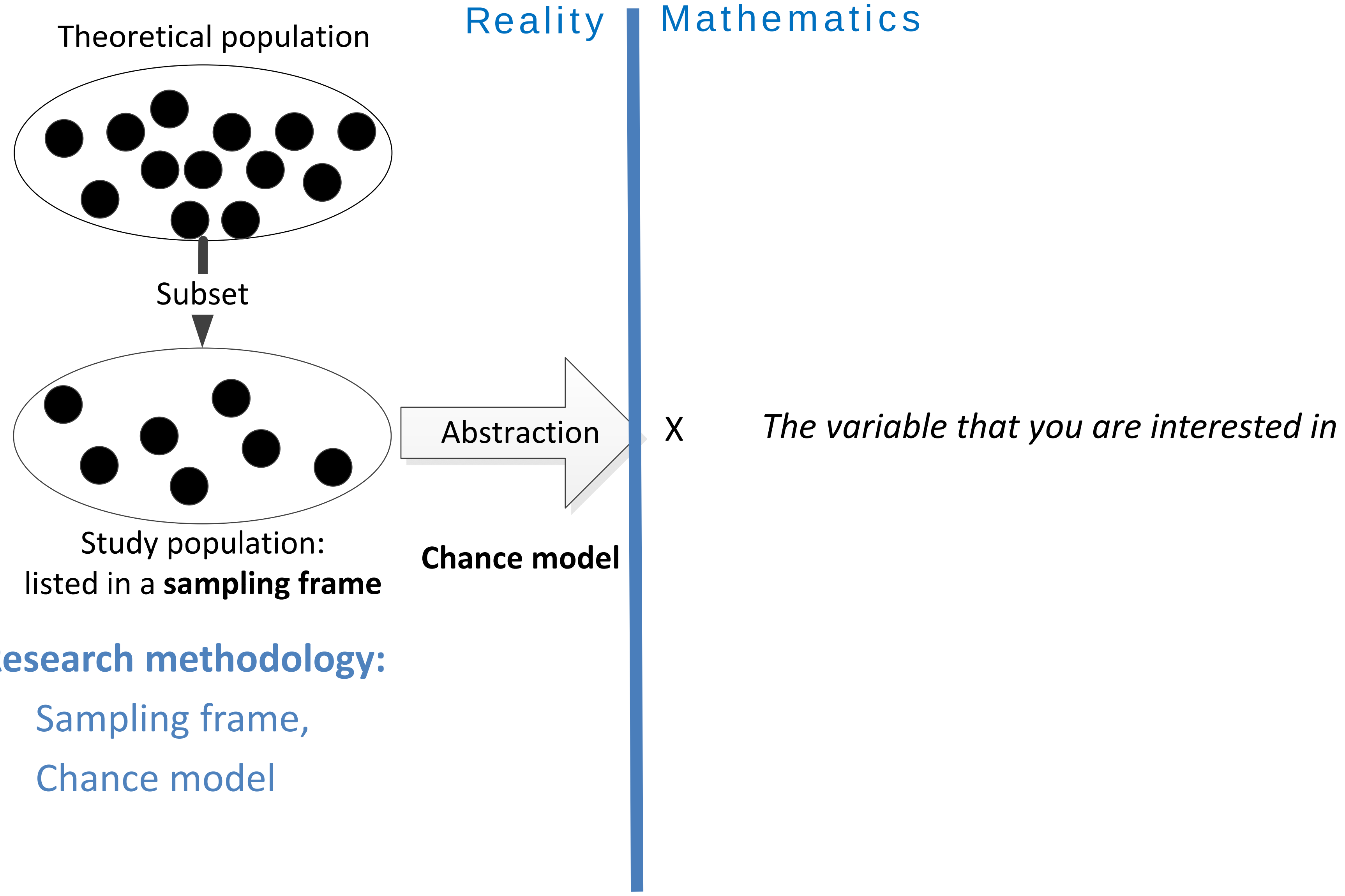
- Sampling frame

E.g.

- *The set of all **prototype instances** of an algorithm running in a **laboratory** context;*
- *The set of all global SE projects **engaged in by company A**;*
- *Etc.*

