Notes Advanced Research Methods

Wieringa – Chapter 1 – What is Design Science?

Every design **science investigates an artifact in context by the means of two activities: designing and investigating.**

- The design activity must have the **social context** of **stakeholders** and **goals** of the project clear
- The investigative activity needs to have **clear the knowledge context** of the project.

<u>The object of study of design science</u>

Design science is the *design and investigation of artifacts in context.* The artifacts are designed to interact with a problem context in order to improve something in that context (e.g. a method to improve compliance of X in Y is an artifact).

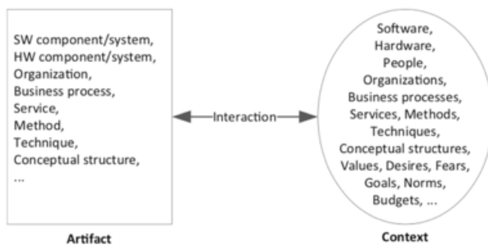


Figure 1. Example of artifact and context

Design science problems are therefore improvement problems → each of the problems has a context in which some improvement is aimed for, thus the context also has to be understood.

<u>Research problems in Design Science</u>

Design and investigation correspond to two research problems:

- Design problems
- Knowledge questions

Design problems (DP) need a change in the real world and **require an analysis of actual/hypothetical stakeholder goals**. A solution is a design, and it needs to be evaluated by their utility.

Fig. 1.2 Design science research iterates over two problem-solving activities

Table 1.1 Some example design science research problems. In the top half of the table, the knowledge questions are motivated by the design problems. In the lower half, the design problems are motivated by the knowledge questions

Design problem	Knowledge question
Design a DOA estimation system for satellite TV reception in a car	Is the DOA estimation accurate enough?
Design an assurance method for DLC for cloud service providers	Is the method usable and useful for cloud service providers?
Design a DOA prototype Design a simulation of plane wave arrival at a moving antenna	Is the DOA estimation accurate enough?
Design a usability and usefulness test with consultants as subjects	Is the method usable and useful for cloud service providers?

Knowledge questions (KQ) ask for knowledge about how it the world around you, they do not need a solution. One must pay attention to not camouflage a design problems as a knowledge question.

Table 1.2 Heuristics to distinguish design problems from knowledge questions

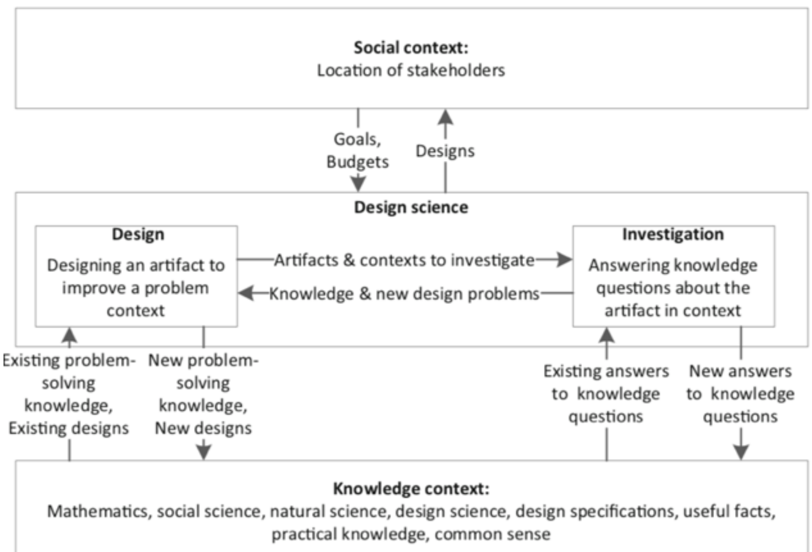
Design problems	Knowledge questions
Call for a change of the world	Ask for knowledge about the world
Solution is a design	Answer is a proposition
Many solutions	One answer
Evaluated by utility	Evaluated by truth
Utility depends on stakeholder goals	Truth does not depend on stakeholder goals

Figure 2. How to distinguish between DP and KQ

A framework for design science

Problem context of an artifact

It can be extended with the stakeholders of the artifact and with the knowledge used to design the artifact. The resulting picture is a framework for design science.

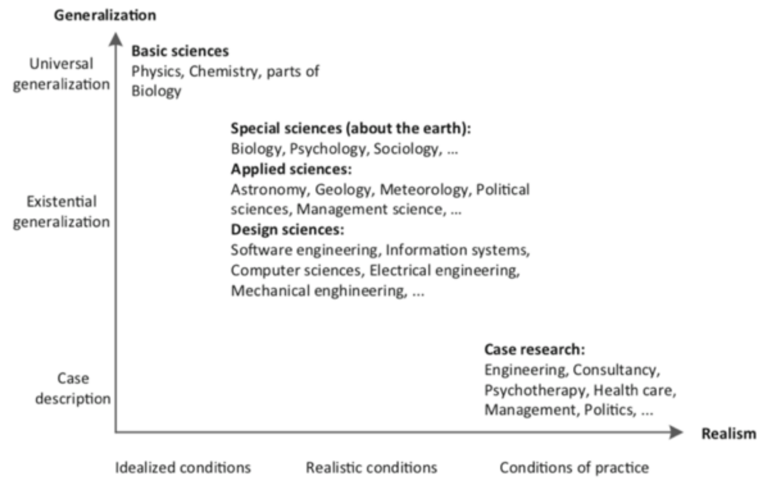


The knowledge context consists of **existing theories** from data science and engineering, facts and lessons learned. The design science project uses this knowledge and can add to it creating new designs or answering knowledge questions. Prior knowledge available before the beginning of the design science project is called **posterior knowledge** (e.g. scientific literature, technical literature, Professional literature, oral communication)**.**

Science of the middle range

The knowledge context of the design science research can be expanded by identifying **different knowledge disciplines**:

- **Basic science:** the study of *fundamental forces and mechanisms of nature* → **aim at universal generalizations at the cost of idealizations**, which are abstraction known to be false in the real word (e.g. economics rational assumptions)
- **Case research:** the other extreme: it **cannot idealize from the real world and must deals with the conditions of practice**. Primary aim is to help the cases it is working with directly.



