DIT590 Research Methods & Technical Writing

Lecture 5: Software Experiments

Richard Berntsson Svensson
http://www.rbsv.eu/courses/rmtw

Re-cap

- Surveys
  - Qualitative (open-ended) or quantitative (closed)
    - Questionnaires
    - Interviews
- Two basic types of surveys
  - Cross-sectional
  - Longitudinal

Experiment

- True experiments mean that the researcher controls all of the factors and conditions

- With an experiment the researcher is trying to learn something new about the world, an explanation of ‘why’ something happens

Experiment

”The basic intent of an experiment is to test the impact of a treatment (or an intervention) on an outcome, controlling for all other factors that might influence that outcome” (Creswell)

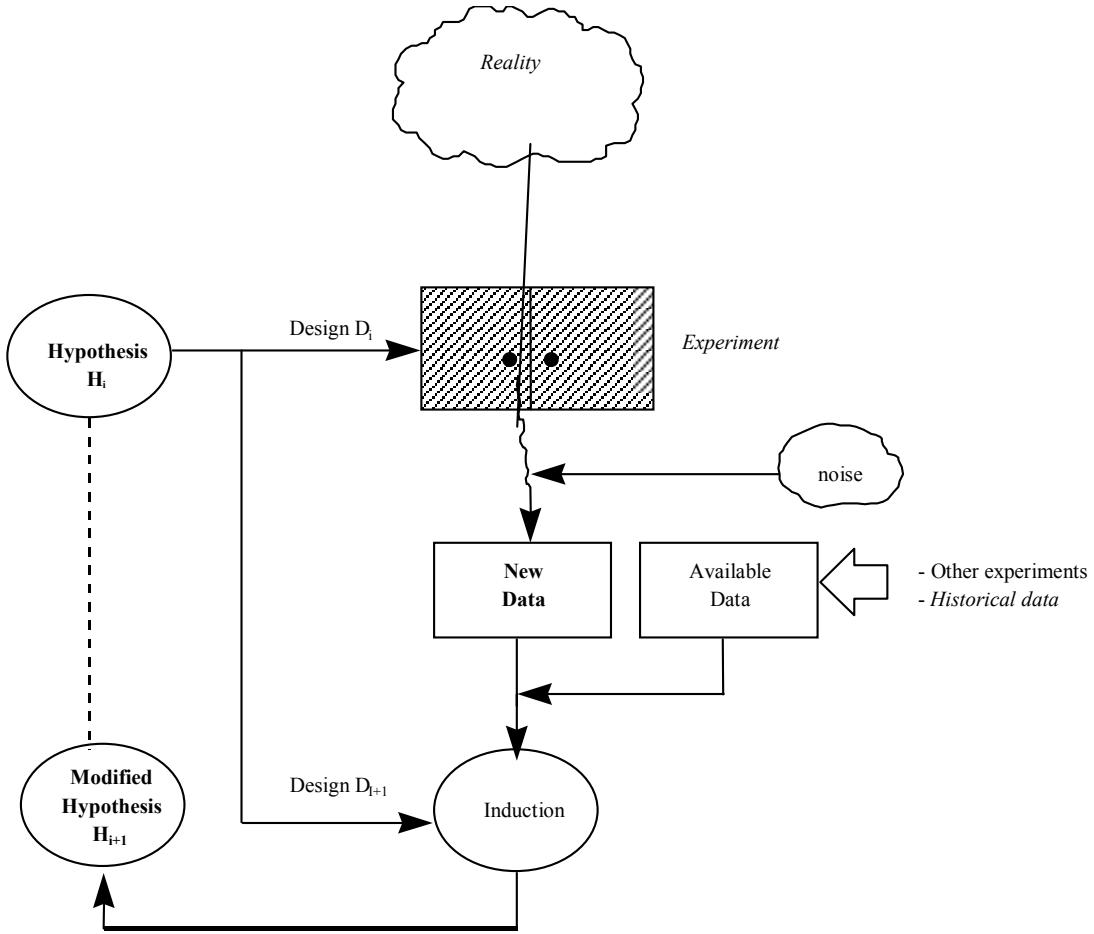
Building knowledge through Experiments

- Experimentation in Software Engineering is necessary
- *"Progress comes when what is actually true can be separated from what is only believed to be true"* (Basili & Lanubile, 1999)
- Allows us to conduct well-defined, focused studies, with the potential for statistically significant results

