Lecture
Empirical Model Building and Methods
(Empirische Modellbildung und Methoden)

Dr. Andreas Jedlitschka

SS 2017

Chapter 3.2 – Definition

Subchapter objectives

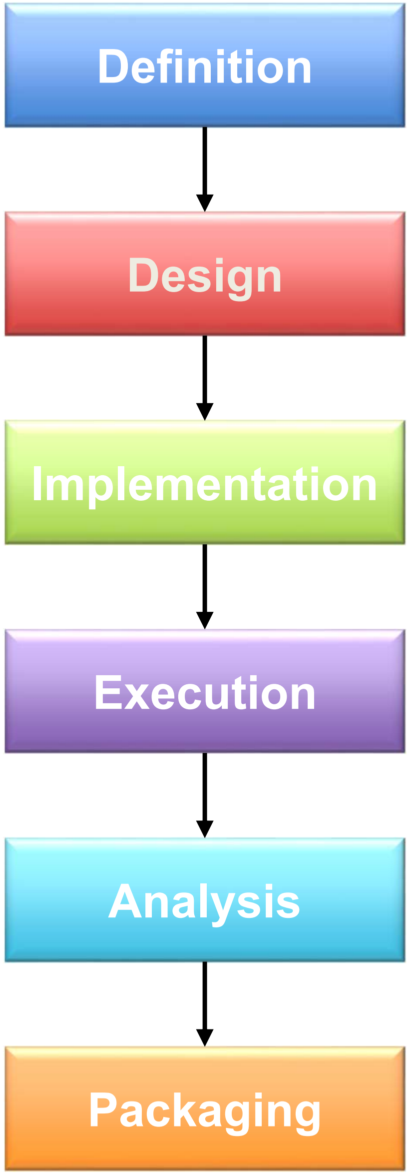
3.1 Introduction
3.2 Definition
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

At the end of this chapter, you should …

- know the purpose of the **definition phase**
- know the tasks to be performed during the **definition** of an empirical study
- understand which decisions should be made during the **definition** of an empirical study and the criteria to be taken into consideration.
- be able to formulate and evaluate research goals, research questions and hypotheses

Empirical process - Definition

3.1 Introduction
Definition
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging



Determine study goal(s) and research hypotheses.

Select type of empirical study to be employed.

Formulate research hypothesis.

Definition

- 3.2.1 Problem Statement
- 3.2.2 Research goals and questions
- 3.2.3 Types of empirical studies
- 3.2.4 Select empirical study method
- 3.2.5 Formulate hypotheses
- 3.2.6 Example

Problem statement

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example

3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

At the end of the definition phase, we need to know:

- Why should the empirical study be conducted (GQM purpose)?
- What will be investigated?
  - Which object(s) to evaluate (GQM object)?
  - Which attribute(s) is (are) to be evaluated (GQM context)?
  - Which effect(s) is (are) expected to be observed (GQM quality focus)?

Steps

- Problem identification
  - How to identify relevant problems?
- Problem selection
  - How to select a problem?

Problem statement

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Problem identification

- Analyze the state of the art
  - How?
    - Systematic literature review
  - What should be identified?
    - Claims, assumptions, investigated or contradictory hypotheses or theories
- Analyze the state of the practice
  - How?
    - Post-mortems, interviews, surveys, observations
  - What should be identified?
    - Problems, needs for improvements

Is there a need / justification for the research?

Details Problem Identification

Problem statement

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Problem selection

- Scientific criteria
  - Is the problem description precise?
  - Can it be investigated empirically?
  - Does its investigation contribute to science?
  - Does its investigation contribute to practice?
- Ethical criteria
  - Does the investigation imply benefits, disadvantages, risks or harms for subjects?
  - Is it possible to collect and keep data anonymous?
  - Does data collection break data protection laws?
  - How will the data be stored and processed?
  - Who will have access to the raw data?
  - How long will the date be stored?
  - Will the results be published?

