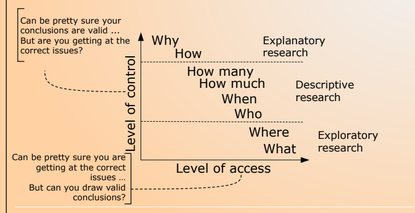
Empirical Research Methods Poster

Check for the latest version at: http://researchmethods.itposter.net

Define research question



Research question examples:

- What are the key success factors of object-oriented frameworks?
- Does the proposed software improvement increases the efficiency of its users?
- How does software development methodology and team size influence developers productivity?

Consider threats to the research

Threats to the research are related to operationalization and measurement issues:

- **Operationalization issues** – The validity of the operationalization
- **Measurement issues** – Reliability, validity, sensitivity (see below)

Reliability threats – refers to the question whether the research can be repeated with the same results.

- **Stability reliability** – Does the measurement vary over time?
- **Representative reliability** – Does the measurement give the same answer when applied to all groups?
- **Equivalence reliability** – When there are many measures of the same construct, do they all give the same answer?

Validity threats

- **Face validity –** Research community »good feel«.
- **Content validity** – Are all aspect of the conceptual variable included in the measurement?
- **Criterion validity** – validity is measured against some other standard or measure for the conceptual variable.
- **Predictive validity** – The measure is known to predict future behavior that is related to the conceptual variable.
- **Construct validity** – A measure is found to give correct predictions in multiple unrelated research processes. This confirm both the theories and the construct validity of the measure.
- **Conclusion validity** – is concerned with the relationship between the treatment and the outcome of the research &choice of sample size, choice of statistical tests).
- **Experimental validity** – (see reliability)

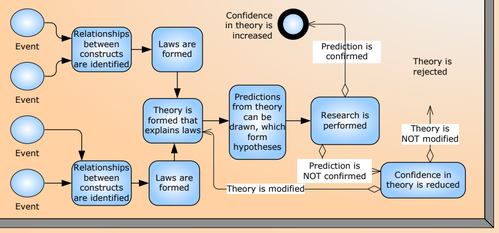
Sources of invalidity

- **Internal** – Is concerned with the validity within the given environment and the reliability of results. It relates to validity of research process design, controls and measures.
- **External** – Is the question of how general the findings are. Can you carry over the research results into actual environment?

Sensitivity

- How much does the measurement change with the change of the conceptual variable?

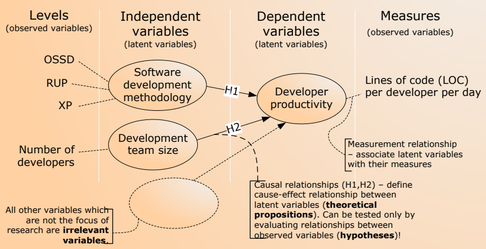
How facts become theories



Create theoretical model

Research question: »How does software development methodology and team size influences developers productivity?«

Theoretical model is based on research question and represents set of concepts and relationships between them!



Hypothesis testing

- Hypotheses are tested by comparing predictions with observed data
- Observations that confirm a prediction do not establish the truth of a hypothesis
- Deductive testing of hypotheses look for disconfirming evidence to falsify hypotheses

Measurement issues

- **Reliability** - does the measurement give the same results under the same conditions (consistency)?
- **Validity** - does the measurement method actually provide information about the conceptual variable?
- **Sensitivity** - how much does the measurement change with the changes on the conceptual variable?

Variables types

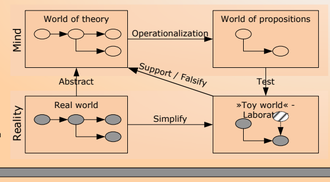
Represent the »effect«	Dependent	Developer efficiency Software reliability
		LOC Mean time between failure
Represent the »cause«	Independent	Requirements change Development team size
		{OSSD, RUP, XP} Number of developers

Latent: Describe abstract theoretical concepts. They cannot be directly measured.

Observed: Define ways of measuring latent variables. Each latent variable may have multiple empirical indicators.

Positivist research model

Positivism is a philosophy that states that the only authentic knowledge is scientific knowledge, and that such knowledge can only come from positive affirmation of theories through strict scientific method.