Experiments in Software Engineering

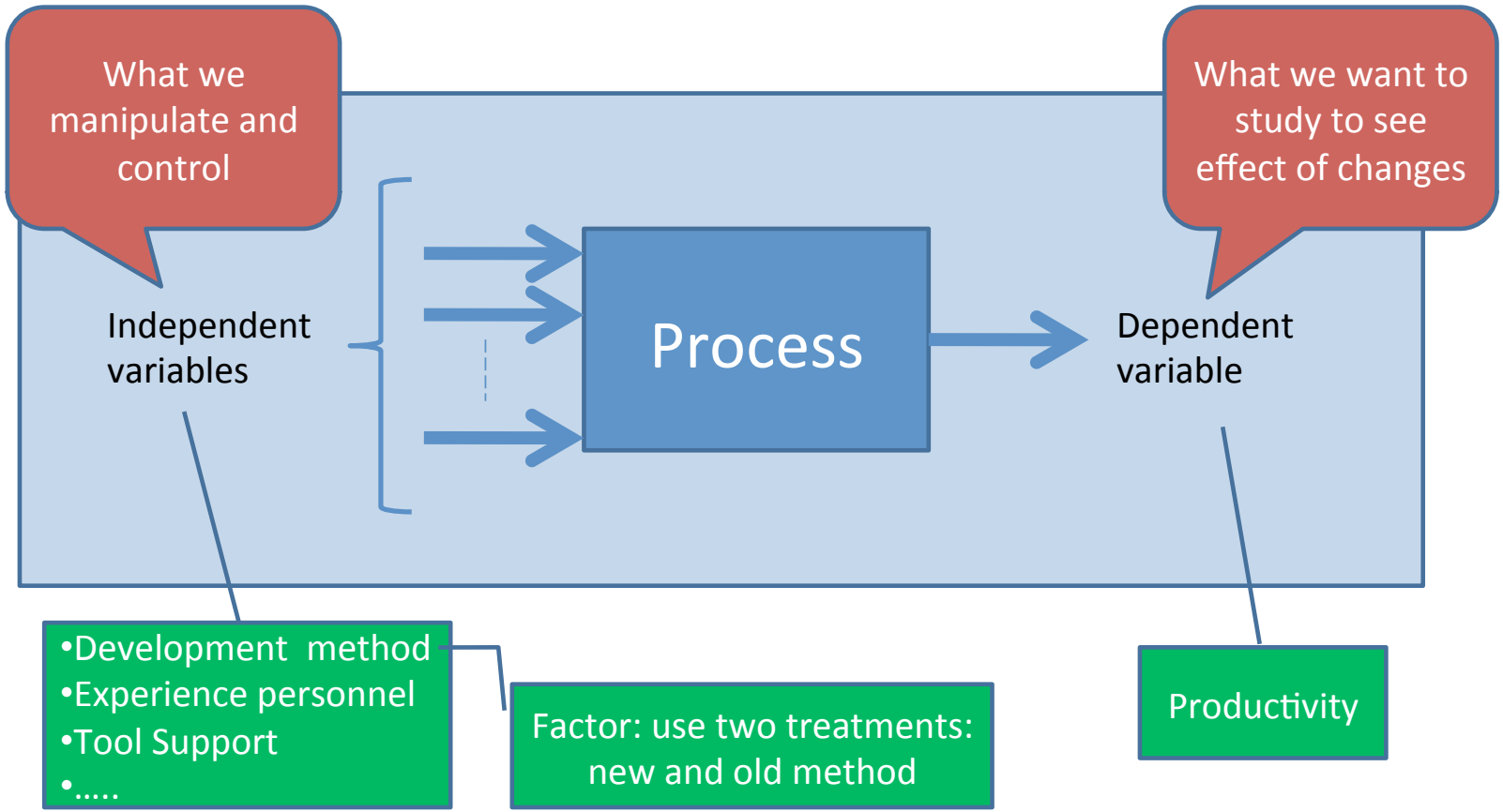
- Often limited by the situation
- Goals vary:
  - understand the effects of processes and techniques to evaluating the product characteristics;
  - predicting costs to trading off environmental constraints
- When introducing any form of process, method or tool the organization needs to evaluate its feasibility, usability and effectiveness in context
- Helps to build knowledge about the discipline as well as help the practitioners understand how to build SW better.

