Empirical Methods: Experiments and Case Studies with Human Subjects

Claes Wohlin

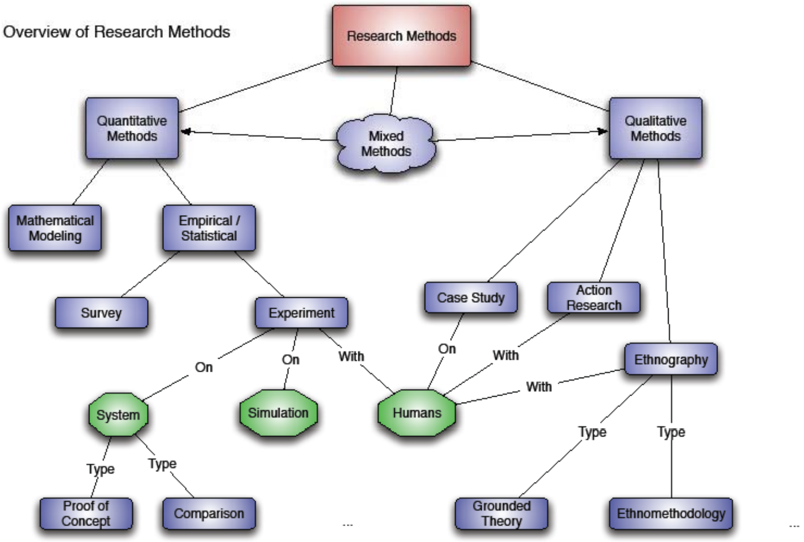
School of Engineering
Blekinge Institute of Technology
Sweden

Today's objective

- An introduction to empirical research methods.
- A brief introduction to case studies.
- A more in-depth introduction to the experimental process for running experiments with human subjects.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Overview of Research Methods



Quantitative methods

- Survey
- Experiment
  - Human subjects (main focus in the lecture)
  - Simulation
  - Real system (proof of concept, comparison)
- Case study
  - Often both qualitative and quantitative

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Why empirical methods?

Evaluation and validation of methods, techniques and tools mean focusing on applying empirical methods, including:

- Controlled experiments
- Case studies
- Surveys

In many cases we perform statistical analysis.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Some noteworthy things

- Proposal by Barbara Kitchenham
  - Research-in-the-typical: case study
  - Research-in-the-large: survey
  - Research-in-the-small: experiment

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

A comparison

- Survey
  - Sample from population
- Case study
  - Often in-depth study of an object, for example development project or an activity
- Experiment
  - Treatment to outcome
  - Cause and effect
  - Although, a sample is still used

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Process/method prototyping

To understand, evaluate and validate a product solution, a prototype may be built. We may do the same for processes/methods using empirical studies.

Especially, experiments using human subjects or simulation may be used for process evaluation.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Empiricism meets engineering 1(3)

- ***Confirmation*** of more or less accepted hypotheses**.** *For example*: object-orientation is good for reuse.
- ***Evaluation*** of methods, models, languages and tools**.** *For example*, whether Java produces higher quality code than C++.
- ***Identification*** of relationships**.** *For example*: find a relationship between fault-prone components and design concepts.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Empiricism meets engineering 2(3)

- ***Validation*** of models or measures. *For example*, to validate a specific cost estimation model.
- ***Understanding*** of methods, techniques and models. *For example*, to understand the relationship between inspections and test.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Empiricism meets engineering 3(3)

- ***Guidance/control*** to help in management. *For example*, as input to planning of personnel to software inspections.
- ***Change/improve*** to support decision-making with respect to changes. *For example*, the result of a study can help us to decide whether or not to introduce a new development tool.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Case Study

"Case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between the phenomenon and context are not clearly evident"

Robert K. Yin, Case Study Research

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Case studies in software engineering

An example: software process improvement

- Involves complex activities
- success or failure depends on many interrelated factors
- cannot be fully studied in isolation
- needs empirical studies in real world settings.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Different types of case studies

Some possible dimensions:

- Snapshot (situation now)
- Longitudinal (over time)
- Pre-post (before and after a change)
- Patchwork (combining related case studies)
- Comparative (cross-unit comparison)

The different types of case studies provide different types of evidence.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Within a case study

A case study may often be conducted as a combination of other methods, for example:

- Archival analysis (documentation)
- Interviews
- Survey
- ...

