



Experiment Planning

Eduardo Figueiredo

<http://www.dcc.ufmg.br/~figueiredo>

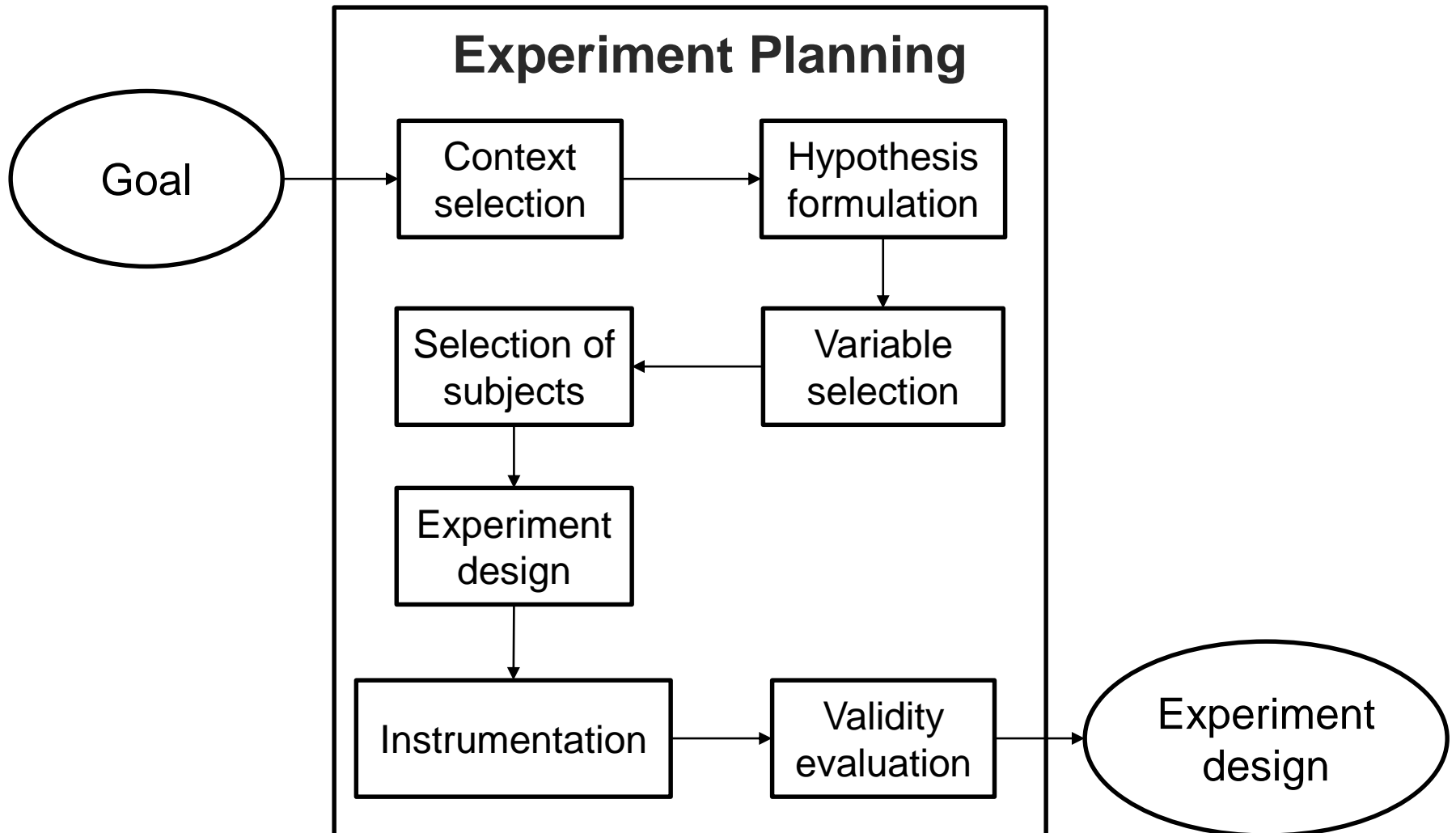
[Scoping and Planning]


- Scoping determines the foundations for the experiment
 - Why the experiment is conducted?
- Planning prepares the experiment
 - How the experiment is conducted?