Experimentation/learning cycle



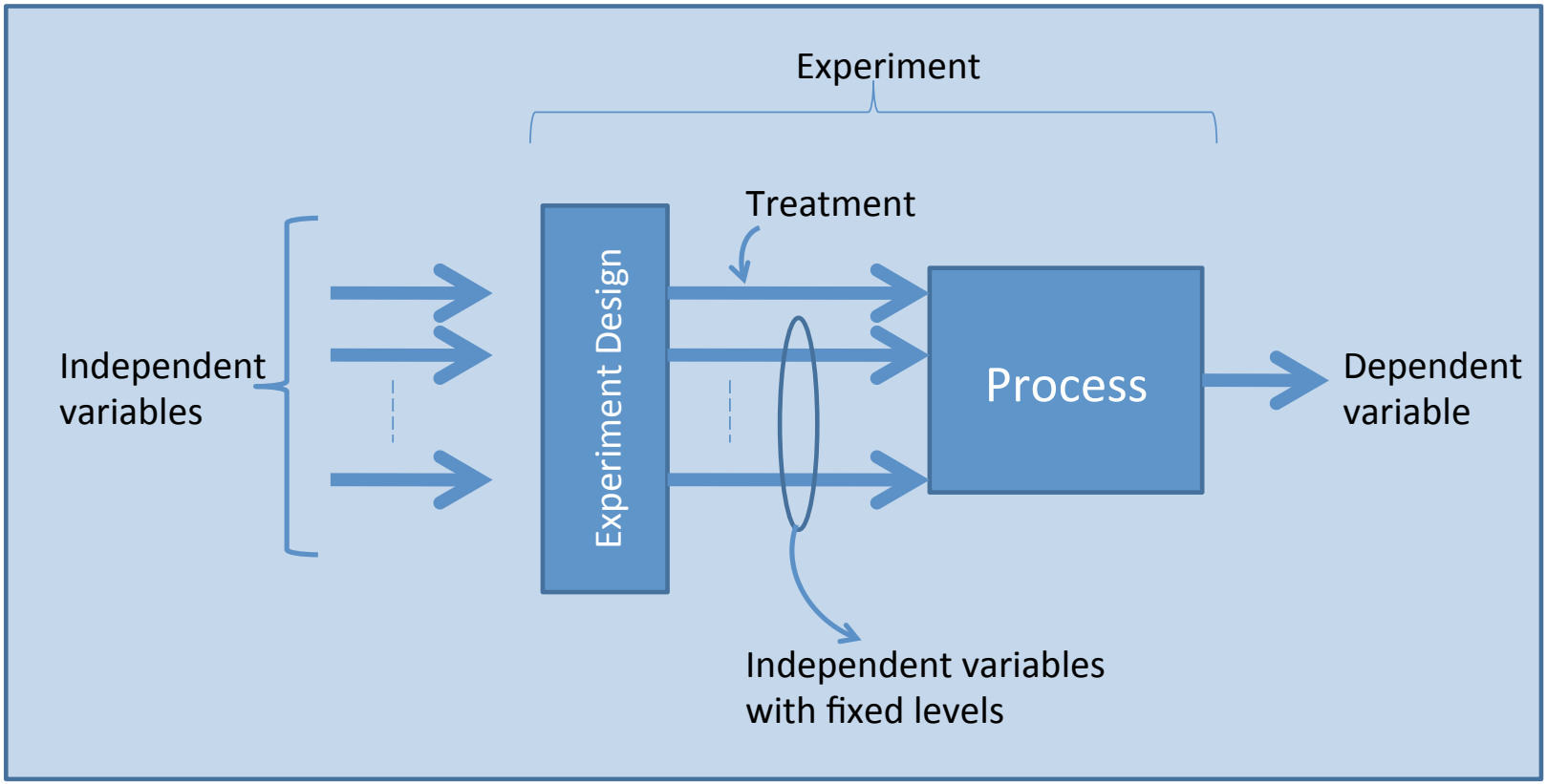
[Expe]