The population elements from which you will select a sample

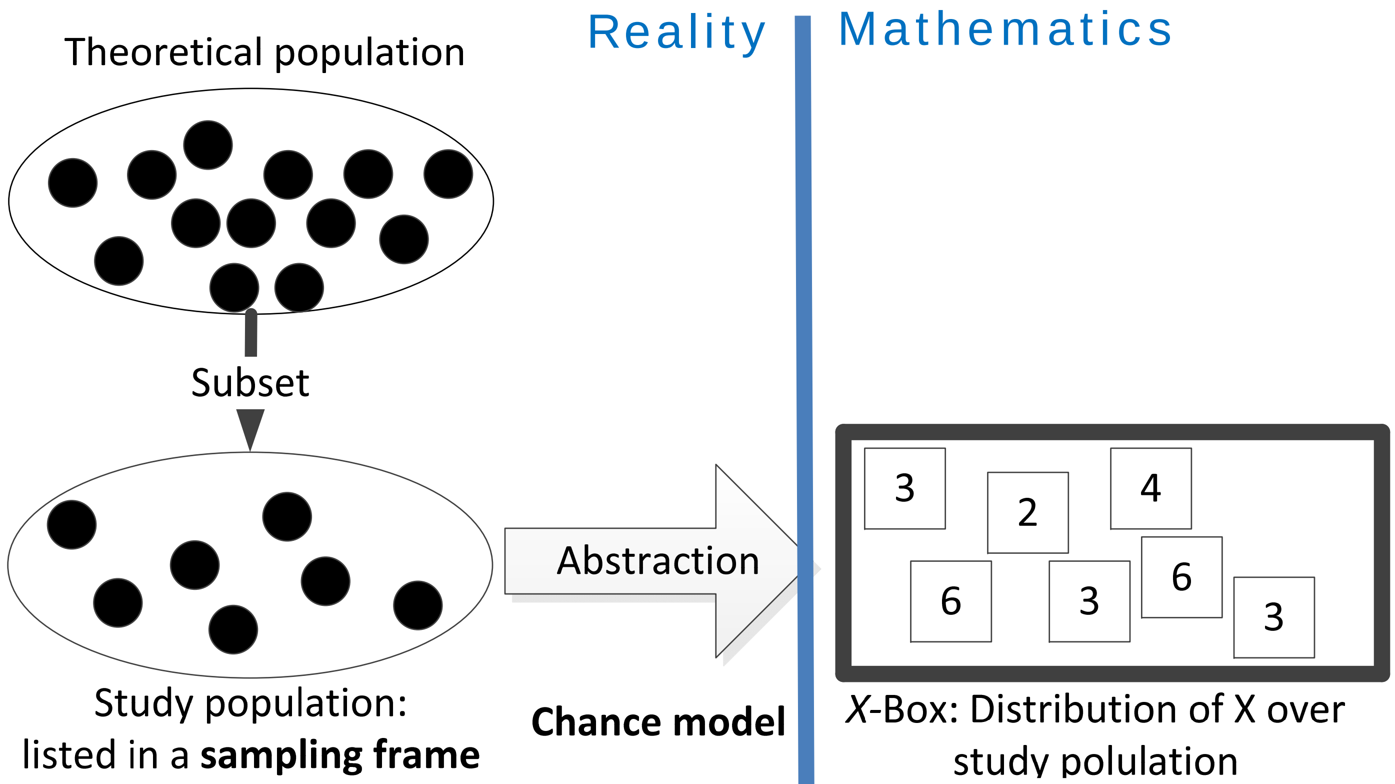
Methodology of statistical inference



Research methodology:

- Sampling frame,
- Chance model

Methodology of statistical inference



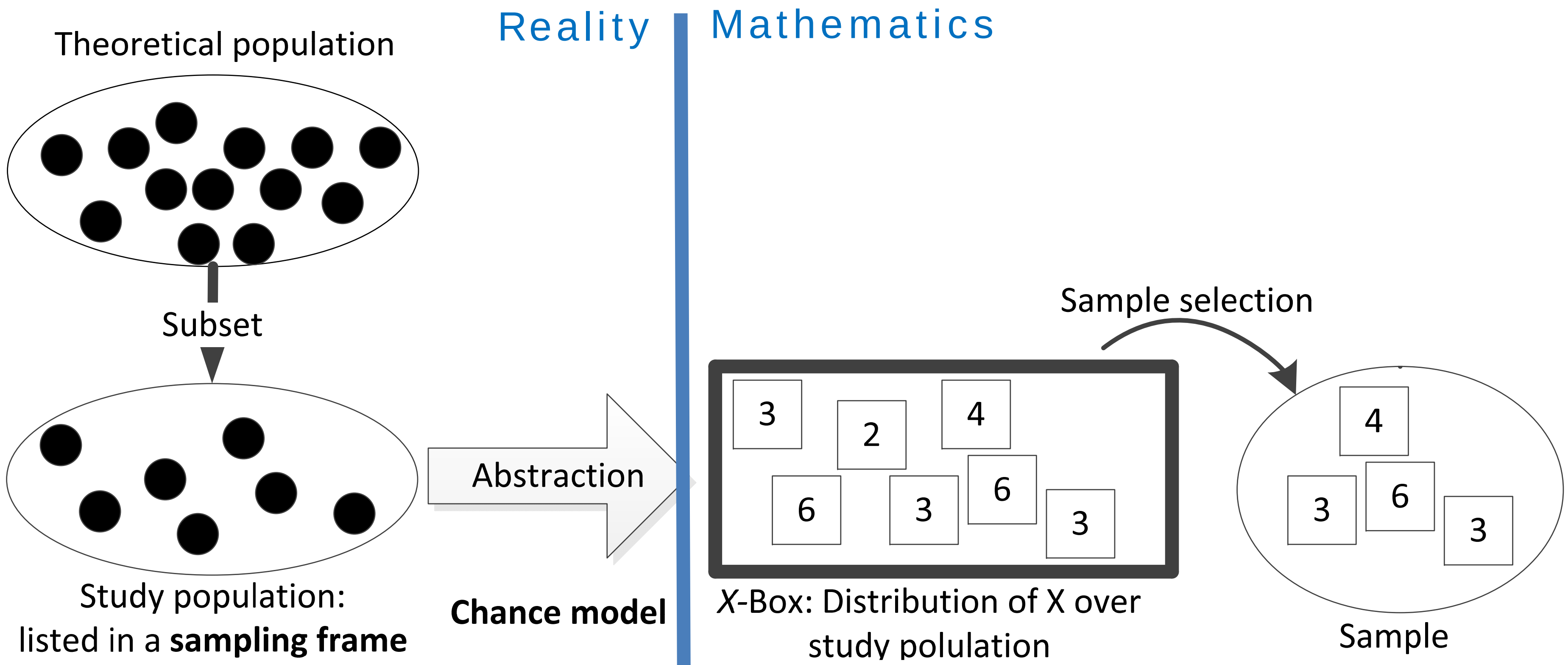
Statistical inference.

- Unobservable distribution of numbers

Research methodology.

- Sampling frame,
- Chance model

Methodology of statistical inference



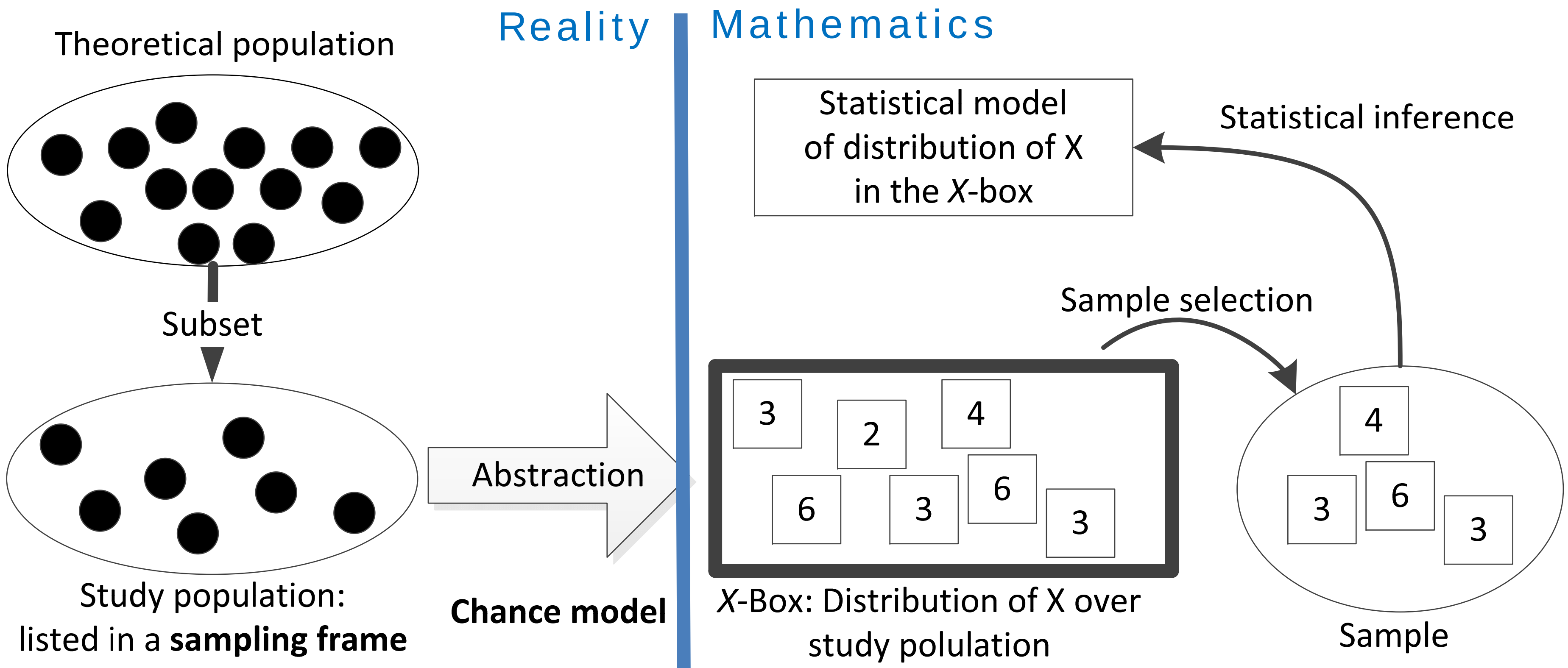
Research methodology.