The middle range sees **special sciences:** that are *happy to generalize about interesting set of objects but do not aim at generalizing the whole universe;* **applied sciences** which apply results from other sciences but have developed knowledge of their own; **design sciences** like software engineering and information systems research. They are in the middle range because they aim at *existential generalization making realistic assumptions about their object of study* – so called **middle-range generalizations.**

On the bottom left part there are no sciences that make idealized description of individual cases, same as the top right part where there are no sciences which make universal generalization about cases without making any abstraction.

⇒*Our aim is to produce knowledge about the real world that does not make any unrealistic abstractions and that has a scope of validity that is as large as possible (top right corner).*

Wieringa – Chapter 2 – Research goals and Research Questions

To **frame** a research project, one has to **specify its research goal**. A design science *project iterates over **designing** and **investigating** and therefore its research goal can be refined **into design goals** and **knowledge goals**.* This chapter gives a template for design problems and a classification of different kinds of knowledge goals.

<u>Research goals</u>

One must distinguish between:

- **Goals of the researchers**: **Researchers** will have a desire and curiosity to improve society and to do research,
- **Goals of an external stakeholder** while the stakeholder will more likely allocate the budget to the research in order to achieve some goals and receive back useful designs and knowledge about those designs.

* For sponsors design science research projects are utility driven and budget constrained (sometimes also exploratory)
* For researches design science research projects are curiosity- and fun-driven.

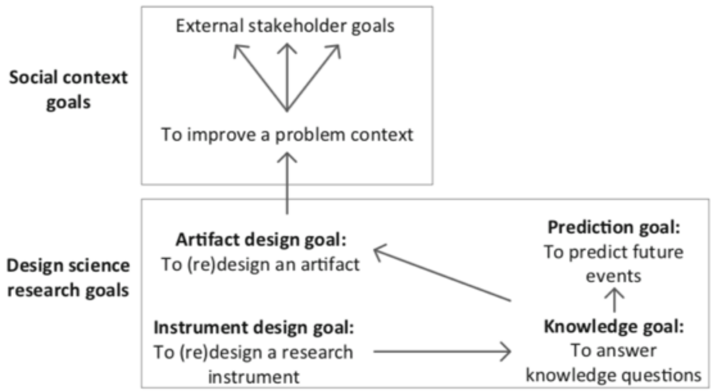


Figure 3. Goal structure of a design science research project. The goals on the left concern improvement of the real world, and those on the right concern our beliefs about the world. In an exploratory project, there may be no higher-level improvement goals

Design science researchers often have a **prediction goal.** A prediction is a belief about what will happen in the future.

* **Knowledge goals** instead describe phenomena and explain them. To answer the knowledge questions one may have to design **instrument design goals** (lowest level design goals).
* **Artifact design goals** or **technical research goal** have the objective of solving, migrating or improving some problems in the social context of the project.

DOA example (3)

The DOA project is market driven. The lowest level goal was to build simulations and prototypes of DOA algorithms and of an antenna array. This is an instrument design goal. These instruments were used to answer knowledge questions about the performance of different DOA algorithms—a knowledge goal. This knowledge was generalizable and could be used to predict the performance of all implementations of the algorithm—another knowledge goal. Answering these questions also contributed to the artifact design goal of designing a DOA estimation component. This in turn contributes the goal of problem context improvement. The DOA estimation component will be part of a directional antenna for satellite TV signal reception, which is to be used in a car to allow passengers on the backseat to watch TV. The sponsor's goal is to develop and sell components of the IT infrastructure needed for this.

<u>Design problems</u>

A design problem is **a problem to (re)design an artifact so that it better contributes to the achievement of some goal**. Fixing the goal put us at some level in the goal hierarchy. Design problems assume a context and stakeholder goals and call for an artifact such that the interactions of help stakeholder to achieve their goals.

Template for design problems
1. Improve <*a problem context*>
2. by <*(re)designing an artifact*>
3. that satisfies <*some requirements*>
4. in order to <*help stakeholders achieve some goals*>.

DOA example (4)
1. Improve satellite TV reception in cars
2. by designing a DOA estimation algorithm
3. that satisfies accuracy and speed requirements
4. so that passengers can watch TV in the car.

Not all elements of the design problem template may be known at the start of the project! Stating the problem can help you identifying missing pieces of information that are needed to bound the research problem.

Knowledge questions

The knowledge goals of a project should be refined into knowledge questions.
All the knowledge here are empirical questions that require data about the world to answer them. Some could also be analytical KQ which can be answered by conceptual analysis.
One must distinguish between KQ and prediction problems. Prediction problems are about the future and we cannot *know* the future.

Descriptive and explanatory questions
- **Descriptive questions:** ask for what happened without asking for explanations (journalistic) – *what, where, when, who*
- **Explanatory questions:** ask *why* something happened – *why*
  - Causation of event
  - Production of event
  - Reason behind event

Open and closed questions
- **Open question** contains no specification of its possible answers, it is exploratory
- **Closed question** contains hypotheses about its possible answers

Therefore there are 3 empirical knowledge questions:
1. **Effect questions: (**artifact * context) produce effects? E.g. what effect are produced by the interaction?
   i What effect an artifact in a context has
2. **Trade-off questions**: (alternative artifact * context) produce effect? E.g. what effect do similar artifact produce?
   i What is the difference between effects of different artifacts in the same context
3. **Sensitivity questions**: (artifact * alternative context) produce effect? E.g. what effect produced in different context?
   i What is the difference between same artifacts in different context
4. **Requirements satisfaction questions**: do effects satisfy requirements? E.g. performance satisfies nonfunctional requirements?

Wieringa – Chapter 3 –The Design Cycle

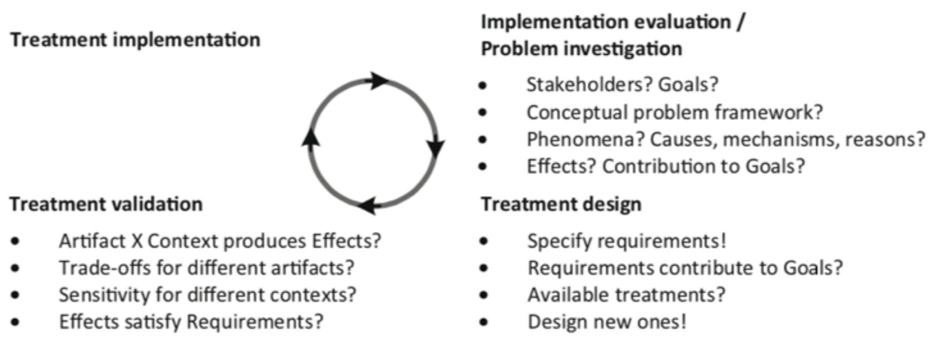
A design science project iterates over the activities of:

- Designing, composed of a cycle of three activities
  - Problem Investigation
  - Treatment design
  - Treatment validation
- Investigating

The Design and Engineering Cycles

The engineering cycle is a rational problem solving process that consists of:

- Investigating the problem (what phenomena can be improved and why?
- Treatment design: design some artifacts that could treat the problem
- Treatment implementation: treat the problem with one of the designed artifacts
- Implantation of the evaluation: how successful was the treatment?