Independent & Dependent Variables

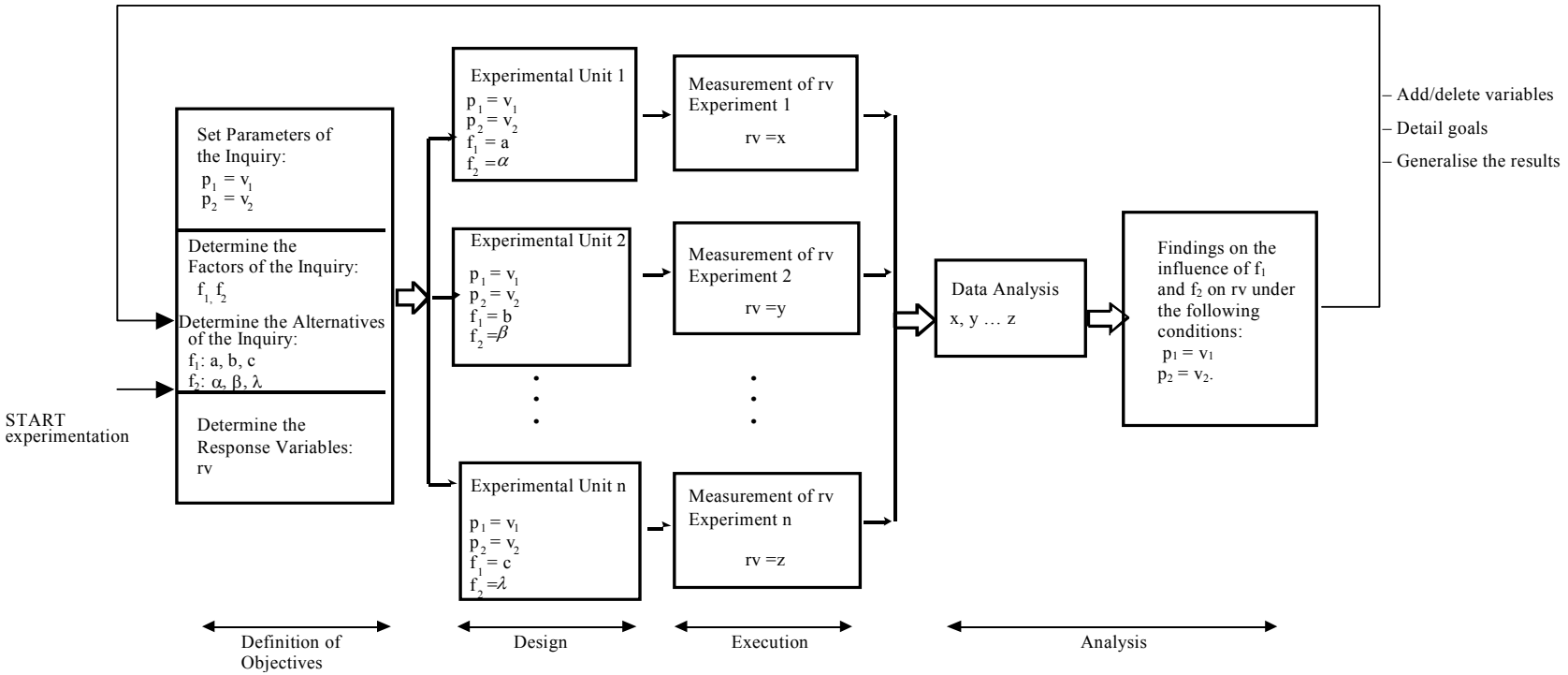


C. Wohlin, P. Runeson, M. Höst, C. Ohlson, B. Regnell, and A. Wessle´n, Experimentation in Software Engineering: An Introduction. Kluwer Academic, 2000.