- Sampling frame,
- Chance model

Statistical inference.

- Unobservable distribution of numbers,
- Selection of observable sample,

Methodology of statistical inference



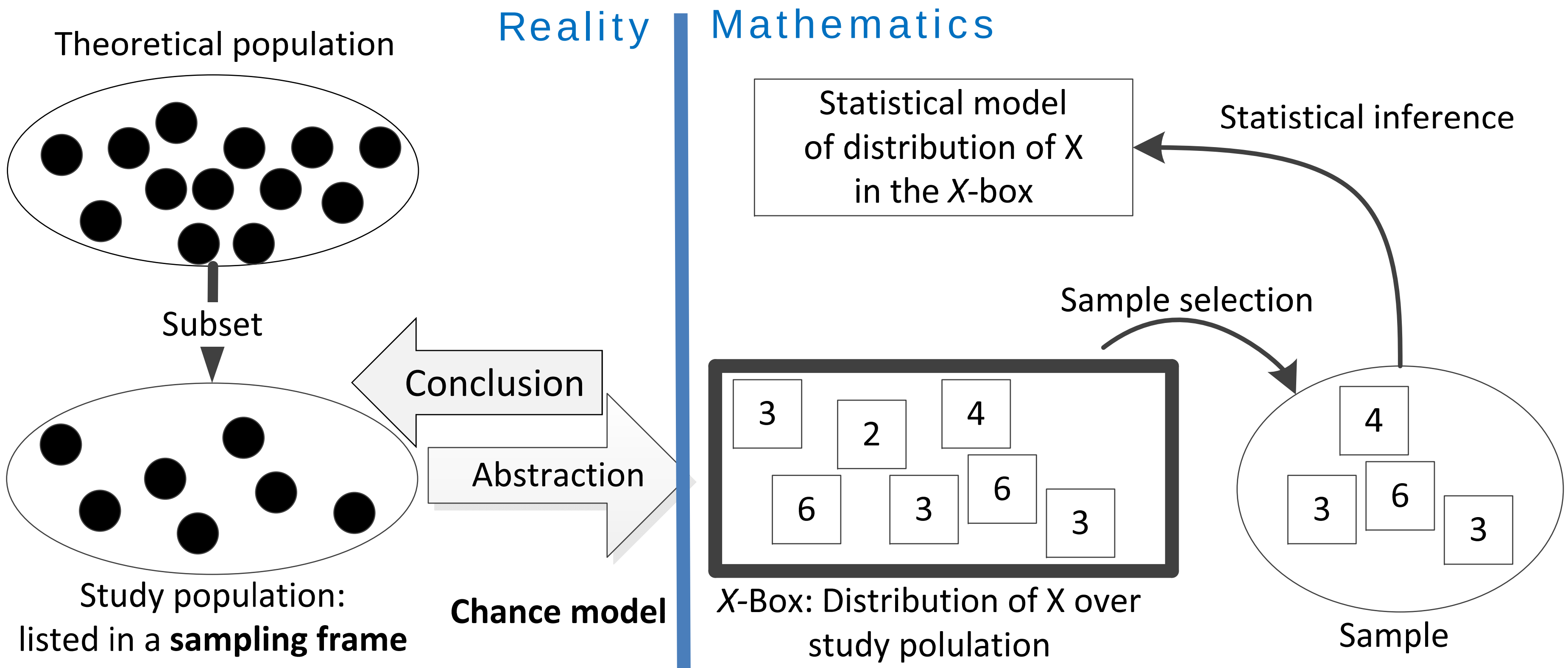
Research methodology.

- Sampling frame,
- Chance model

Statistical inference.

- Unobservable distribution of numbers,
- Selection of observable sample,
- Conclusion about unobservable distribution of numbers