The goal of the implementation evalutation is to evaluate a treatment after it has been applied to the problem context. The goal of the problem investigation is instead to prepare for the design of a treatment by learning more about the problem.

Treatment

Avoid say designing solution because the artifact could only maybe solve partially the problem. It is therefore preferred to use the term ***interventions***. As it suggests that the artifact will only interact with the context.

- ***Treatment*** here is the interaction between the artifact (the intervention) and the problem context.

Artifacts

They are used when designing, developing, implementing, maintaining information systems and software systems.

- An artifact Is something created by people for some practical purpose. It is used by people, and it interacts with a context that contains people.

Design and specification

Treatments (which are artifacts) are designed and these designs documented in a specification.

- **A design** is a decision about what to do
- **A specification** is a documentation of the design decision

Implementation

- **An implementation** of a treatment is the application of the treatment to the original problem context.

An implementation is the transfer to the problem context – when we use the artifact (what we think of a solution) and we put it into working mode somewhere.

Validation and Evaluation

- **To validate a treatment** means justifying that would contributed (positively/negatively or just contribute?) to stakeholder goals
    - **Goal:** predicting how an artifact will interact – usually done in lab
- **Evaluating a treatment** means on the other hand investigating the treatment as applied by stakeholders in the field – done after implementation
    - **Goal:** investigating how implemented artifact interact with real world context

Engineering Processes

The engineering and design cycles provide a logical structure of tasks but do not prescribe the process of engineering or of designing. They tells us that we have to understand, justify by validating and implementing and evaluating. They do not tell us how to manage these processes.

Possible execution sequences include:

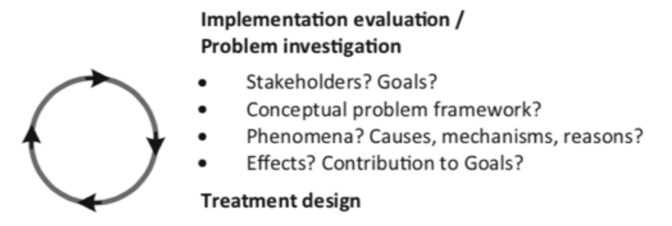
- **Waterfall development process:** one sequential pass on engineering cycle without backtracking. Only possible if problem is understood
- **Agile development:** many sequential passes through the cycle. In each pass a problem is treated of a size small enough to pass through the entire engineering cycle
- **Scaling up:** many passes are made simplifying assumption and testing at some level of realism the artifact
- **System Engineering:** number of iterations through the design cycles are performed and then validated conceptually without making any implementation attempts

Summary

- The design cycle consists **of problem investigation, treatment design, and treatment validation.**
- **The design cycle is part of the engineering cycle**, in which a designed and validated treatment is implemented in the problem context, and the implementation is evaluated.
- Implementation evaluation may be the problem investigation of a new engineering cycle.
- Managing the research and development process includes deciding what to do when, how to align with stakeholders, and how to use finite resources to achieve the research goal. This is out of scope of the engineering and design cycles.

Wieringa – Chapter 4 – Stakeholders and Goal Analysis

Even in exploratory projects is useful to understand who are the people who could be interested in our project, the stakeholders.



<u>Stakeholders</u>

- A person, group or institution affected by treating a problem.
- They are source of goals and constraints of the project – source for requirements in the treatment

⇒ *The whole goal of the treatment design is that some stakeholders are better off when the problem is treated.*

<u>Desires and goals</u>

There are multiple levels of awareness of a problem in which stakeholders can be:

1. No awareness of the problem or of the need of a treatment
2. Awareness of an improvement possibility but no interest in carrying out the improvement
3. Awareness of the improvement possibility and desire but not able to commit resources into realizing it

⇒ ***Stakeholder goal:*** *the desire for which the stakeholder has committed resources (finite)*

<u>List of possible stakeholder of an artifact</u>

- *Normal operator*: give routine commands to the artifact: end users
- *Maintenance operators*: interact with the system
- *Operational support staff*: help the system
- *Functional beneficiaries*: benefit from the output – user of the artefact
    - Interface the system and have an interest
- Financial beneficiary and political beneficiary draw benefits from it
- *Negative stakeholder*: would be worse off when the artifact is applied
- *Threat agent*: wants to hurt the system
- *Sponsor*: initiate and provides a budget
    - Purchases: terminates the development successfully
- *Developers*: not normal operator of the system
- *Consultants and suppliers*: support and deliver components of the artifact

<u>Desires and conflicts</u>

2 desires are in **logical conflict** if it is logically impossible to realize them both

- *Physical conflict* when there would be a violation of the laws of nature to satisfy them both
- *Technical conflict* when it would be physically possible to realize them but there are no technical means to achieve this
- *Economic conflict*: technically possible but no budget
- *Legal conflict*: illegal to realize them
- *Moral conflict*: morally wrong to realize them

Wieringa – Chapter 5 – Implementation evaluation and problem investigation

Once treatment are implemented there may be new properties provided that could trigger a new iteration on the engineering cycle.

Research goals

- **Implementation evaluation**: evaluate an implementation of a treatment **after** it has been applied in the original problem
    - **An evaluation question** compares an observed phenomena with a norm
- **Problem investigation:** research goal is to investigate an improvement problem **before** an artifact is designed and when no requirements have been identified yet

Theories

The goal of implementation evaluation and problem investigation is to build a scientific theory of real-world implementations and problems.

⇒ *A scientific theory is a belief about a pattern in phenomena that has survived testing against empirical facts and critical reviews by peers. It has a conceptual framework to frame the problem, describe phenomena and analyse their structure. Generalization can be useful to explain the causes or mechanisms or reasons for phenomena.*

Design science therefore is a scientific theory that **studies the interaction between an artifact and its context** – developing design theories. The phenomena studies in implementation evaluation and problem investigation are real-world phenomena.