Experiment Illustration



C. Wohlin, P. Runeson, M. Höst, C. Ohlson, B. Regnell, and A. Wessle´n, Experimentation in Software Engineering: An Introduction. Kluwer Academic, 2000.


START experimentation
Set Parameters of the Inquiry:
$p_1 = v_1$
$p_2 = v_2$
Determine the Factors of the Inquiry:
f_1, f_2
Determine the Alternatives of the Inquiry:
f_1: a, b, c
f_2: α, β, λ
Determine the Response Variables:
rv
Experimental Unit 1
$p_1 = v_1$
$p_2 = v_2$
$f_1 = a$
$f_2 = \alpha$
Experimental Unit 2
$p_1 = v_1$
$p_2 = v_2$
$f_1 = b$
$f_2 = \beta$
Experimental Unit n
$p_1 = v_1$
$p_2 = v_2$
$f_1 = c$
$f_2 = \lambda$
Measurement of rv Experiment 1
rv =x
Measurement of rv Experiment 2
rv =y
Measurement of rv Experiment n
rv =z
Data Analysis
x, y … z
Findings on the influence of f_1 and f_2 on rv under the following conditions:
$p_1 = v_1$
$p_2 = v_2$.
– Add/delete variables
– Detail goals
– Generalise the results
Definition of Objectives
Design
Execution
Analysis

GOAL	FACTORS	PARAMETERS	REFERENCE
Studying the effect of different testing techniques on the effectiveness of the testing process	• Software testing techniques (code reading, functional testing, structured testing) • Program types: three different programs • Subject level of expertise (advanced, intermediate, junior)	• Testing process (first training, then three testing sessions and then a follow-up session) • Program size • Familiarity of subjects with editors, terminal machines and programs implementation language (good familiarity) • High-level language for implementing programs	(Basili, 1987)
Studying the effect of testing techniques on the effectiveness and efficiency at revealing failures	• Inspection technique (code reading, functional testing, structured testing) • Program types (three different programs) • Subjects (six groups of similar subjects) • Order of applying techniques	• Order in which subjects inspect programs (first program 1, then program 2, and then program 3). • Implementation language (C) • Problem complexity (low) • Subjects from a university lab course	(Kamsties, 1995)

Examples of response variables

Development process	Schedule deviation, budget deviation, process compliance
Methods	Efficiency, usability, adaptability
Resources	Productivity
Products	Reliability, portability, usability of the final product, maintainability, design correctness, level of code coverage

[Expe]

Pre-Experimental designs

- Does not include a control group
- Either a single group or multiple groups are observed subsequent to some agent or treatment presumed to cause change.
- Example: The effect of work experience (during one semester) on university grade.
- Low validity (trustworthiness)