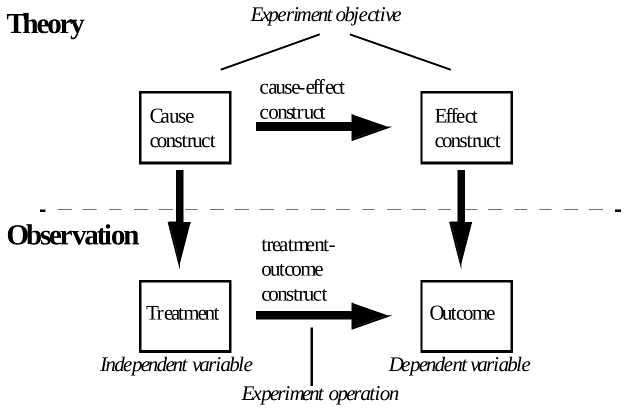
Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Experiment

- Experiments are carefully planned and fully controlled. An experiment should be replicable, i.e. somebody else should be able to repeat it.
  - This type of method will be used to exemplify empirical studies with human subjects.

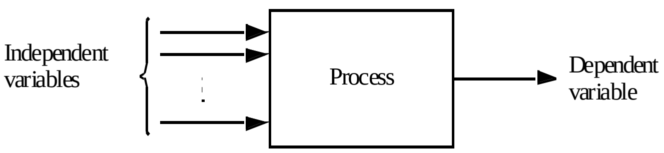
Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Experiment principle



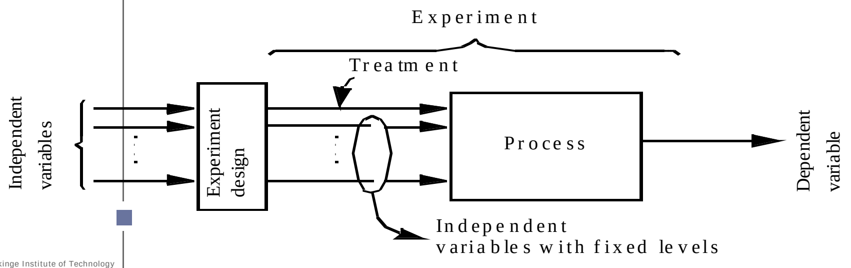
Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Independent and dependent



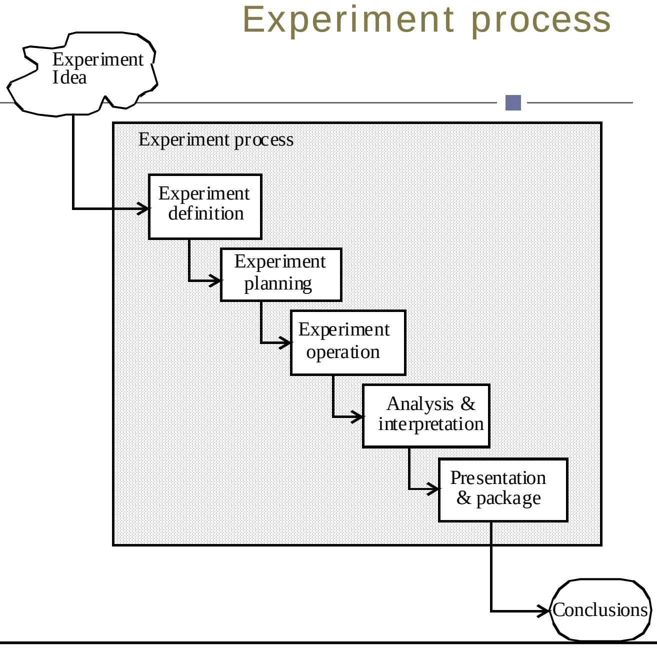
Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Illustration of experiment



Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Experiment process



Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Experiment definition

- The goal template is:

  Analyse *<Object(s) of study>*

  for the purpose of *<Purpose>*

  with respect to their *<Quality focus>*

  from the point of view of the *<Perspective>*

  in the context of *<Context>*.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

An example definition

Analyse *the PBR and checklist techniques*

for the purpose of *evaluation*

with respect to *effectiveness and efficiency*

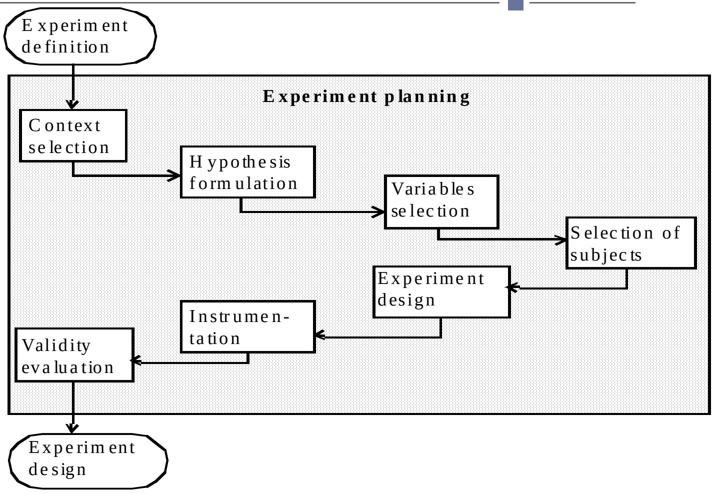
from the point of view of *the researcher*

in the context of *students reading requirements documents*.

PBR – Perspective-based reading

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Experiment planning



Planning phase overview

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Steps in planning 1(4)

- Context:
  - Off-line vs. On-line
  - Students vs. Professionals
  - Toy vs. Real problems
- Hypothesis formulation:
  - Null hypothesis (no real underlying trend or pattern) and alternative hypothesis. The objective is to reject the null hypothesis with as high significance as possible.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Steps in planning 2(4)

- Variables:
  - Independent (input)
  - Dependent (output)
- Subjects
  - Sampling strategy, sampling from population
- Design principles
  - Randomization
  - Blocking (e.g. on experience)
  - Balancing (same number of subjects in groups)

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Steps in planning 3(4)

- Design types

A large number of standard designs exist, and we should select an appropriate design type depending on treatments and number of subjects and of course the objective (hypothesis) of the experiment.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

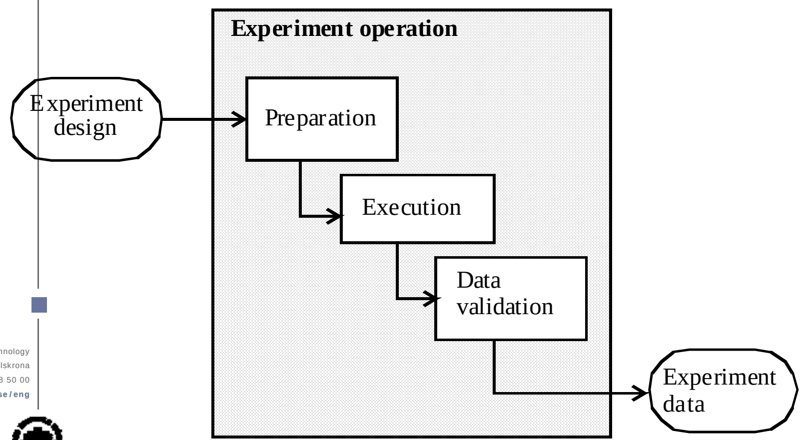
Steps in planning 4(4)

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

- Instrumentation
  - Objects
  - Guidelines
  - Measurement instruments
- Validity evaluation
  - Conclusion validity: treatment to outcome (observation), ability to draw correct conclusions
  - Internal validity: treatment causes outcome, threat to causal relationship
  - Construct validity: relationship between theory and observation
  - External validity: generalization