Now that we have the problem, we want to address …

Research goals and questions

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example

3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Exploring
- Understanding events, decisions, processes, …, and their meaning in specific context based on subjects' insight

Describing
- Drawing accurate descriptions of events, decisions, processes,… , and the relations among them

Precise descriptions are the **foundation** for theories and hypotheses

Explaining
- Testing causal-relationships among events

Confirmed causal-relationships are the foundations for prediction

Evaluating
- Assessing goals achievement w.r.t. normative criteria (e.g., baseline or a standard)

Changing
- Identifying required improvements in events, processes, … for achieving goals

Research goals and questions

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example

3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Research goal definition (according to GQM)

- Analyze
  (object of study: process, product, people)
- with the purpose to
  (purpose: understand, describe, explain, evaluate, change, …)
- with respect to
  (focus: cost, correctness, reliability, usability,...)
- from the point of view of
  (stakeholder: user, customer, manager, developer, corporation,...)
- in the following context
  (context: problem factors, people factors, resource factors, process factors,...)

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example

3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Examples for research questions

- Non-causal research questions
  - What is X? What does X mean?
  - How does X work? Why (not)?
  - How do you select/adopt/use/estimate/…. X?
  - Why does a subject support/select/adopt/use/…. X?

- Correlation
  - What is the correlation between X and Y

- Causal research question
  - Does X cause Y?
  - Does X1 cause more of Y than X2 causes of Y?
  - “Difference” research question
    - What are the differences between X1 and X2?
    - What is the effect of X on Y?
  - “Change” research questions
    - How does Y change (over time) because of X?

(abstracted from M. Miles and A. Huberman, 1994, p.24)

Research goals and questions

Example

PBR Experiment (Ciolkowski et al., 1997)

Research goal

Analyze PBR and Ad-hoc reading techniques for the purpose of their evaluation with respect to their effectiveness from the viewpoint of the researcher in the context of the Software Engineering (SE) lecture at the University of Kaiserslautern (UKL).

Research questions

Q1. Are individuals using PBR more effective than using Ad-hoc reading?	Q2. Are teams using PBR more effective than using Ad-hoc reading ?	Q3. Is PBR more effective in detecting defect classes than Ad-hoc reading?	Q4. Is PBR more effective in detecting differing defects than Ad-hoc reading?

Research subgoals

G1: Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness for individuals ...	**G2:** Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness for teams ...	**G3:** Analyze PBR perspectives for comparing them w.r.t their effectiveness to detect different defect classes ...	**G4:** Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness to detect differing defects

Types of Empirical Studies

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example

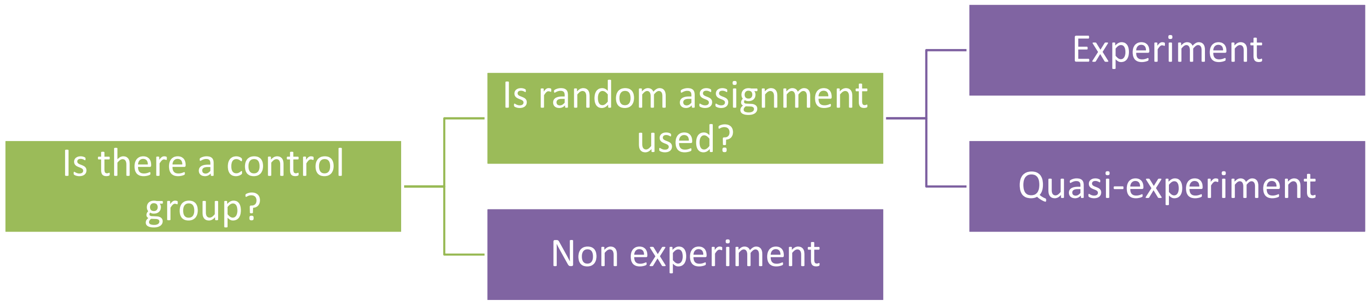
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