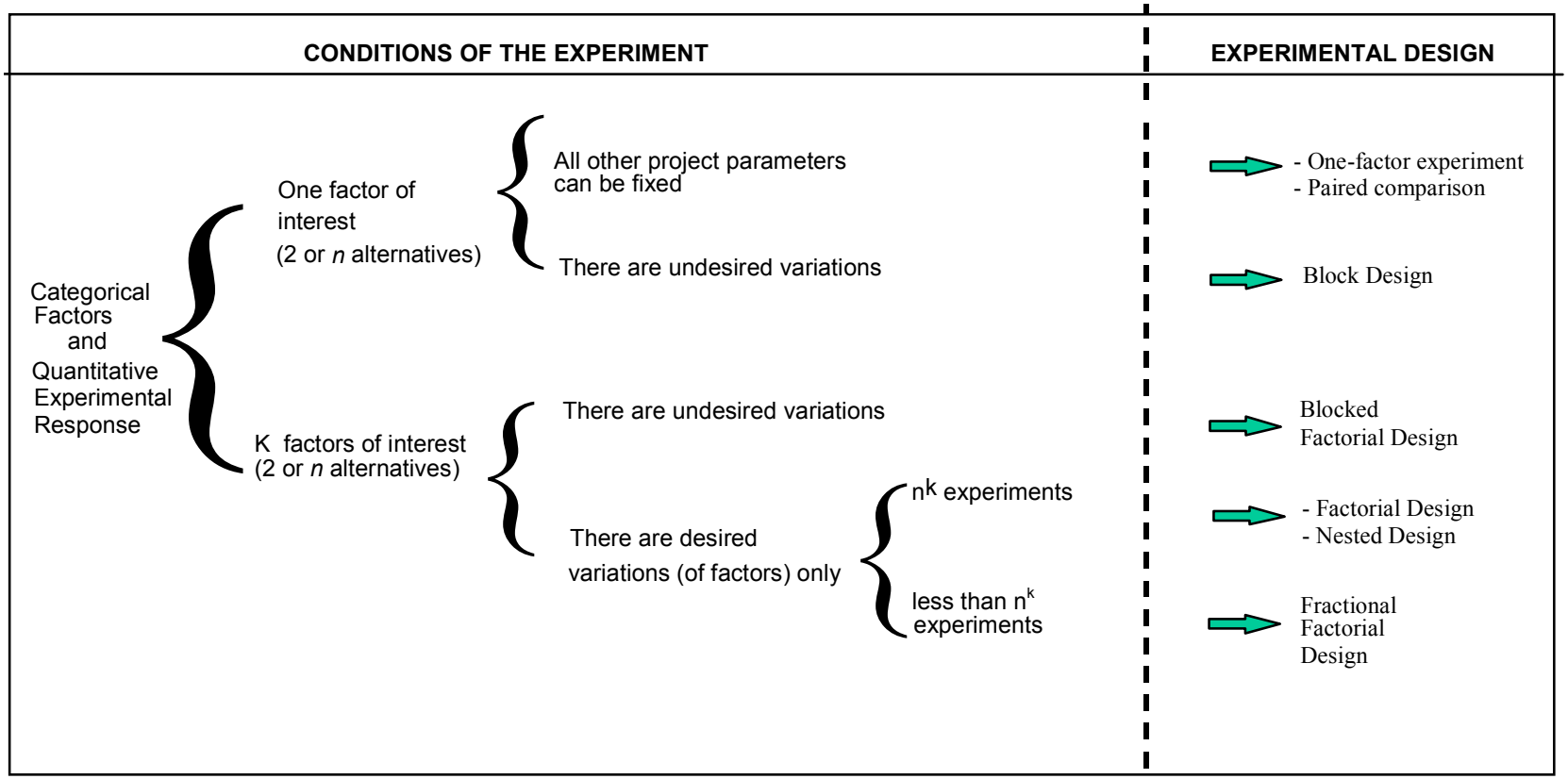
Quasi-Experimental designs

- Quasi-Experimental designs *(field experiments)*. This type of experimental design does include a control group, but the design does not include randomization.
- Can be used to reinforce findings in case studies.
- Example: Comparing two different software development methodologies on perceived product quality.
- High external validity

True Experimental designs

- When individuals can be randomly assigned to groups.
- Less threat to internal validity.
- Example: Laboratory experiment to study behaviors, efficiency etc. *'Does exposure to classical music improve memory'?.*
- High internal validity; low external validity.