Research methods

- Surveys
- Observational case studies
    - May give access to all aspects of the studied phenomena. However, they may disturb the phenomena that one is observing.
- Single-case mechanism Experiments
    - Test of a single case in which the researcher applies stimuli to the case and explains the responses in terms of mechanisms internal to the case
    - The researcher actually intervenes in the experiment
- Statistical difference making experiments
    - Apply a treatment to a sample and compare the average outcome with the average of another treatment.
    - Sample drawn at random
    - Used in implementation evaluation and problem investigation

Summary

- **Implementation evaluation is the investigation of artifacts that have been transferred** to their intended real-world problem context. **The research goal is to evaluate them** with respect to actual stakeholder goals.
- **Problem investigation is the investigation of real-world problems as a preparation for the design of a treatment** for the problem. The research goal is to learn about stakeholder goals and to understand the problem to be treated.
- Both kinds of empirical study ask the same research questions that are about phenomena, causes and effects, and the contribution of phenomena to stakeholder goals. Questions can be open or closed and descriptive or explanatory.
- Implementation evaluation and problem investigation aim to develop theories of phenomena. A theory consists of a conceptual framework and generalizations. If the theory is about the interaction of an artifact with its context, we call it a design theory.
- Implementation evaluation and problem investigation are real-world research. Different research methods can be used, including surveys, observational case studies, single-case mechanism experiments, and statistical difference-making experiments.

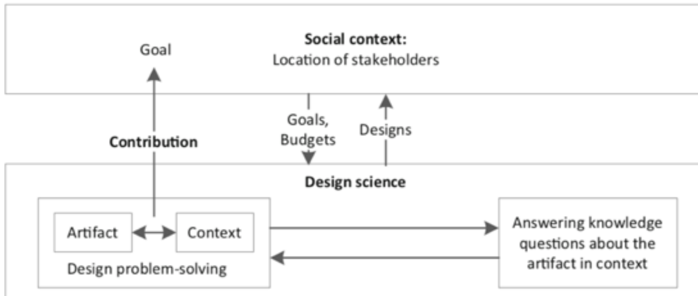
Wieringa – Chapter 6 – Requirements Specification

Think about the **desired properties of a treatment** before designing one.

- A requirement Is a **property of the treatment desired by some stakeholder with committed resources**
- It is a goal for the to-be-designed treatment

Contribution Arguments

- **Specifying requirements is it a duty of the design researcher**
- To justify the choice of requirements one has to give a contribution argument → an argument that an artifact that satisfies the requirements would contribute to a stakeholder goal in the problem context



A contribution argument is a prediction: it argues that the artifact, once inserted in its problem context, would interact with it in a way that contributes to stakeholders goals.

Kind of requirements

- Classified in different ways: according to the **importance of the goal** (priority) or to the **deadline** (urgency)
- Constraints on the **internal** composition of the artifact are distinguished from requirements on the **external** properties
- **Functional** and **non-functional** requirements
    - **Function** of an artifact is a terminating part of the interaction between an artifact and its context
    - **Non-functional property** is called also *quality property,* and it is any property that is not a function (usually flobal properties e.g. *utility* for a stakeholder, *accuracy* of output, …)
- **Indicators and norms**
    - An **Operationalization** of a property is a measurement procedure to establish the presence of a property
        - Functional properties are operationalized by specifying tests for them
        - Non-functional properties are operationalized by defining one or more indicators (variables to be measured that indicate the presence of a property - metrics)
    - **Norm** for an indicator is a set of required values of the indicators (also acceptance criteria)
        - Norms operationalize requirements
        - Norms are a set of required values of the indicator
        - Examples: "The artifact should recognize that an action happened" (ability) or "the estimation of an outcome should be less than x" (accuracy)
    - **Indicator** operationalize properties

Summary

- Requirements are treatment goals. They are desired by some stakeholder who has committed a budget
- Requirements are specified by the design researcher. To justify requirements, there should be a contribution argument
- Between functional and nonfunctional requirements:
    - To make a nonfunctional property measurable, it must be operationalized by indicators.
    - To make a nonfunctional requirement measurable, it must be operationalized by indicators and norms.
    - Sometimes a norm is crisp, but often it just indicates the direction of improvement.

Wieringa – Chapter 7 – Treatment Validation

To validate a treatment means justifying that it would contribute to the aforementioned stakeholder goals when implemented in the context defined. When the requirements are defined, then we can validate the treatment by showing that it satisfies its requirements.

⇒ *We want to understand and investigate how does the treatment contribute to stakeholder goals –> build validation models of the artifact in context, investigate them and build a design theory of the artifact in context to predict effect in the real world*

<u>The validation research goal</u>

Goal: develop a **design theory of an artifact in context** that allows us **to predict what would happen** if the artifact were transferred to its intended problem context.

The central problem of validation research is that is done **before implementation** so how can we investigate if an implementation in the real world does not exist?

We have to ask knowledge questions of the kind already shown in ch 2:

Effect questions: (artifact × context) produce Effects?

- What effects are produced by the interaction between the artifact and context?
- How does the artifact respond to stimuli?
- What performance does it have in this context? (Different variables)

Trade-off questions: (alternative artifact × context) produce effects?

- What effects do similar artifacts have in this context?
- How does the artifact perform in this context compared to similar artifacts?
- How do different versions of the same artifact perform in this context?

Sensitivity questions: (artifact × alternative context) produce Effects?

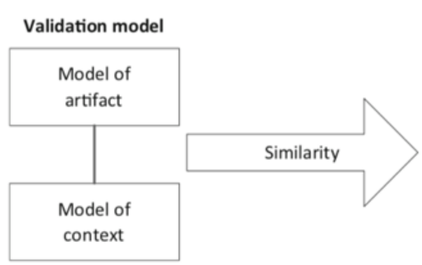
- What effects are produced by the artifact in different contexts?
- What happens if the context becomes bigger/smaller?
- What assumptions does the design of the artifact make about its context?

Requirements satisfaction questions: do effects satisfy requirements?

- Does the stimulus-response behavior satisfy functional requirements?
- Does the performance satisfy nonfunctional requirements?