Methodology of statistical inference



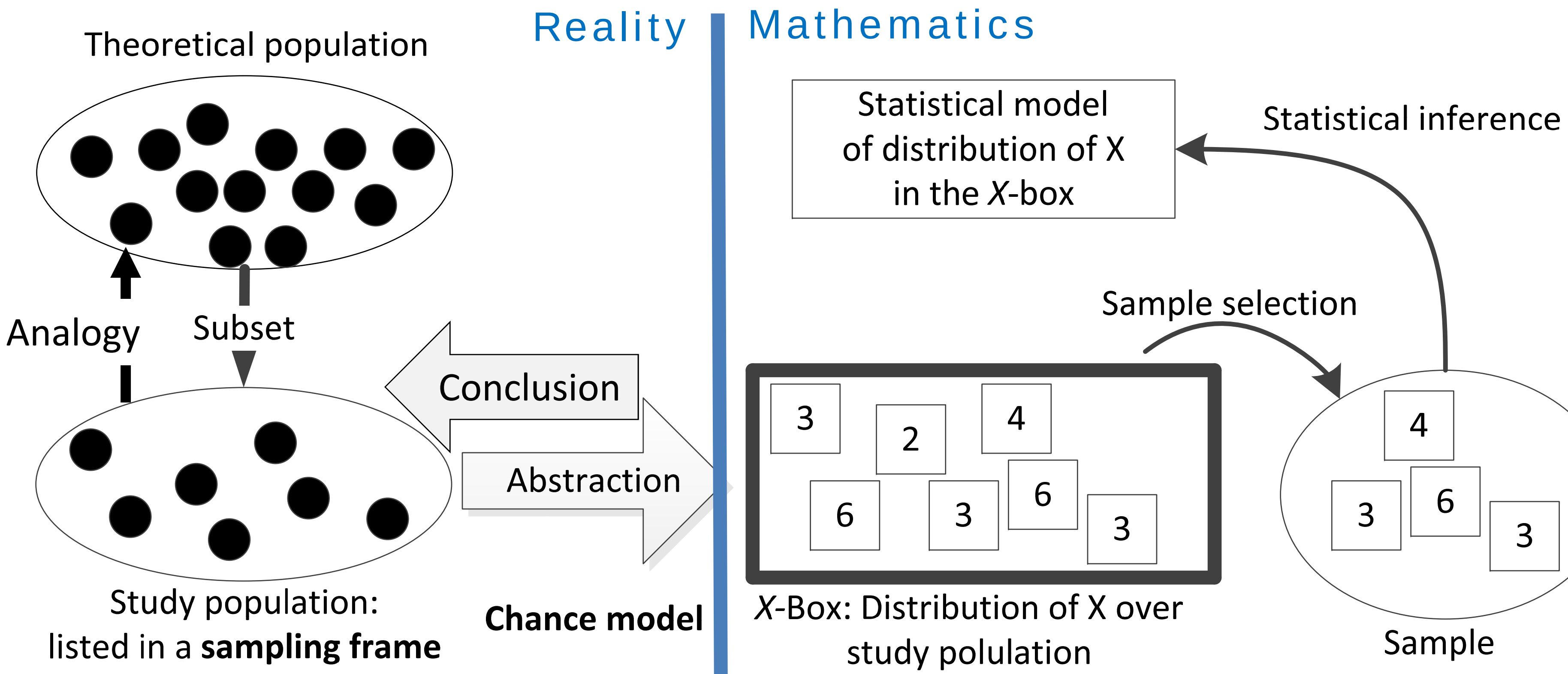
Research methodology.

- Sampling frame,
- Chance model,
- Conclusion about study population

Statistical inference.

- Unobservable distribution of numbers,
- Sample selection,
- Conclusion about unobservable distribution of numbers

Methodology of statistical inference



Research methodology.

- Sampling frame,
- Chance model,
- Conclusion about study population,
- Conclusion about theoretical population.

Statistical inference.

- Unobservable distribution of numbers,
- Sample selection,
- Conclusion about unobservable distribution of numbers

Statistical inference

1. **By big data:** If the sample is almost the size of the population, then the population probably has similar statistics.
   – Only true if the sample is random. Law of large numbers.

- **The sample is very large.**
- **Skip statistical inference, but not the other steps**
- **Take care of assumptions**

2. **By statistical learning:** Use a sample of (X, Y) values to estimate Y as a function of X in the population.
   – E.g. regression. Different methods come with different assumptions.

- **The sample is of sufficient size and possiblty very large.**
- **Do statistical inference**
- **Take care of assumptions**

3. **Bayesian inference.** Use a sample to update a hypothesized distribution of a variable over the population.
   – Need to start with an initial hypothesized distribution.
4. **Frequentist statistical inference:** In repeated random sampling from the same population, the sample averages are approximately normally distributed around the population mean.
   – Central-limit theorem. Assumes random samples.

Fallibility

- Generalization based on statistical inference gives uncertain conclusions
- Explicitly describe this uncertainty

- Reduce this uncertainty by replication!

Statistical inference

1. **By big data:** If the sample is almost the size of the population, then the population probably has similar statistics.
   – Only true if the sample is random. Law of large numbers.
2. **By statistical learning:** Use a sample of (X, Y) values to estimate Y as a function of X in the population.
   – E.g. regression. Different methods come with different assumptions.
3. **Bayesian inference.** Use a sample to update a hypothesized distribution of a variable over the population
   – Need to start with an initial hypothesized distribution.
4. **Frequentist statistical inference:** In repeated random sampling from the same population, the sample averages are approximately normally distributed around the population mean.
   – Central-limit theorem. Assumes random samples.

Four varieties of frequentist statistical inference

- **Fisher:** Test a null hypothesis that is unlikely, given what you know
- **Neyman-Pearson:** Decide between alternative hypotheses, based on a previously set of error rates
- **Neyman:** Estimate a confidence interval of a distribution parameter
- **Social sciences:** ~~Null Hypothesis Significance Testing (NHST).~~
  ~~Misconceived and logically incorrect mix of Fisher & Neyman-Pearson~~

- R.B. Kline. *Beyond Significance Testing. Statistics Reform in the Behavioral Sciences.* Second edition. American Psychological Association, 2013.
- G. Cumming. *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis.* Routledge 2012.

Null-hypothesis significance test (NHST)

- Null hypothesis: Assume the population mean of X is 0.
- Compute your observed sample mean m.
- If the probability of observing at least m is less than 5%, then reject H0.

Computed using CLT

Misconception 1 of NHST: Fixed decision rule

- Why 5%? What if p-value = 4.9% or 5.1%?
  - Outcome of hypothesis test should be combined with what we know from earlier tests and from established theory.
- Impact of NHST rule:
  - Published p-values crowd just below 5% (“p-hacking’’).
  - Just above 5% they are sparse (‘’publication bias’’)

Misconception 2 of NHST: Probabilistic falsification

- Rule of falsification
  - If $p \rightarrow q$ and we observe $\neg q$, then $\neg p$.
- There is no valid rule of probabilistic falsification
  - If p *implies that* q is improbable and we observe q, then no conclusion.

Misconception 3 of NHST: If H_0 is false, then H_1 is true

- There are many alternatives to H_0!
- In NHST, H_0 is **not a substantial hypothesis** but a hypothesis of no difference between treatment group and control group
  - If we reject H_0 then we can only conclude that "something is going on"
  - But we knew this already.