These three categories are not meant to be exclusive

Classification criterion 1: Purpose

- explorative, descriptive, and causal
  - In the PBR example we are evaluating different reading techniques with regard top their effectiveness -> difference - > causal

Classification criterion 2: Design



Classification criterion 3: Method

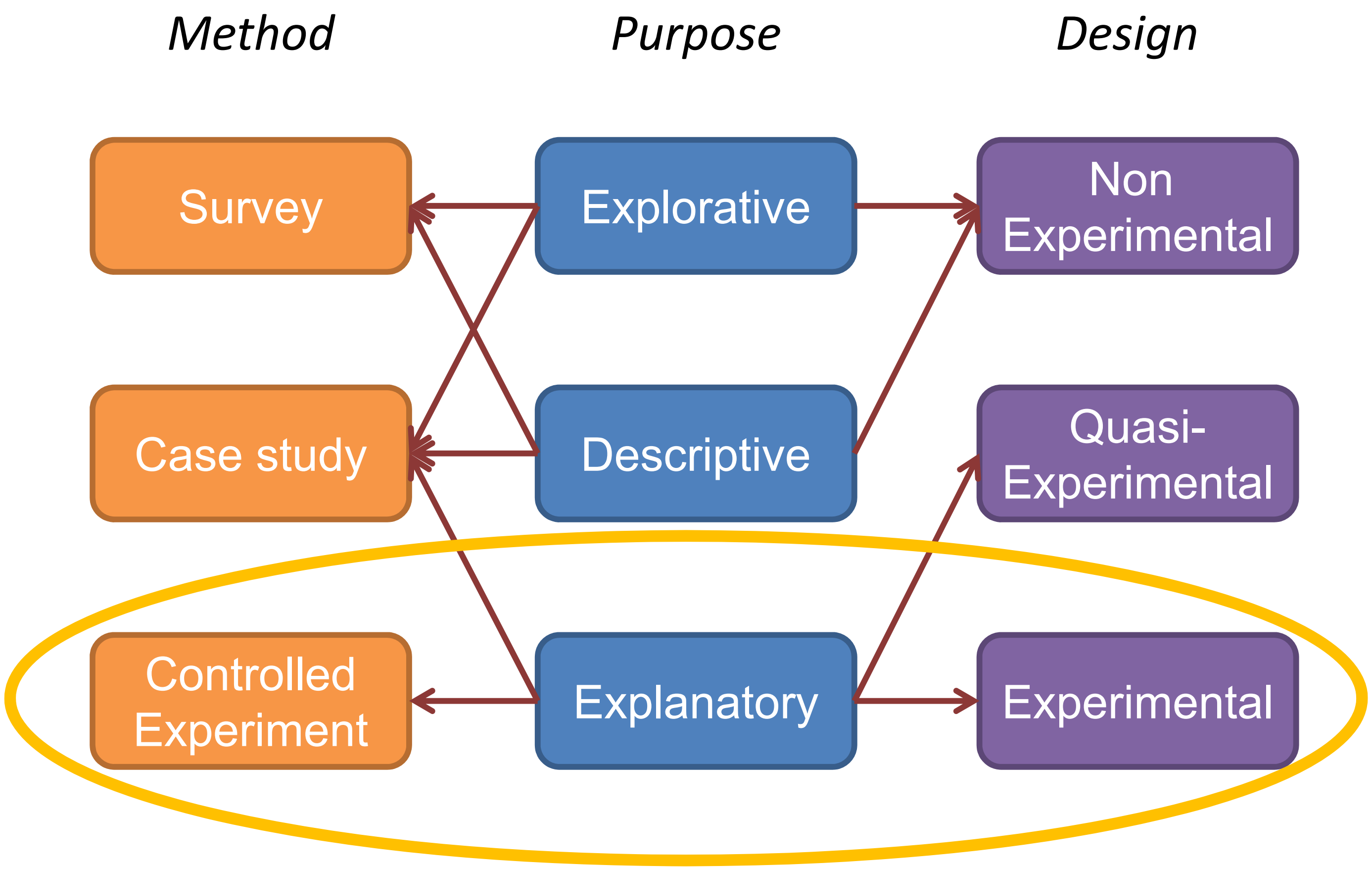
- in SE we usually distinguish: **(controlled) experiment, case study, and survey**