<u>Validation model</u>

- **An analogic model:** is an entity that represents entities of interest, called **targets** (examples is the use of electrical network as model of a network of water pipes)
- **A validation model** consists of model of the artifact interacting with a model of the problem context (model of the artifact interacting with a model of the problem context)



<u>Design Theories</u>

Once we have a theory that has successfully survived tests and criticism we use it to predict

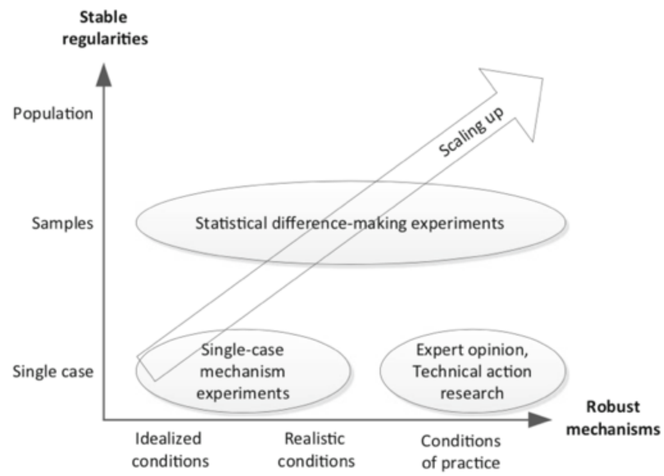
<u>Research Methods</u>

To validate models we have different methods:

- **Expert Opinion**: design is submitted to a panel of experts → is their requirements are not satisfied it has to be redesigned. Negative opinions are more useful than positive ones
- **Single case Mechanism Experiments** is a test in which the **researcher applies stimuli to a validation model** and explains the response in terms of mechanisms internal to the model. It is useful for validation research and allows us to expose the model to controlled stimuli
- **Technical Action Research**: use of an **artifact prototype** to help a client and to learn from this. Usually at last stage
- **Statistical difference making experiment**: compare the average outcome of treatments applied to sample

Scaling up to stable regularities and robust mechanism

In the scaling up approach two lines of reasoning are followed:



- **Horizontally:** the research goal is to test an increasingly realistic model of the artifact (more practice, more real world) – **case-based inference**
- **Vertically:** effect exists on the average in a population. Statistic to average out the nondeterminism by observing average outcomes in samples and make plausible conclusions. **Sample-based inference**

Summary

- In treatment validation, we develop a design theory of the interactions between the artifact and its context.
- Important validation research questions are **what the effects of the interaction are and whether these satisfy requirements.** *Trade-off questions test the* ***generalizability over different versions*** of the artifact; *sensitivity questions test the generalizability* ***over differences in context.***
- Since we cannot, by definition, validate real-world implementations, we investigate validation models. A validation model consists of a model of the artifact interacting with a model of the context.
- Research methods to investigate validation models include expert opinion, single-case mechanism experiments, technical action research, and statistical difference-making experiments

Wieringa – Chapter 10 – The empirical Cycle

It is a **rational way to answer scientific knowledge questions**. It is structured as **a checklist of issues** to decide when a researcher designs a research setups and has to reason about the data

⇒ *It is a checklist of questions that help to find justifiable answers to scientific knowledge questions, not a sequence to be performed in order.*

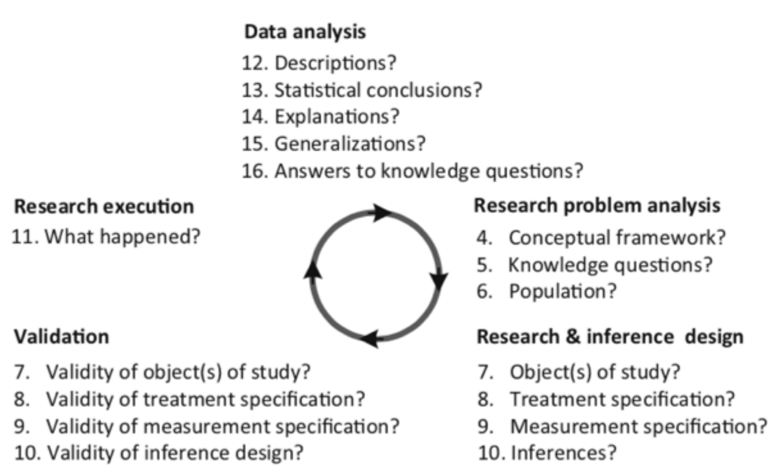
<u>Research context</u>

1. Knowledge goal(s)
   - What do you want to know? Is this part of an implementation evaluation, a problem investigation, a survey of existing treatments, or a new technology validation?
2. Improvement goal(s)?
   - If there is a higher-level engineering cycle, what is the goal of that cycle?
   - If this is a curiosity-driven project, are there credible application scenarios for the project results?
3. Current knowledge
   - State of the knowledge in published scientific, technical, and professional literature?
   - Available expert knowledge?
   - Why is your research needed? Do you want to add anything, e.g., confirm or falsify something?
   - Theoretical framework that you will use?

17. Contribution to knowledge goal(s)
    - Refer back to items 1 and 3
18. Contribution to improvement goal(s)?
    - Refer back to item 2
    - If there is no improvement goal, is there a potential contribution to practice?

- **(1) Knowledge goal** summarizes in one phrase all knowledge questions that you want to answer in the study
  - Typical are investigate an improvement problem
  - Survey possible treatments
  - Validate a design
  - Evaluate an implementation in the field
- **(2) Working in the context of a higher-level design or engineering cycle with an improvement goal**
- **(3) Summarize current knowledge**

<u>The Empirical Cycle</u>



Validation in the empirical cycle is about the **match between the research setup and the inferences form the data**. There are three kinds of validity questions about a research design:

- *Inference support*

- *Repeatability*
- *Ethics*

The research problem

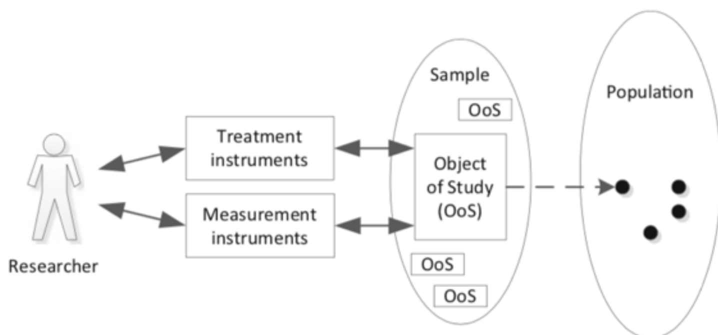
A checklist to frame an empirical research problem

4. Conceptual framework
   - Conceptual structures? Architectural structures, statistical structures?
   - Chance models of random variables: semantics of variables?
   - Validity of the conceptual framework? Clarity of definitions, unambiguous application, avoidance of mono-operation and mono-method bias?
5. Knowledge questions
   - Open (exploratory) or closed (hypothesis-testing) questions?
   - Effect, satisfaction, trade-off, or sensitivity questions?
   - Descriptive or explanatory questions?
6. Population
   - Population predicate? What is the architecture of the elements of the population? In which ways are all population elements similar to each other and dissimilar to other elements?
   - Chance models of random variables: assumptions about distributions of variables?