Misconception 4 of NHST

- If the (probability of a difference ≥ d, given H_0) is 5% or less, then
  – we **cannot** conclude that the probability of H_0 is at most 5%;
  – And we **cannot** conclude that the probability of H_1 is at least 95%;

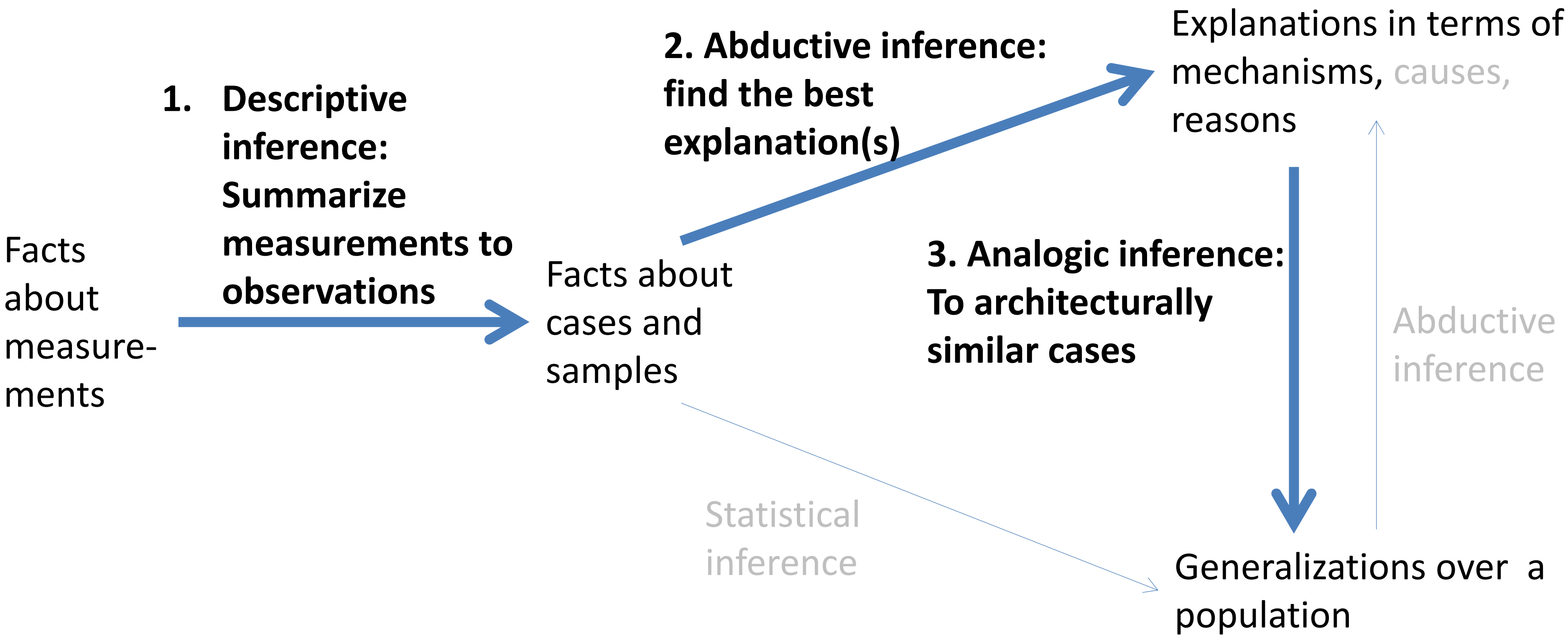
- 80% of Dutch people are blond ↔
- 80% of blond people are Dutch

Misconception 5 of NHST

- Between treatment group and control group there is always a non-zero difference!
  - it would be a miracle it the two sample means were identical.
- If the two samples are large enough, any non-zero difference can be discerned statistically. Mathematical theorem.

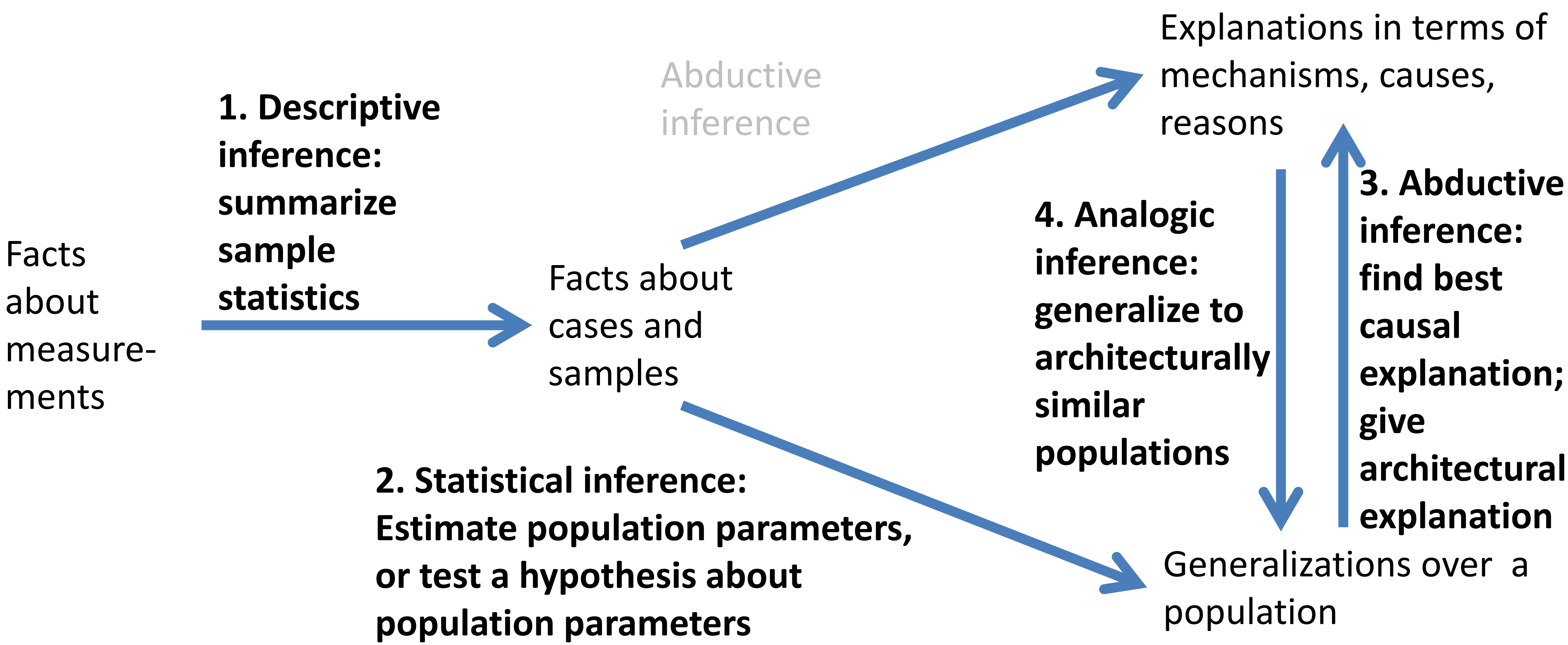
Summary of inferences

Case-based inference



- Analogic inference to similar cases must be based on architectural explanations (in terms of mechanisms or reasons)