- Other methods include: Action research, observational study, interviews, focus groups, phenomenology, ethnography, …

Criteria for Selecting Empirical Study Types

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example

3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging



Select Empirical Study Method

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example

3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Criteria

Survey | Case study | Controlled Experiment

- Purpose
  - What is known of the problem (State of the art)?
  - What is the need? Exploration, description or explanation!
- Validity of results
  - Internal validity: Is it possible to explain unequivocally the changes in the dependent variable due to the influence of the independent variables?
  - External validity: Is it possible to generalize the results of the sample to other populations or situations?
- Degree of control
  - Can we 'arrange' the real world?
  - Can we randomize the assignment of subjects to treatments?
- Cost
  - To which extend is it possible to fund the study?
- Risk
  - How likely is it that the study might fail?
  - Which are the consequences?

Select Empirical Study Method

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Controlled Experiment | Experimental | Quasi-Experimental

When to select a Controlled Experiment?

- **When appropriate**: (1) Experimental and control groups can be defined; (2) Control on who is using which technology, when, where and under which conditions. (3) Investigation of self-standing tasks where results can be obtained immediately.
- **Validity**: High internal validity, limited external validity.
- **Level of control**: High.
- **Randomization**: Random assignment of treatments to subjects.
- **Data analysis**: Statistical analysis, e.g. testing causal relationships or comparing central tendencies of treatments/groups.
- **Pro’s**: High internal validity, i.e. it is possible to establish statistically significant causal relationships and to confirm theories. Usually, low execution costs and risks.
- **Con’s**: Limited external validity, i.e. it is difficult to generalize result to other contexts due to limited representativeness. Application in industrial context requires additional efforts and compromises. Usually, high design costs.

Select Empirical Study Method

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Case study | Non Experimental

When to select a Case Study?

- **When appropriate**: change (new technology) is wide-ranging throughout the development process, want to assess a change in a typical situation, understand a problem or current situation.
- **Level of control**: Low
- **Validity**: Limited internal validity, external validity often considered high in comparison to an experiment
- **Data analysis:** Statistical analysis, e.g. comparing (1) results to a baseline / standard (sister project, company baseline), (2) results over time or (3) multi-case study results.
- **Pro's:** High external validity, i.e. results are representative if real circumstances/projects/…. Provide (qualitative) insights regarding why and how questions. Eventually, low design costs.
- **Con's:** Low internal validity, i.e. it is difficult to determine causal relationships because of cofounding variables. Results are context specific. Usually, high execution costs and risks.

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example

3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Select Empirical Study Method

Survey | Non Experimental

When to select a Survey?

- **When appropriate:** Early exploratory and descriptive studies, e.g. technology change implemented across a large number of projects, description of results, influence factors, differences and commonalities
- **Level of control:** Low
- **Validity:** Internal validity: It depends e.g. on instrument reliability. External validity: It depends e.g. on sample representativeness (and response rate)
- **Data analysis**: Statistical analysis regarding frequency distributions, e.g. describing or comparing knowledge, attitudes, behavior, decisions, preferences, … (over time). Examine association and trends, consistency of scores
- **Pro's:** If the sample is representative and the survey is valid, then the generalization of results is usually easier than case study. It is easy to access to a large sample. Applicable in practice. Usually, low execution costs and risks.
- **Con's:** Little control of cofounding factors. High design costs.