The research setup

To answer a knowledge question about a population there needs to be one or more objects of study. These have some relationship with the population.

- **In case-based research:** the object of study is called a case
- **In sample based research**: the researcher studies a sample of cases



Differences

- In **case-based** research we **study individual objects** and **aim to generalize** to similar object
  - Case-based uses an analytical **induction** process (from one case tries to generalize to the whole population
- **In sample based we study samples and aim to generalize from samples**
  - Studies samples as a whole and is **deductive** (from general to a particular theory) – uses statistical inference to draw conclusions about the population

Experimental research vs observational research

- **Experimental:** researcher applies an experimental treatment to the Oos and then measures what happens
- **Observational:** reserachers refrains from intervening and measures phenomena in the Oos

	Observational study (no treatment)	Experimental study (treatment)
Case-based research	• Observational case study (17)	• Single-case experiment (18) • Comparative-cases experiment (14) • Technical action research (19)
Sample-based research	• Survey • Quasi-experiment (20)	• Randomized controlled trial (20) • Quasi-experiment (20)

Inferences from Data

- **Inference:** is the process of drawing conclusions from data
  - **Ampliative:** *conclusions may be false but premises are true*
  - **Deductive:** *conclusions are guaranteed to be true when premises are true*
- A discussion regarding the validity, which is the degree of support of the inferences, must accompany both of them

Some categorizations:

- **Descriptive inference**: summarize data into descriptions, subject to constraints of descriptive validity
- **Statistical inference**: inference of population characteristics from sample statistics; subject to conclusion validity
- **Abductive inference**: most plausible explanations for observations; subject to internal validity
- **Analogic inference**: generalization of explanations; subject to external validity

Execution and Data Analysis

Oos may interact even if no plan to do so, and there could be may more points of failure in the execution of the research. Therefore the validity of these inference designs is affected by unplanned events during research execution meaning that the validity discussion may be revisited during data analysis under the light of the new events that came up.

The Empirical cycle is not a research process

It is only a grouping of questions and not a temporal sequence of task that must be performed. Almost anything can be changed, repeated or redone. The only two prohibited things are:

- **Rule of posterior knowledge:** knowledge created by the research is present only after execution and cannot be used before executing the research
  - **Rule of prior ignorance:** any knowledge present before the research may influence the outcome of the research.
- **Rule of full disclosure:** all events that could have influenced research conclusions must be reported

Summary

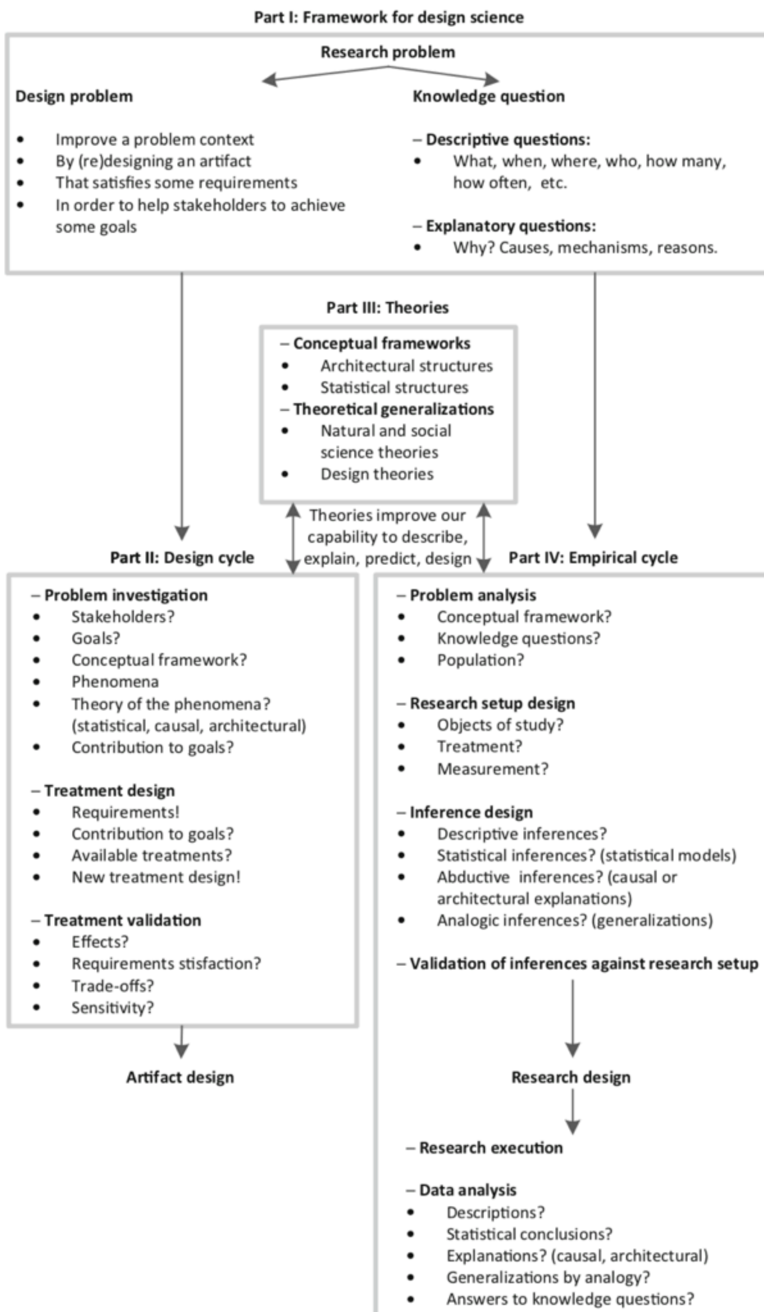
- In design research, we try to answer knowledge questions about implementations, problems, or treatments. The research context may be utility driven or curiosity driven.
- Empirical research is expensive, and before you attempt it, you should check other ways of answering knowledge questions, including literature study and asking experts.
- The research problem to be solved consists of a list of knowledge questions that presuppose a conceptual framework and assume a population of interest. The goal of research is to find support for a generalization to that population.
- The empirical cycle is a logical grouping of questions to ask about your research. You may also use it to decide what to put in a report and to analyze a report.
- The empirical cycle is not a list of tasks to be performed in a certain order. Almost anything goes, but two things are forbidden:
  - Rule of posterior knowledge: Knowledge created by the research is present *after* execution of the research, and it is absent before executing the research.
  - Rule of prior ignorance: Any knowledge present *before* doing the research may influence the outcome of the research.
- To ensure this, publication is required:
  - Rule of full disclosure: All events that could have influenced research conclusions must be reported.
- Research setup and inference are closely coupled, because the setup must support the inferences to be done later. Both must be designed before you start the research.