Experiment operation

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

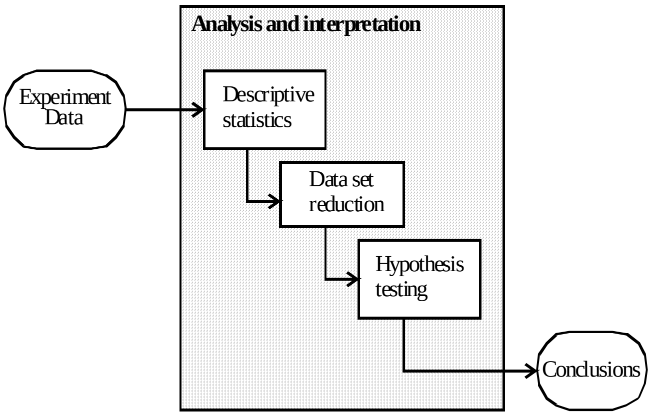


Steps in operation

- Preparation
  - Commit participants
  - Instrumentation (availability)
- Execution
  - Data collection
  - Experimental environment
- Data validation (general check)

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Analysis and interpretation



Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Steps in analysis 1(2)

- Descriptive statistics
  - Scale types (nominal, ordinal, interval and ratio)
  - Measures of central tendency, dispersion and dependency
  - Graphical visualization
- Data set reduction
  - Outliers

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Steps in analysis 2(2)

- Hypothesis testing
  - Parametric tests (assumes a specific distribution, usually a normal distribution)
  - Non-parametric tests (no assumption on distributions)

The different types of tests are related to the standard design types. The intention is to be able to reject the null hypothesis with a statistical significance.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Interpretation

The statistical analysis forms the basis for interpretation.

The interpretation is the foundation for decision-making based on engineering principles.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Packaging

Report outline:

- Introduction
- Problem statement
- Experiment planning
- Experiment operation
- Data analysis
- Interpretation of results
- Discussion and conclusions
- Appendix

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Additional concerns

- Triangulation
- Replication
- Lab packages
- Meta-analysis

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Simplistic example: experiment

- Problem: We want to evaluate reading techniques for inspections.
- We have two competing methods. State a null hypothesis, for example method A and method B show no difference in defect detection for requirements specifications.
- Let people inspect a requirements specification with a known number of defects.
- Use a statistical method to evaluate the hypothesis. Could the null hypothesis be rejected?
- Determine which method is the best.
- Decide whether or not to start using the method.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Some conclusions

- There is a lack of validated results in the field,
- Empirical studies is needed in software engineering to evaluate and validate development process activities,
- Empirical studies mean that the human dimension in software development can be included in the analysis.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Some sources of information

- ISERN: International Software Engineering Research Network (www.iese.fhg.de/ISERN/)
- BESQ: Large research project together with industry at BTH (www.bth.se/besq)
- ESEM: International Symposium on Empirical Software Engineering and Measurement (www.esem-conferences.org). This year in Kaiserslautern, Germany and in USA in 2009.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng

Selected references

- Basili, Selby and Hutchens, "Experimentation in Software Engineering", IEEE Transactions on Software Engineering, Vol. SE-12, No. 7, pp. 733-743, 1986.
- Wohlin et al., "Experimentation in Software Engineering – An Introduction", Kluwer Academic Publishers, 1999.
- Juristo and Moreno, "Basics of Software Engineering Experimentation", Kluwer Academic Publishers, 2001.
- Fenton and Pfleeger, "Software Metrics. A Rigorous and Practical Approach", International Thomson Computer Press, 1996.
- Höst, Wohlin and Thelin, "Experimental Context Classification: Incentives and Experience of Subjects", Proceedings ICSE '05, pp 470-478, 2005.
- Yin, "Case Study Research. Design and Methods" (3rd edition) London: Sage, 2003.

Blekinge Institute of Technology
SE-371 79 Karlskrona
+46 455 38 50 00
www.bth.se/eng