[Planning Steps]

- The planning phase can be divided into seven steps
 - Context selection
 - Hypothesis formulation
 - Variable selection
 - Selection of subjects
 - Experiment design
 - Instrumentation
 - Validity evaluation

Planning Phase Overview





Context Selection

Context Selection

- The context of the experiment can be characterized by four dimensions
 - Offline vs. online
 - Student vs. professional
 - Toy vs. real problems
 - Specific vs. general
- Cheaper: offline with students using a toy problem in a specific domain



Hypothesis Formulation

[Hypothesis Formulation]

- Hypothesis formulation is the basis for the statistical analysis of an experiment
- If the hypothesis can be rejected, conclusions can be drawn
- We aim to reject the null hypothesis
 - A hypothesis is stated formally
 - Data are collected

[Null and Alternative Hypotheses]


- The experiment is formalized into two hypotheses
 - Null hypothesis (H_0)
 - Alternative hypothesis (H_1)
- An alternative hypothesis is in favor of the null hypothesis rejection

[Null Hypothesis]

- A null hypothesis (H_0) often states that there are no real underlying trends or patterns in the experiment setting
 - Therefore, the only reason for differences are coincidental
- The experimenter wants to reject this hypothesis

[Types of Errors]


- Hypothesis testing involves two main types of risks (or types of errors)
 - Type-I-error: the risk of rejecting a true hypothesis
 - Type-II-error: the risk of not rejecting a false hypothesis



Variable Selection

[Variable Selection]

- The independent variables we can control and change in the experiment
 - The choice of independent variables also includes choosing the measurement scales
- The effect of treatment is measured in the dependent variable
 - There is usually only one dependent variable
 - The dependent variable is derived from the hypothesis



Selection of Subjects

[Selection of Subjects]

- The selection of subjects is related to the generalization of the results
 - In order to generalize, the selection must be representative for the population
- The selection (or sampling) can be of two types
 - A probability sample
 - A non-probability sample

[Types of Sampling]

- Probability Sampling: the probability of selecting each subject is known
 - Example: subjects are selected from a list of the population at random
- Non-Probability Sampling: the probability of selecting each subject is unknown
 - Example: the nearest and most convenient persons are selected as subjects



Experiment Design

[Experiment Design]

- To get the most out of the experiment, it must be carefully planned and designed
 - Design is closely related to interpretation
- An experiment consists of a series of tests of the treatments
 - A design describes how the tests are organized and run

[General Design Principles]

- A proper design allows replication
- The general design principles are
 - Randomization
 - Blocking
 - Balancing

[Randomization]

- All statistical methods require observations from independent random variables
- Randomization applies to the allocation of objects, subjects, and the order of the tests
 - It is also used to select subjects that are representative of the population

[Blocking]

- Blocking is used to systematically eliminate the undesired effect in the comparison among treatments
 - Within one block (group), the undesired effect is the same
- Example:
 - Subjects with different experience in OOP
 - They are grouped into two groups: one of people with and one without experience

[Balancing]

- We have a balanced design if each treatment has equal number of subjects
 - That is, the same number of subjects in each group (block)



Instrumentation

[Instrumentation]

- The goal of instrumentation is to provide means for performing and monitoring the experiment
- The instruments for an experiment are of three types
 - Objects
 - Guidelines
 - Measurements

[Experiment Objects]

- When planning an experiment, it is important to choose appropriate objects
- Example
 - In an experiment for bad smell detection, the number of bad smells must be known
 - You can either introduce bad smells or use objects with known bad smells

[Guidelines]

- Guidelines are needed to guide the participants in the experiment
 - Participants may also need training in the methods or tools to be used
- Guidelines include, for instance, process descriptions and checklists

[Measurement]

- Measurements in an experiment are conducted by data collection
 - Data is usually collected via manual forms and interviews
 - Automated data collection tends to be more reliable than manual collection