Wieringa – Chapter 16 – A Road Map of Research Methods

- Design science research iterates over solving design problems and answering knowledge questions.
- Design problems that need novel treatment are dealt with the design cycle
- Knowledge questions that require empirical research are dealt with the empirical cycle
- Design and empirical research require a theoretical knowledge in terms of a framework and generalization.

The outcome of the design cycle is a validated artifact design but not an implementation in the context of use. Implementation in the social context of use is transfer of technology to a stakeholder context → it is not part of design science but may be a sequel to it.

The **empirical cycle goes beyond research design by executing the design and analysing the results**, *which are then used to answer the knowledge questions that triggered the empirical cycle.*



Four empirical Research Methods

- **Observational Case Studies:** are about individual case to investigate how the phenomena in case are produced by the architecture of the case. There is not intervention of the researcher. Are done in the field
- **Single-case mechanism experiments:** individual cases are studied – to investigate how phenomena in the case are produced by the architecture of the case. There is researcher intervention as they are often done in the laboratory. Used and useful to validate new technology and investigating problems in the fields
- **Technical Action Research:** researcher experiments with single cases – not only to answer a knowledge question but also to help a client.
- **Statistical difference-making experiments:** two sample of population elements different treatments in order to find out if the difference in treatments causes a difference

One checklist

Is a list of possible choices. It starts by questioning the positioning of the researcher in the goal structure of the design science project then continues with all the questions we already asked ourselves:

- **What is studied:** know which population of interests and what are its elements
- **Where is studied:** laboratory or field?
- **How it is studied:** intervene or not to intervene?
- **Why it is studied:** research goals

Possible example of this:

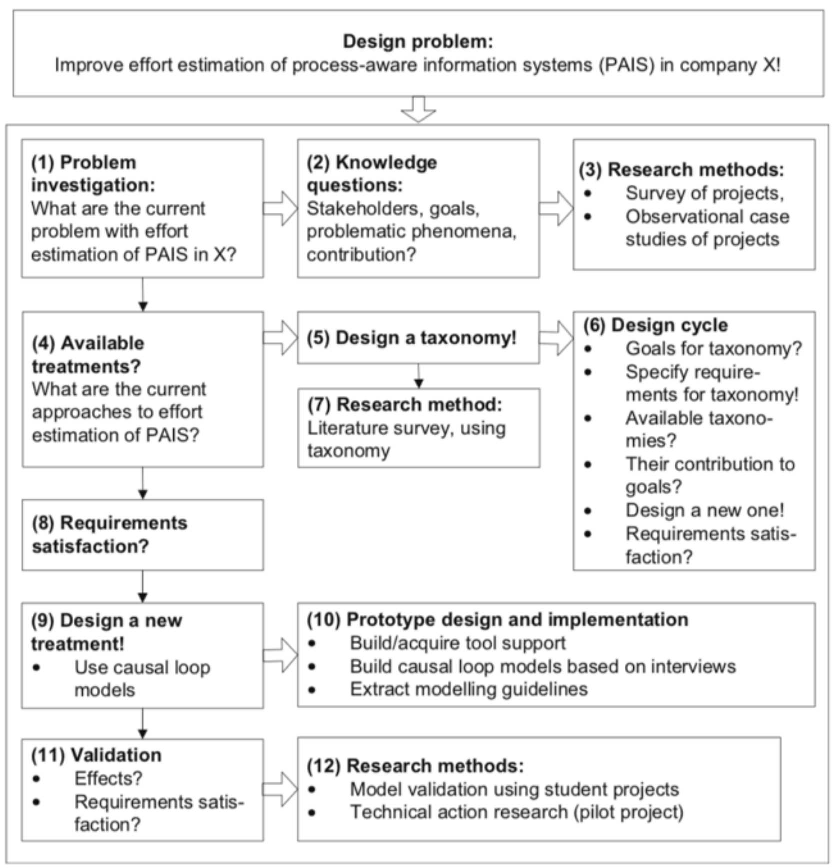


Fig. 16.3 Methodological structure of a PhD thesis. The *hollow arrows* are problem decompositions, the *single arrows* roughly indicate temporal sequence of problem-solving, and the numbers indicate the sequence in which tasks were performed

Wieringa – Chapter 17 – Observational Case Studies

An observational case study is a study of a real-world case without performing an intervention. In all forms of research the researchers tires to minimize its intervention on the case. **The goal is not to acquire knowledge about samples, but about individual cases.**
These case studies are needed to study phenomena that cannot be produced in the laboratory. There are many checklists and handbooks for doing case study research in information systems.

The case study should have
- **Case study protocol**: that contains the description of the context, research problem and design
- **Case study log/ diary**: to collect details of data analysis

The checklist for the research context of observational case study:

1. Knowledge goal(s)
   – What do you want to know? Is this part of an implementation evaluation, a problem investigation, a survey of existing treatments, or a new technology validation?
2. Improvement goal(s)?
   – If there is a higher-level engineering cycle, what is the goal of that cycle?
   – If this is a curiosity-driven project, are there credible application scenarios for the project results?
3. Current knowledge
   – State of the knowledge in published scientific, technical and professional literature?
   – Available expert knowledge?
   – Why is your research needed? Do you want to add anything, e.g. confirm or falsify something?
   – Theoretical framework that you will use?