Different experimental designs



Threats to validity

- Internal validity
  - Concerns the cause effect relationship, that is, if the measured effect is due to changes caused by the researcher or due to some other unknown cause.
- External validity
  - The possibility to generalize the results
- Conclusion validity
  - Ability to draw statistically correct conclusions based on the measurements
- Construct validity
  - Ability to measure what we are interested in measuring

C. Wohlin, P. Runeson, M. Höst, C. Ohlson, B. Regnell, and A. Wessle´n, Experimentation in Software Engineering: An Introduction. Kluwer Academic, 2000.

Experiment Phases

- Definition of the objectives of the experimentation
  - Hypothesis, variables
- Design of the experiments
  - Design, conditions, determine affected variables
- Execution of the experiment
  - Run as indicated by the selected design
- Analysis of the results/data collected from the experiments
  - Analyze the data

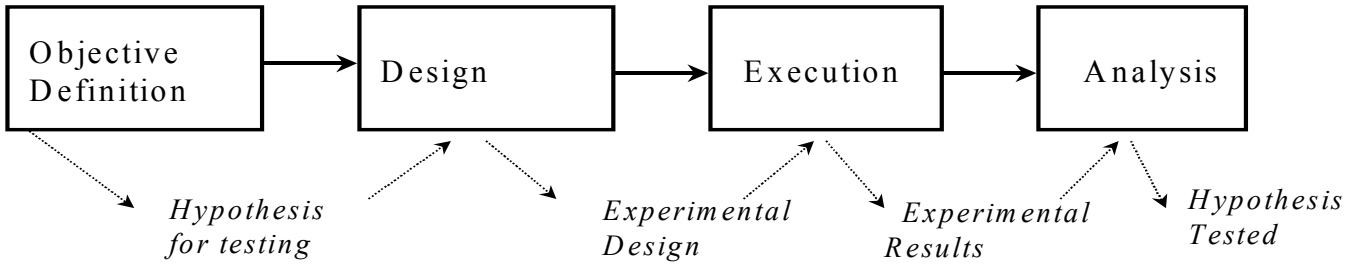


Figure 3.1. Process of experimentation in SE

[Expe]

Experiment Components

- Participants, Selection process
  - Population, random/convenience
- Variables
  - Independent variables (identify all)
  - Treatment variable
  - Dependent variable(s)
- Instruments
  - Observations and measures
  - Materials
- Procedure
  - Type of experiment (pre-. True, quasi, single-subject)
  - Identify what is being compared
  - Design, observations, treatment, timeline of activities
- Validity threats
  - Internal, external, conclusion, construct

Design and Steps

- 1. Identify the factors
- 2. Identify the response variables
- 3. Identify the parameters
- 4. Identify the blocking variables
- 5. Determine the number of replications
- 6. Select the kind of experimental design
- 7. Select the experimental objects
- 8. Select experimental subjects
- 9. Identify the data collection process

Example

- Does listening to music while driving lead to stressful driving behavior?

- **What are the variables?**
  - ‘Music’ (independent) and ‘driving behavior’ (dependent)

- **Who to study?**
  - Experiment group, receives treatment (music)
  - Control group; receives no treatment

Example

- The most important factors to control
  - Individual: control through randomization
  - Contextual: control by conducting experiment in a lab

- Worth conducting a pilot to test the experimental design

- **Experimental unit:** The objects on which the experiment is run are called experimental units or experimental objects.

- **Experimental subjects**: The person who applies the methods or techniques to the experimental units is called experimental subject.

- **Response variable**. The outcome of an experiment is referred to as a response variable.

- **Parameters**. Any characteristic (qualitative or quantitative) of the software project that is to be invariable throughout the experimentation will be called parameter.

- **Factors**. Each software development characteristic to be studied that affects the response variable is called a factor.

- **Levels**. The possible values of the factors during each elementary experiment are called levels.

[Expe]

To do …

- Work on Assignments 1 and 2 (if you didn't attend L4 and L5)
  - Assignment 1 is due February 9, 8.00am
- Theory: Read [Case], [DSIS], [AR], [EMSE]
- Attend Seminar 1 on February 11
- Attend lecture about CS/DS/AR on February 16