Sample-based inference

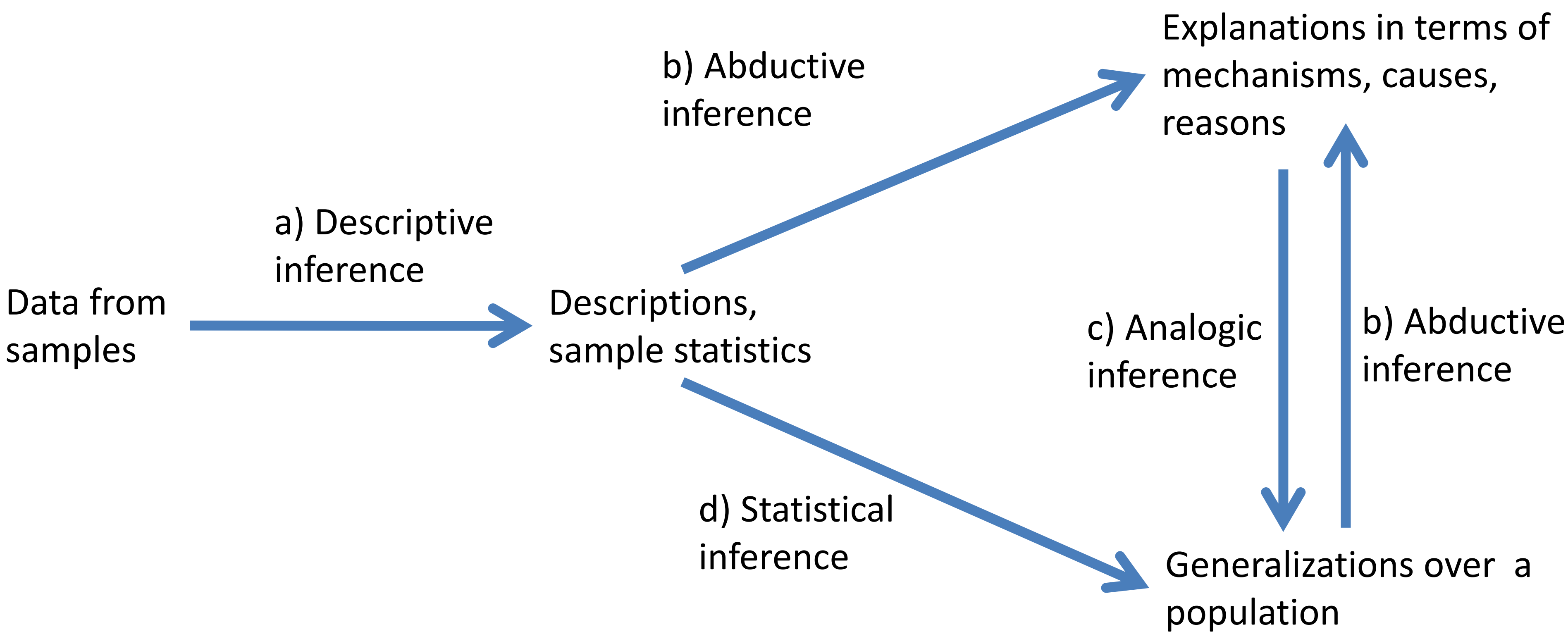


- Statistical inference yields descriptive generalization over a study population.
- Analogic generalization to similar populations must be based on architectural explanation of those causes.

Validity

- Outside mathematics there is no certainty
- All conclusions of empirical research are fallible
  - All conclusions of empirical research are improvable
  - We need to indicate to how and why are conclusions could be wrong!
  - Validity: degree of support for a conclusions