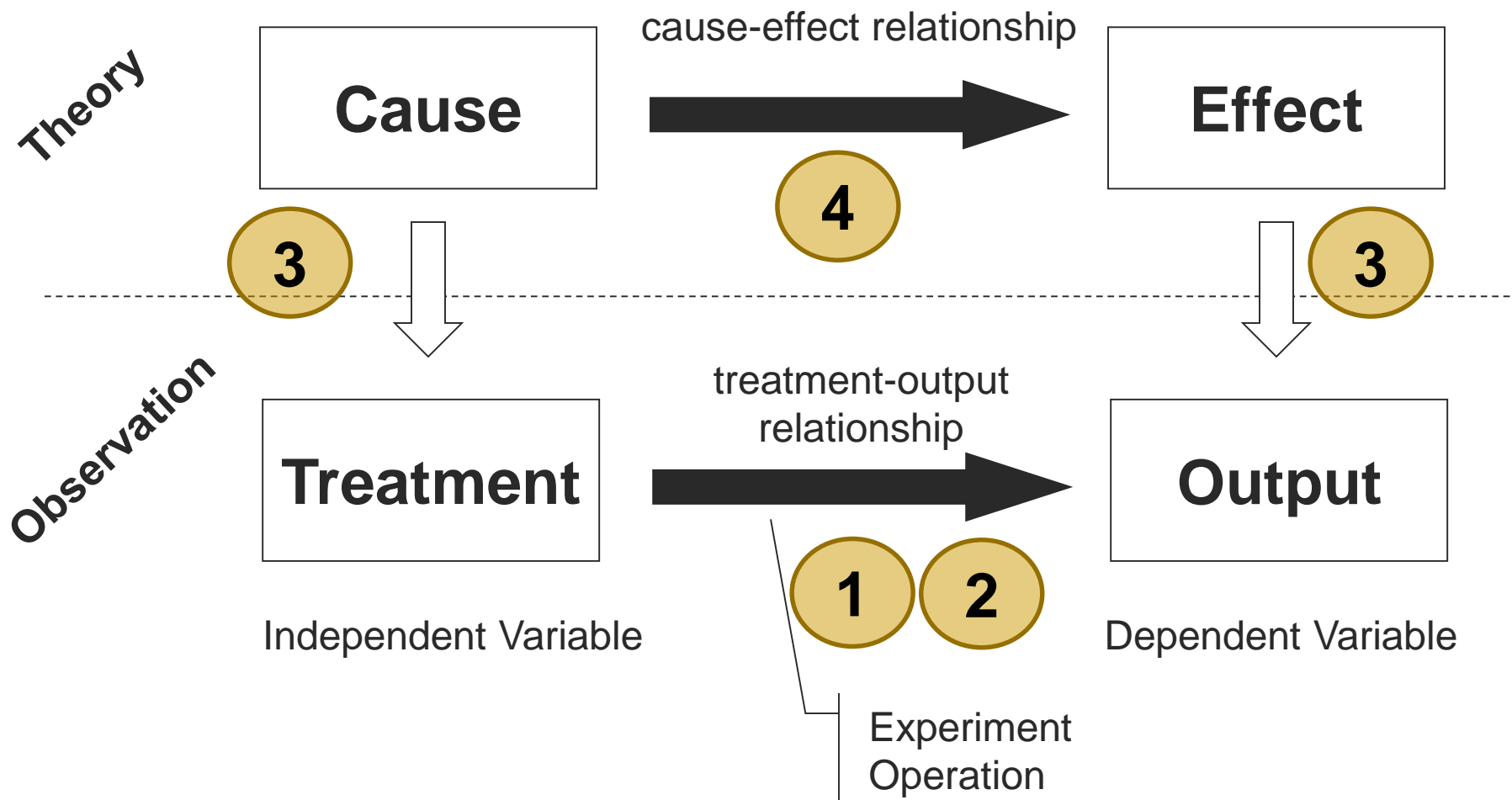


Validity Evaluation

[Validity Evaluation]

- It is important to consider the experiment validity already in the planning phase
- The results must be valid for the population of interest
 - If the population is one organization, results have to be valid for this organization
- The four main types of validity threats
 - Conclusion, Internal, Construct, External

[Experiment Principles]



[Four Categories of Threats]

1. Conclusion Validity: relationship between treatment and outcome is explainable
2. Internal Validity: there is a causal relationship between treatment and outcome
3. Construct Validity: concerned with the relationship between theory and observation
4. External Validity: there is a causal relationship between cause and effect (i.e., theory holds)

[Bibliography]

- C. Wohlin et al. **Experimentation in Software Engineering**, Springer. 2012.
 - Chapter 8 – Planning