Qualitative vs. Quantitative analysis

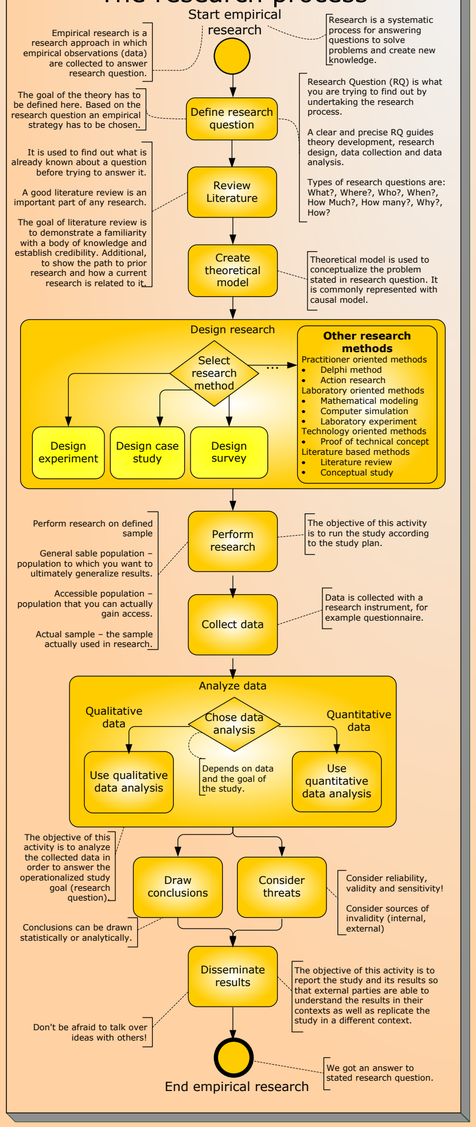
Qualitative (»Judgments«)

- Tends to be the poor relation.
- Problems of opinion and perception when making the judgment.
- The data collected is more likely to create differences of opinion over interpretation.
- Not easily measurable.
- As the benefits are longer term, they can be outweighed by shorter term costs.
- Can lead to inconsistent assessments of performance between places over time and between project elements.
- Subjective opinions tend to be given less status than quantitative ones.

Quantitative (»Hard numbers«)

- Easier to implement and collect data. Tick boxes.
- Easier to make comparisons over time and between places.
- Can be a quick fix when organizations need performance data to justify project investment.
- Easier to process through a computer.
- Easier for other stakeholders to examine and comprehend.
- Trends and patterns easier to identify.
- Can distort the evaluation process as we measure what is easy to measure.
- Can lead to simplistic judgments and the wider more complex picture is ignored.

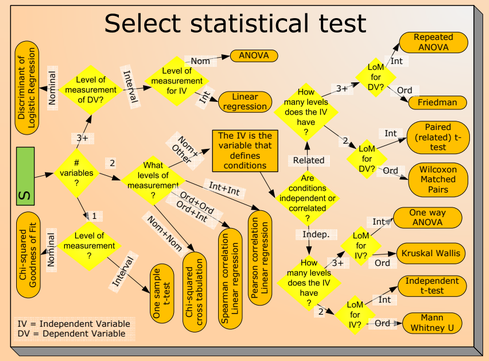
The research process



Select research method

	Experiment	Case study	Survey
Purpose	Establishes causal relationships, confirm theories.	Investigate a typical »case« in realistic representative conditions.	Investigate information collected from a group of people, projects, organizations or literature.
Control	Requires high control	Requires medium control	Requires low control
When appropriate	• Control on who is using which technology, when, where, and under which conditions is possible. • To investigate self standing tasks from which results can be obtained immediately.	• Change to be assessed (e.g., new technology) is wide-ranging throughout the development process. • Assessment in a typical situation required.	• Technology change is implemented across a large number of projects. • Description of results, influence factors, differences and commonalities is needed.
Pro's	• Can establish causal relationships. • Can confirm theories.	• Can be incorporated in normal development activities. • Already scaled up to life size if performed on real projects. • Can determine whether expected effects apply in studied context. • Easy to plan. • Help answer why and how questions. • Can provide qualitative Insights.	• Can use existing experience. • Can confirm an effect generalizes to many projects/organizations. • Allow to use standard statistical techniques. • Enable research in the large. • Applicable to real world projects in practice. • Generalization usually easier. • Good for early exploratory analysis.
Con's	• Application in industrial context requires Compromises.	• With little or no replication they may give inaccurate results. • Difficult to interpret and generalize (e.g., due to confounding factors). • Statistical analysis usually not possible. • Few agreed standards on procedures for undertaking case studies.	• May rely on different projects/organizations keeping comparable data. • No control over variables methods. • Can at most confirm association but not causality. • Can be biased due to differences between respondents and nonrespondents. • Questionnaire design may be tricky (validity, reliability).
Data collection	• Process and product measurement. • Questionnaires.	• Process and product measurement. • Questionnaires. • Interviews	• Questionnaires. • Interviews. • Project measurement. • Literature survey.
Analysis types	Parametric and nonparametric statistics, compare central tendencies of treatments, groups.	Compare case study results to a representative comparison baseline: sister project, company baseline, project subset with no change.	Comparing different populations among respondents, association and trend analysis, consistency of scores.
Major threats	• Conclusion validity • Internal validity • Construct validity • External validity	• Internal validity • Construct validity • External validity • Experimental validity or reliability	• Internal validity • Experimental validity or reliability • Construct validity • External validity

Select statistical test



Analyze quantitative data

Statistics is a mathematical science pertaining to the collection, analysis, interpretation or explanation, and presentation of data.