Validity of inferences: degree to which they are justified

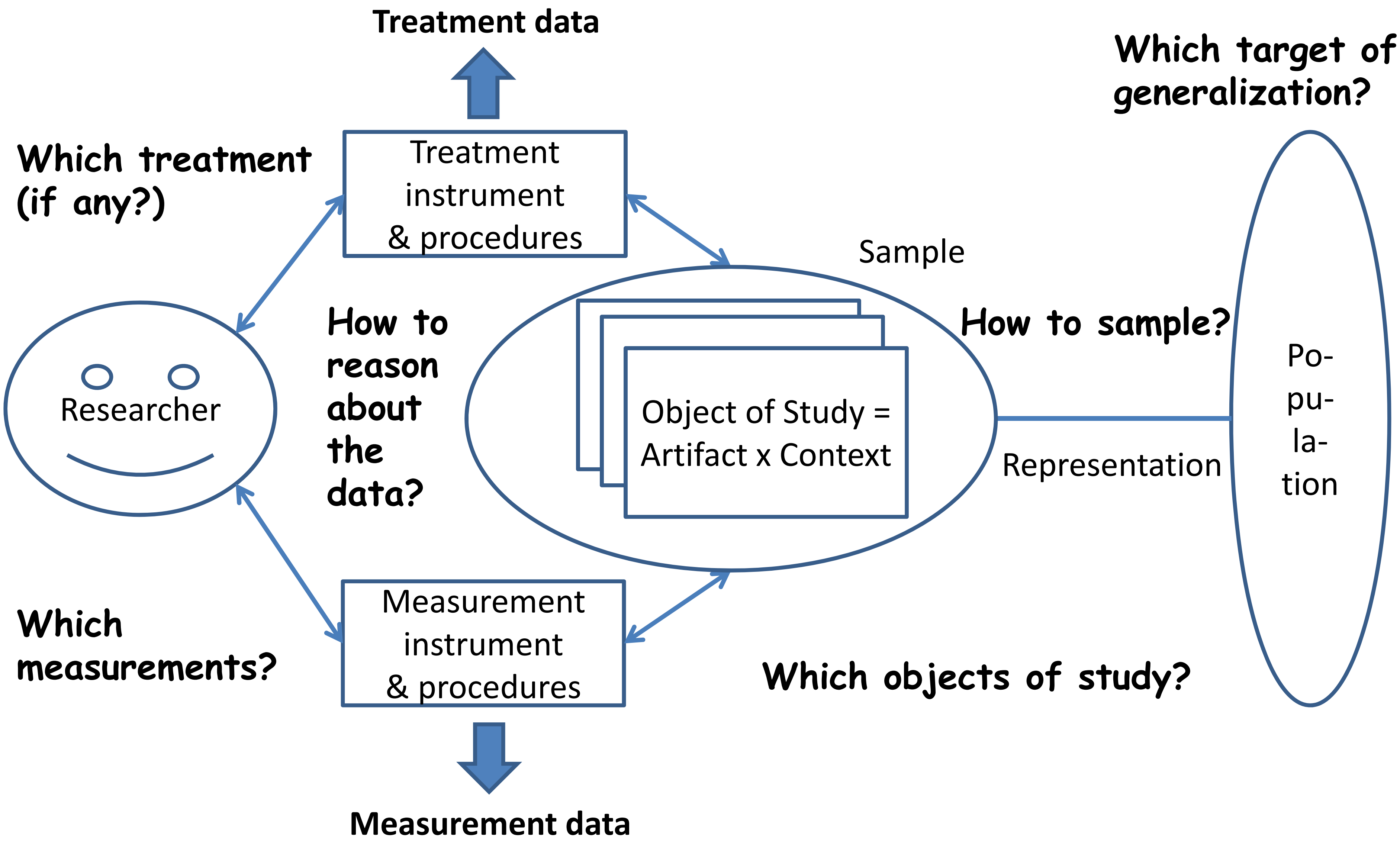


a) **Descriptive validity:** no information added in the descriptions
b) **Internal validity:** degree of support for explanations
c) **External validity:** degree of support for analogic generalizations
d) **Statistical conclusion validity:** degree of support for statistical inference

Outline

1. Design problems versus knowledge questions
2. The design cycle
3. Design theories
   - Scientific theories
   - Scientific inference: from data to theories
4. The empirical cycle

Design decisions for research setup



Research designs and inferences and their role in the design cycle

	Observational study (no treatment)	Experimental study (treatment)
Case-based: investigate single cases, look at architecture and mechanisms. <u>**Inference: Architectural explanation, generalization by analogy**</u>	**Observational case study**	• **Expert opinion** (mental simulation by experts), • **Case-based experiment** (simulations, prototyping), • **Technical action research** (experimental use of the artifact in the real world)
Sample-based: investigate samples drawn from a population, look at averages and variation. <u>**Inference: Statistical inference, causal explanation, possible architectural explanation and analogy**</u>	**Survey** **Real-world problem investigation / implementation evaluation methods**	• **Sample-based experiment** (treatment group – control group experiments) **Treatment validation methods (depends on budget)**

Checklist for research design: context

This is a checklist for
- **research design,**
- **research reporting,**
- **reading a report.**

App. B in my book & my web site

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

Design cycle

Empirical cycle

4. …

….