Types of Empirical Studies: Design

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Design	Key Aspects of the design
Experimental	• Researchers manipulate systematically potential causal variables and then seek to observe or measure the effects. • It involves: - Comparison of (at least) 1 experimental and 1 control group - Random assignment of subjects to treatments - Control of independent and cofounding variables
Quasi-Experimental	• Experimental design in which pre-existing groups are studied (not randomized)
Non Experimental	• It aims at studying a phenomenon as it exists at the time of the study • No use of control or comparison groups • Purposeful sampling (not randomized)

Formulate Hypotheses

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

A hypothesis (from Greek ὑπόθεσις; plural hypotheses) is a proposed explanation for a phenomenon (Wikipedia, May 2012)

Classification

- Descriptive Hypothesis
  - contains only one variable thus it is also called univariate hypothesis
- Relational Hypothesis
  - Correlation
    - Proposes a relationship between the variables (do not indicate cause and effect)
    - As X increases/decreases, Y increases/decreases
- Causal (Explanatory) Hypotheses
    - Proposes a cause-effect relationship
  - Difference Hypothesis
    - X is more effective than Y with regard to Z
  - Change Hypothesis
    - “Reading a certain newsletter regularly changes the political perspective of the reader”

Formulate Hypotheses

3.1 Introduction
Definition
3.2.1 Problem Statement
3.2.2 Research goals and questions
3.2.3 Types of empirical studies
3.2.4 Select empirical study type
3.2.5 Formulate hypotheses
3.2.6 Example
3.3 Design
3.4 Implementation
3.5 Execution
3.6 Data analysis
3.7 Packaging

Classification

- Non directional (two-tailed hypothesis)
  - It predicts a difference/relationship between two variables, but not the direction or the nature of their relationship….
  - e.g., Effectiveness(PBR) <> effectiveness(CBR)
- Directional (one-tailed hypothesis)
  - It predicts the
    - direction of the difference between two variables
    - nature of the corresponding relationship, i.e. a positive or negative correlation.
  - It can be specific or unspecific, i.e., the difference is quantified
  - Requires previously obtained knowledge about the effect
    - E.g., theory
  - e.g., EaseOfUse(structDocs) > EaseOfUse(unstructDocs)

Research goals and questions

Example (repeated)

PBR Experiment (Ciolkowski et al., 1997)

Research goal

Analyze PBR and Ad-hoc reading techniques for the purpose of their evaluation with respect to their effectiveness from the viewpoint of the researcher in the context of the Software Engineering (SE) lecture at the University of Kaiserslautern (UKL).

Research questions

Q1. Are individuals using PBR more effective than using Ad-hoc reading?	Q2. Are teams using PBR more effective than using Ad-hoc reading ?	Q3. Is PBR more effective in detecting defect classes than Ad-hoc reading?	Q4. Is PBR more effective in detecting differing defects than Ad-hoc reading?

Research subgoals

G1: Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness for individuals ...	**G2:** Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness for teams ...	**G3:** Analyze PBR perspectives for comparing them w.r.t their effectiveness to detect different defect classes ...	**G4:** Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness to detect differing defects

Formulate Hypotheses: Example

Example

PBR Experiment (Ciolkowski et al., 1997)

Research subgoals

G1: Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness for individuals ...	**G2:** Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness for teams ...	**G3:** Analyze PBR perspectives for comparing them w.r.t their effectiveness to detect different defect classes ...	**G4:** Analyze PBR and Ad-hoc reading techniques for comparing them w.r.t their effectiveness to detect differing defects

Research hypotheses

H_1: Individuals using PBR are more effective than individuals using Ad-hoc reading techniques.	**H_2:** Teams using PBR are more effective than individuals using Ad-hoc reading techniques.	**H_3:** PBR is more effective in detecting different defect classes than Ad-hoc reading techniques	**H_4:** PBR is more effective in detecting differing defects than Ad-hoc reading techniques