Statistical methods can be used to summarize or describe a collection of data; this is called **descriptive statistics**. In addition, patterns in the data may be modeled in a way that accounts for randomness and uncertainty in the observations, and then used to draw inferences about the process or population being studied; this is called **inferential statistics**. Both descriptive and inferential statistics comprise **applied statistics**.

Descriptive statistics

Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data.

Measures of central tendency
A measure of central tendency is a single number that is used to represent the average score in the distribution.
Mode – the most common score in a frequency distribution
Median – the middlemost score in a distribution
Mean – the common average

Measures of variability
A single number which describes how much the data vary in the distribution.
Range – The difference between the highest and lower score in a distribution.
Variance – The average of the squared deviations from the mean.
Standard deviation – the square root of the variance, a measure of variability in the same units as the scores being described.

Correlation and regression
Determine associations between two variables.
Correlation – The strength of the relationship between two variables.
Regression – Predicting the value of one variable from another based on the correlation.

Inferential statistics

Inferential statistics or statistical induction comprises the use of statistics to make inferences concerning some unknown aspect of a population.

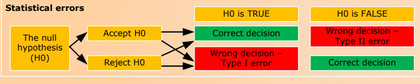
Sampling distribution – the distribution of means of samples from a population.
Sampling distribution has three important properties:

- It has the same mean as the population distribution.
- It has smaller standard deviation as the population distribution.
- As the sample size becomes larger, the shape of the distribution approaches a normal distribution, regardless of the shape of the population from which the samples are drawn.

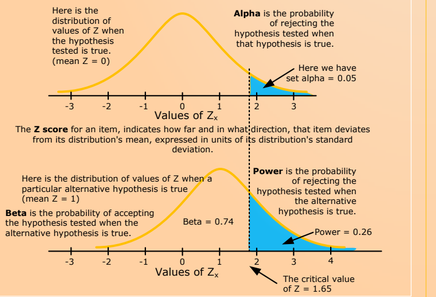
Hypothesis testing - is the use of statistics to determine the probability that a given hypothesis is true. The usual process of hypothesis testing consists of four steps.

- Formulate the **null hypothesis H0** (the hypothesis that is of no scientific interest) and the **alternative hypothesis Ha** (statistical term for the research hypothesis).
- Identify a test statistic that can be used to assess the truth of the null hypothesis.
- Compute the P-value, which is the probability that a test statistic at least as significant as the one observed would be obtained assuming that the null hypothesis were true. The smaller the P-value, the stronger the evidence against the null hypothesis.
- Compare the p-value to an acceptable significance value alpha (sometimes called an alpha value). If p<=alpha, that the observed effect is statistically significant, the null hypothesis is ruled out, and the alternative hypothesis is valid.

Statistical errors



Statistical significance – the probability that an experimental result happened by chance.



Design Experiment

- Analyze <Object(s) of study> (what is studied/ observed?)
- for the purpose of <Purpose> (what is the intention?)
- with respect to their <Quality focus> (which effect is studied?)
- from the point of view of the <Perspective> (whose view?)
- in the context of <Context> (where is the study conducted?).

Variables selection:

- Independent and dependent variables
- Observed variables
- Measurement scales (nominal, ordinal, interval, ratio)

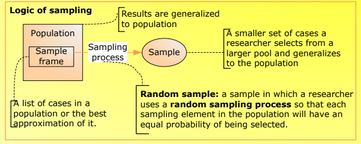
Selection of subjects:

- Profile description
- Quantity
- Separation criteria

Context selection:

- Online vs. Offline
- Student vs. Professional
- Specific vs. general

Logic of sampling



Experiment design:

- Define the set of tests
- How many tests (to make effects visible)
- Link the design to the hypothesis, measurement scales and statistics
- Randomize, block(a construct that probably has an effect on response) and balance(equal number of subjects)

Hypothesis formulation:

- Hypoth. statement
- H_0: positive
- H_1: One or two tailed

Experiment design	Random assignment	Pretest	Posttest	Control group	Experimental group	Design notation
Classical	Yes	Yes	Yes	Yes	Yes	R O X O O
One shot case study	No	No	Yes	No	Yes	X O
One group pretest posttest	No	Yes	Yes	No	Yes	O X O
Static group comparison	No	No	Yes	Yes	Yes	X O O
Two group posttest only	Yes	No	Yes	Yes	Yes	R X O O
Time series design	No	Yes	Yes	No	Yes	O X O X O

Experiment design notation:

X = Treatment

O = Observation

R = Random assignment

Design case study

Case method facts:

- Does not explicitly control or manipulate variables.
- Studies a phenomenon in its natural context.
- Makes use of qualitative tools and techniques for data collection and analysis.
- Case study research can be used in a number of different ways.
- Can be used for description, discovery and theory testing.

Varieties of case study research:

- Case studies can be carried out by taking a positivist or interpretivist approach.
- Can be deductive and inductive.
- Can use qualitative or quantitative methods.
- Can investigate one or multiple cases.

Case research design:

Single case	• Investigate a phenomenon in depth, get close to the phenomenon, provide a rich description and reveal its deep structure.
Multiple case	• Enable the analysis of data across cases, which enable the researcher to verify that findings are not the result of idiosyncrasies of the research setting. • Cross case comparison allows the researcher to use literal or theoretical replication.

- In **exploratory** case studies, fieldwork, and data collection may be undertaken prior to definition of the research questions and hypotheses.
- **Explanatory** cases are suitable for doing causal studies. In very complex and multivariate cases, the analysis can make use of pattern-matching techniques.
- **Descriptive** cases require that the investigator begin with a descriptive theory, or face the possibility that problems will occur during the project.

Case research objectives:

Discovery and induction:

- Discovery is the description and conceptualization of the phenomena.
- Conceptualization is achieved by generating hypotheses and developing explanations for observed relationships
- Statements about relationships provide the basis for the building of theory.

Testing and deduction:

- Testing is concerned with validating or disconfirming existing theory.
- Deduction is a means of testing theory according to the natural science model

Case study research design components:

- A study's question.
- Its propositions, if any.
- The unit of analysis.
- The logic linking the data to propositions.
- The criteria for interpreting the findings.

Design Survey

A survey is a study by asking (a group of) people from a population about their opinion on specific issue with the intention to define relationships outcomes on this issue.

Survey Process:

- Study definition – determining the goal of a survey.
- Design – operationalizing of the study goals into a set of questions (see theoretical model)
- Implementation – operationalisation of the design so that the survey will be executable.
- Execution – the actual data collection and data processing.
- Analysis – interpretation of the data.
- Packaging – reporting about the survey results.

Types of survey

- **Descriptive surveys** are frequently conducted to enable descriptive assertions about some population, i.e., discovering the distribution of certain features or attributes. The concern is not about why the observed distribution exists, but instead what that distribution is.
- **Explanatory surveys** aim at making explanatory claims about the population. For example, when studying how developers use a certain inspection technique, we might want to explain why some developers prefer one technique while others prefer another. By examining the relationships between different candidate techniques and several explanatory variables, we may try to explain why developers choose one of the techniques.
- **Explorative surveys** are used as a pre-study to a more thorough investigation to assure that important issues are not foreseen. This could be done by creating a loosely structured questionnaire and letting a sample from the population answer to it. The information is gathered and analyzed, and the results are used to improve the full investigation. In other words, the explorative survey does not answer the basic research question, but it may provide new possibilities that could be analyzed and should therefore be followed up in the more focused or thorough survey.

Data analysis

- Coding scheme (for open question)
- Data entry
- Checking
- Resolve incomplete data
- Statistical testing of results

Question types

- Open
- Closed

Reporting response rate

- Total sample selected
- Number located
- Number contacted
- Number returned
- Number complete

Literature used

- Bernd Freimut, Teade Punter, Stefan Biffl, & Marcus Ciolkowski 2002, *State-of-the-Art in Empirical Studies*, Virtuelles Software Engineering Kompetenz-zentrum.
- Johnston, R. & Shanks, G. Research Methods in Information Systems. 2003.
- Neuman, W. L. 2005, *Social research methods : qualitative and quantitative approaches*, 5th ed. edn.
- Winston Tellis 1997, "Introduction to Case Study", *The Qualitative Report*, vol. 3, no. 2.
- www.wikipedia.org

About the Research Methods Poster

Authors:
Gregor Polančič
Email: info@itposter.net

University of Maribor
Faculty of Electrical Engineering and Computer Science
Institute of Informatics

Poster version: 0.6 (DRAFT)
http://researchmethods.itposter.net