16. …

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

Designing something useful

Answering a knowledge question

Data analysis

12. Descriptions?
13. Statistical conclusions?
14. Explanations?
15. Generalizations?
16. Answers?

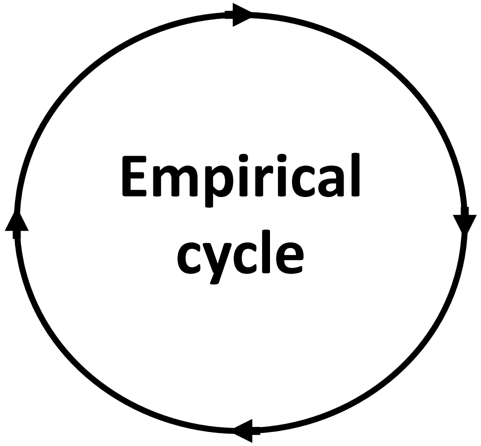
This is a checklist for

- **research design,**
- **research reporting,**
- **reading a report.**

App. B in my book & my web site

Research execution

11. What happened?



Research problem analysis

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

Design validation

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

Research & inference design

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

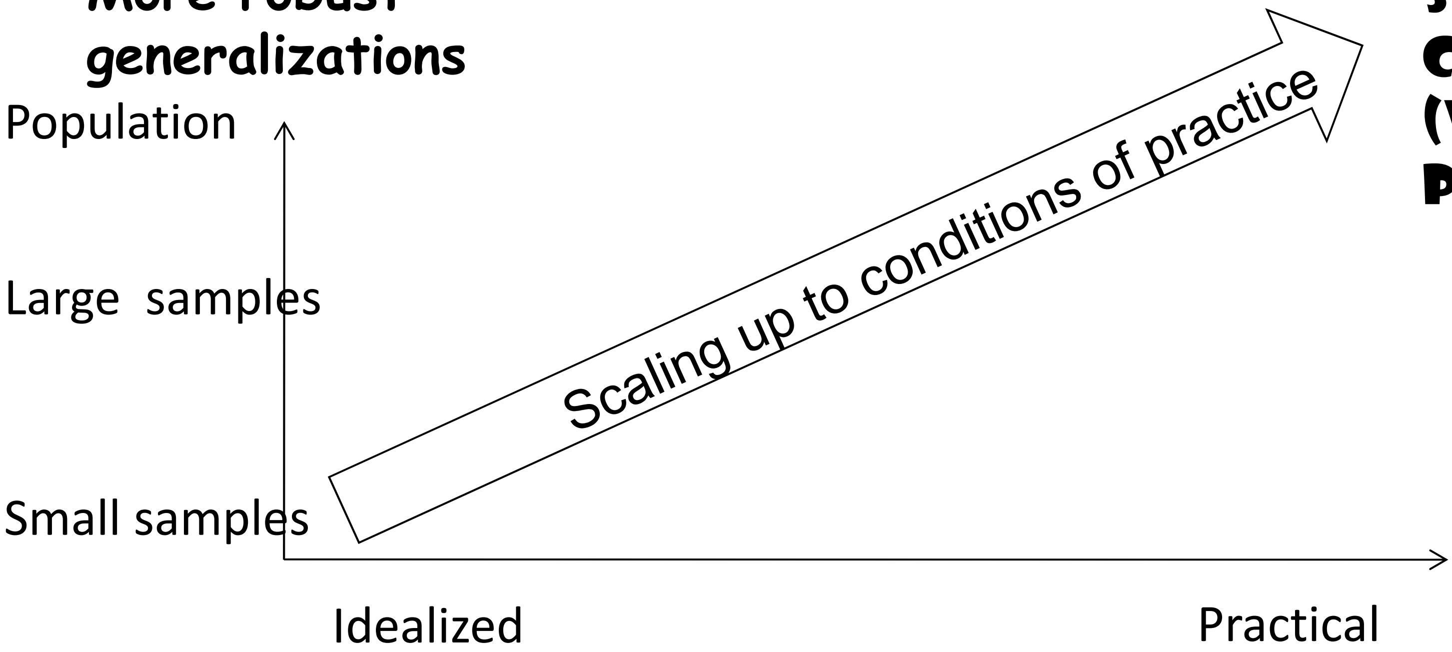
Research setup

Inference

Design science research strategy

More robust generalizations

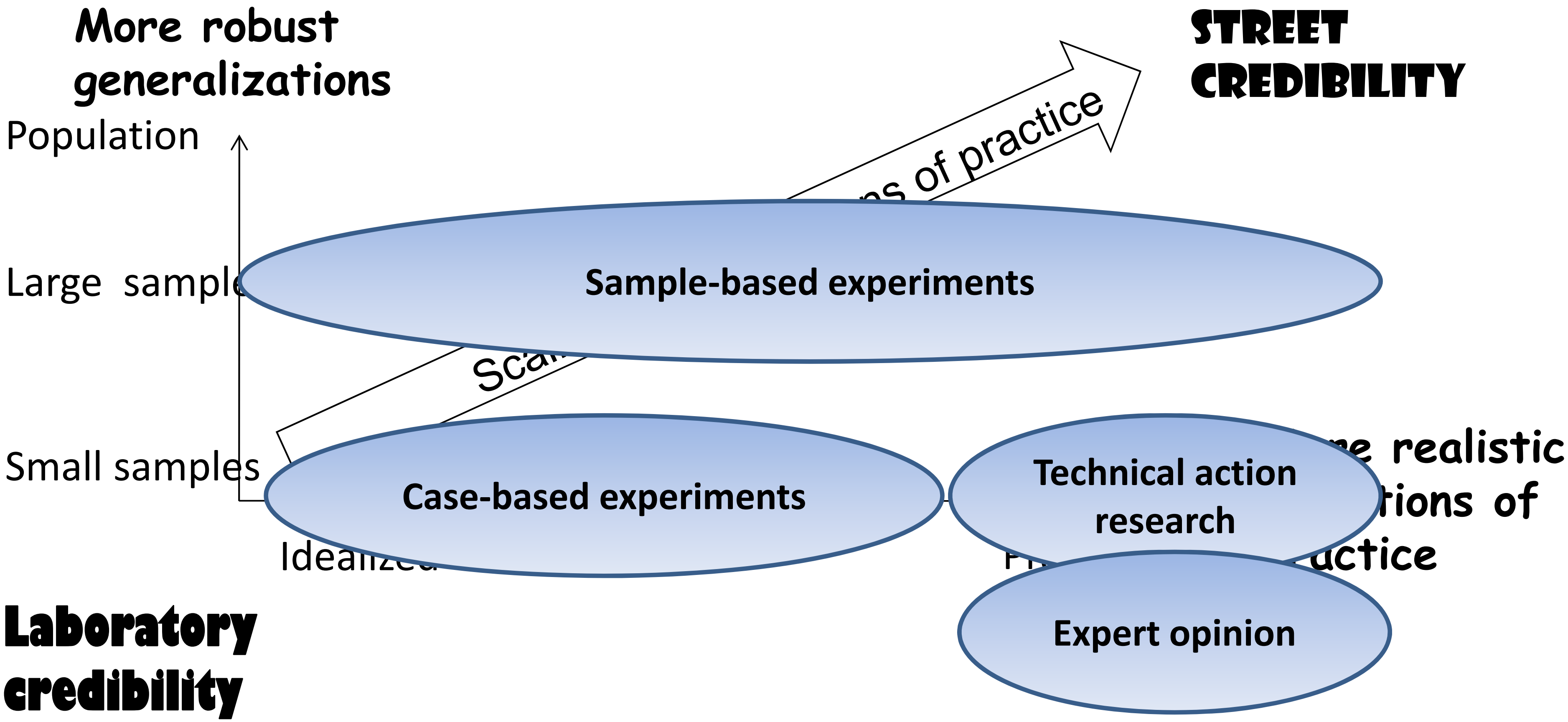
STREET CREDIBILITY (WORKS IN PRACTICE)



More realistic conditions of practice

Laboratory credibility (works in theory)

- Just like New Drug Research



- Scaling up:
  - Case-based experiments (laboratory simulation)
  - Expert opinion
  - Case-based experiments (field simulation)
  - TAR (apply technique in a real-world project)

Outline

1. Design problems versus knowledge questions
2. The design cycle
3. Design theories
   - Scientific theories
   - Scientific inference: from data to theories
4. The empirical cycle

Summary

- What is the problem?
- What diagnoses are there of the problem?
- Who are the stakeholders?
- What are their goals?
- Artifact x Context → Effects?
- Effects satisfy requirements?
- Requirements contribute to goals?

- General problem descriptions & explanations
- General design-in-context descriptions & explanations

Exercise (design-driven thesis): your table of contents

- Make a poster with the outline of the table of contents of your thesis, following this pattern:
  1. Introduction: Societal improvement problem, stakeholders and their goals, current designs, gap with improvement needs.
  2. Research problem: top-level design problem; decomposition into subproblems; **knowledge questions**
  3. State of the art: existing designs
  4. Requirements for a new design; motivation in terms of stakeholder goals; evaluation of current designs against the requirements
  5. New design
  6. **Validation of new design: prototypes, simulations, field experiments, etc**.
  7. (More designs and validations)
  8. Conclusions, recommendations, and further work

Exercise (knowledge-driven thesis): your table of contents

- Make a poster with the outline of the table of contents of your thesis, following this pattern:
  1. Introduction: Societal improvement problem, stakeholders and their goals, current knowledge, gap with desired knowledge.
  2. **Research problem: Top-level knowledge question; decomposition into sub-questions**
  3. State of the knowledge: existing knowledge
  4. **Research methods followed**
  5. **Study: observational study, experimental, case-based, sample-based, etc.**
  6. (More studies)
  7. Conclusions, recommendations, and further work

- Wieringa, R.J. and Daneva, M. (2015) *Six strategies for generalizing software engineering theories.* Science of computer programming, 101. pp. 136-152.

Wieringa, R.J. (2014) *Design science methodology for information systems and software engineering.* Springer Verlag

- Wieringa, R.J. (2014) *Empirical research methods for technology validation: Scaling up to practice.* Journal of systems and software, 95. pp. 19-31.
- 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, 14-15 May 2012, Las Vegas, USA. pp. 220-238. Lecture Notes in Computer Science 7286. Springer.
- Wieringa, R.J. (2010) *Relevance and problem choice in design science.* In: *Global Perspectives on Design Science Research (DESRIST). 5th International Conference*, 4-5 June, 2010, St. Gallen. pp. 61-76. Lecture Notes in Computer Science 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.