DCC / ICEx / UFMG

Empirical Strategies

Eduardo Figueiredo

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

Types of Research

- We can identify two main types of research for empirical studies
  - Exploratory Research
  - Explanatory Research
- The two types are complementary rather than competitive

Exploratory Research

- It is concerned with studying objects in their natural settings
  - Letting the findings emerge from observations
- It has a flexible research design
- It is usually supported by qualitative data

Explanatory Research

- It is mainly concerned with quantifying a relationship or comparing groups
  - The aim often is to identify a cause-effect relationship
- It is usually conducted through a controlled experiment (fixed design)
  - Supported by quantitative data
  - Promotes comparison and statistical analysis

Quantitative vs. Qualitative

- A quantitative investigation can answer
  - How much a new inspection method decreases the number of faults found in tests?
- A qualitative investigation can answer
  - What are the sources of faults between different inspection groups?

Empirical Strategies

- The main empirical strategies are
  - Survey
  - Case Study
  - Experiment / Quasi-Experiment
- Empirical strategies are not orthogonal
  - Some studies may be viewed as a combination of strategies

Survey

Survey

- A survey is an empirical strategy for collecting information from people
  - Used to describe, compare, or explain their knowledge, attitudes, and behavior
- Data is collected by interviews or questionnaires
- Survey is often performed in retrospect
  - Data is analyzed to derive conclusions

A Survey Questionnaire

- In a survey, a questionnaire is answered by a sample of developers
- Collected information can support both quantitative and qualitative analyses
- Be careful: surveys with many questions are tedious for respondents

Generalization

- The purpose of a survey is to understand the population
  - Example: by interviewing 25 developers, we aim to know the opinion of a population
- Conclusions in a survey can often be generalized

Examples of Survey

- Example 1:
  - A tool has been used for a while
  - A survey is conducted to assess its advantages and drawbacks
- Example 2
  - A pool is used to determine how a population will vote in the next election

Case Studies

Case Study

- Case Study is an empirical strategy that draws on multiple sources of evidences
  - It relies on one instance (or small set of instances) within its real-life context
- It normally aims at tracking a specific attribute or at establishing relationships between attributes

Case Study Arrangements

- A case study can be applied as a comparative research strategy
  - A comparison of results using a new method against historic data
  - A sister project: one using a new method and another using the typical method
  - Apply to some components of a larger project and compare results with other components

Confounding Factor

- It is hard to distinguish the effects of two factors from each other in case studies
  - Example: A better result may be either due to a new tool or due to the user experience
- Confounding factors are common in case studies due to low level of control

Advantages and Drawbacks

- Advantages
  - Case studies are easier to plan
  - Results are more realistic
- Drawbacks
  - Data are hard to interpret
  - Results are difficult to generalize

Survey vs. Case Study

- Survey is usually done in retrospect
  - Case study is done while a project is executed
- The purpose of surveys is to understand the population
  - Case Study targets a particular project

Experiments

(Controlled) Experiment

- Experiment is an empirical strategy that manipulates one factor (or variable) of the studied setting
- Different treatments are applied to the variable (or variables)
  - Other variables are kept constant
- Experiments are mostly done in a laboratory
  - They require a high level of control

Quantitative (+ Qualitative)

- Experiments are almost pure quantitative
  - Statistical methods are usually applied
- Qualitative data may be used to help in the interpretation and conclusions

Human vs. Technology

- Experiments can be human-oriented or Technology-oriented
- Human-oriented experiments
  - Humans apply different treatments to objects
  - Two groups use different inspection methods
- Technology-oriented experiments
  - Two different tools (e.g., two testing tools) are applied to different objects

Baseline

- It is common to consider the current (or typical) situation as baseline
  - Baseline (control group) is one level of the independent variable
- The new situation (evaluated group) is the one we want to evaluate
  - Another level of the independent variable
- Values of other variables should stay the same, i.e., controlled

Quasi-Experiment

- Quasi-Experiment is similar to experiment
- However, treatments cannot be based on randomization
  - They emerge from characteristics of the subjects or objects
  - Example: it is hart to randomize programming experience in a class

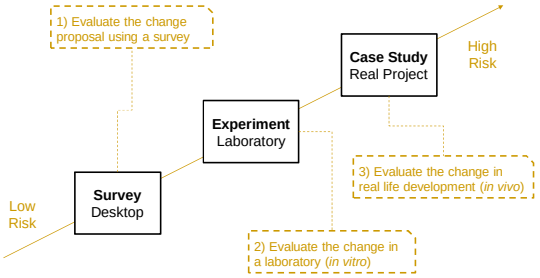
Case Study vs. Experiment

- Different environments
  - Case studies run in real environment
  - Experiments run in controlled environment
- Experiments are more controlled
  - Control is lower in a case study
- Experiments rely on measurements and manipulation of variables
  - Case studies are most observational

Comparative Table

	Survey	Case Study	Experiment
Design Type	Fixed	Flexible	Both
Qualitative / Quantitative	Both	Both	Quantitative
Execution Control	No	No	Yes
Control of Measure	No	Yes	Yes
Costs	Low	Medium	High
Replication	High	Low	High

Strategies Used Together



Bibliography

- C. Wohlin et al. **Experimentation in Software Engineering**, Springer. 2012.
  - Chapter 2 - Empirical Strategies (Sections 2.1 to 2.5)