Context
Each observational case study has a **knowledge goal** and may or may not have an **improvement goal**. If it has an improvement goal, then it is part of an implementation evaluation or problem investigation – otherwise it is a curiosity-driven study that aims to increase our knowledge about a topic.
The checklist asks about current knowledge about this topic.

Research Problem
In order to specify your knowledge questions and population of interest, one needs to define a **conceptual framework**. This should define the architectural structures that the researcher is looking for.

Knowledge questions: prevent you to collect too many data: these questions can be exploratory or can be focused on hypothesis testing (expectations are called propositions). Case study knowledge questions should not only be descriptive but also explanatory – case study are performed in order to gain understanding, not only facts.

4. Conceptual framework

   – Conceptual structures? Architectural structures, statistical structures?
   – Chance models of random variables: Semantics of variables?
   – Validity of the conceptual framework? Clarity of definitions, unambiguous application, avoidance of mono-operation and mono-method bias?

5. Knowledge questions

   – Open (exploratory) or closed (hypothesis-testing) questions?
   – Effect, satisfaction, trade-off or sensitivity questions?
   – Descriptive or explanatory questions?

6. Population

   – Population predicate? What is the architecture of the elements of the population? In which ways are all population elements similar to each other, and dissimilar to other elements?
   – Chance models of random variables: Assumptions about distributions of variables?

A series of case studies is an example of analytical induction → using multiple studies could help improve an architectural theory and get more clarity about the population predicate.

<u>Research Design and Validation</u>

In observational case study research the object is cases as they are found in the real world – the study however has to be designed.

Case selection

This consists in the checklist for selecting a case, that are selected according to some population predicate that specifies architectural properties of the cases in the population of interest.

In case study research the population predicate may not be very clear and the goal of the research may exactly to gather more information about it.

7.1 Acquisition of Objects of Study (cases)

   – How do you know that a selected entity is a case? How do you know it satisfies the population predicate?
   – Validity of OoS
      - *Inference support.* Which inferences would be valid with respect to this design? See checklists for validity of descriptive statistics, abductive and analogic inferences.
      - *Repeatability.* Could other researchers use your report to construct or select a similar OoS?
      - *Ethics.* Are people informed that they will be studied, and do they consent to this? Are they free to stop at any time without giving reasons, and do they know this?

For architectural inference, few consideration are relevant:

- **Analysis**: can enough information be acquired about the case architecture to analyse relevant case mechanisms
- **Variation**: most real-world cases match the population predicate but not completely
- **Abstraction**: the case architecture abstracts away components and mechanisms that may be present in real-world cases

<u>Sampling</u>

Checklist to build a sampling in the case study research

- What is the analytical induction strategy

- Validity of the sampling procedure:
    - Inference support, repeatability and ethics

Measurement Design

Checklist that one must go through when designing measurement in observational case study. Needs first a framework that defines variables and their scales

9. Measurement design

   – Variables and constructs to be measured? Scales, chance models.
   – Data sources? People (e.g. software engineers, maintainers, users, project managers, politically responsible persons, etc.), primary data (e.g. source code, log files, bug tracking data, version management data, email logs), primary documents (e.g., project reports, meeting minutes, organization charts, mission statements), etc.
   – Measurement instruments? Interview protocols, questionnaires, video recorders, sound recorders, clocks, sensors, database queries, log analyzers, etc.
   – What is the measurement schedule? Pretests, posttests? Cross-sectional or longitudinal?
   – How will measured data be stored and managed? Provenance, availability to other researchers?
   – Validity of measurement specification:

     * *Inference support.* Which inferences would be valid with respect to this design? See the applicable parts of the checklists for validity of abductive and analogic inferences.
     * *Repeatability.* Is the measurement specification clear enough so that others could repeat it?
     * *Ethics.* Which company data must be kept confidential? How is privacy of persons respected?

Inference design and Validation

Case-based inference consists of three steps: description, architectural explanation and generalization by analogy.
Descriptive inference from case data may require considerable time and effort because of the amount of data → be sure aout the resources before starting the case study.

10.3 Abductive inference design

   – What possible explanations can you foresee? What data do you need to give those explanations? What theoretical framework?
   – Internal validity

     * *Causal inference*

       - *Sampling influence.* Could the selection mechanism influence the selected cases? Could there be a regression effect?
       - *Measurement influence.* Will measurement influence the case?

     * *Architectural inference*

       - *Analysis:* The analysis of the architecture may not support its conclusions with mathematical certainty. Components fully specified? Interactions fully specified?
       - *Variation:* Do the real-world case components match the architectural components? Do they have the same capabilities? Are all architectural components present in the real-world case?
       - *Abstraction:* Does the architectural model used for explanation omit relevant elements of real-world cases? Are the mechanisms in the architectural model interfered with by other mechanisms, absent from the model but present in the real world case?

     * *Rational inference*

       - *Goals.* An actor may not have the goals assumed by an explanation. Can you get information about the true goals of actors?
       - *Motivation.* A goal may not motivate an actor as much as assumed by an explanation. Can you get information about the true motivations of actors?

10.4 Analogic inference design

- What is the intended scope of your generalization?
- External validity
  * *Object of Study similarity.*
    - *Population predicate.* Does the case satisfy the population predicate? In which way will is it similar to the population elements? In which way will is it dissimilar?
    - *Ambiguity.* Does the case satisfy other population predicates too? What could be the target of analogic generalization?
  * *Representative sampling,* case-based research: In what way will the selected sample of cases be representative of the population?
  * *Treatment.*
    - *Treatment similarity.* Is the specified treatment in the experiment similar to treatments in the population?
    - *Compliance.* Is the treatment implemented as specified?
    - *Treatment control.* What other factors than the treatment could influence the OoS's? Could the implemented treatment be interpreted as another treatment?
  * *Measurement.*
    - *Construct validity.* Are the definitions of constructs to be measured valid? Clarity of definitions, unambiguous application, avoidance of mono-operation and mono-method bias?
    - *Measurement instrument validity.* Do the measurement instruments measure what you claim that they measure?
    - *Construct levels.* Will the measured range of values be representative of the population range of values?

Wieringa – Chapter 19 – Technical Action Research

TAR is the use of an experimental artifact to help a client and to learn about its effects **in practice**. The artifact is experimental (under development); TAR is used to validate the artifact in the field

Example: a researcher may have developed a new effort estimation technique and is now ready to test it in the field. She teaches it to project managers, who then use it in their next project. The researcher observes what happens in order to answer knowledge questions about the technique.

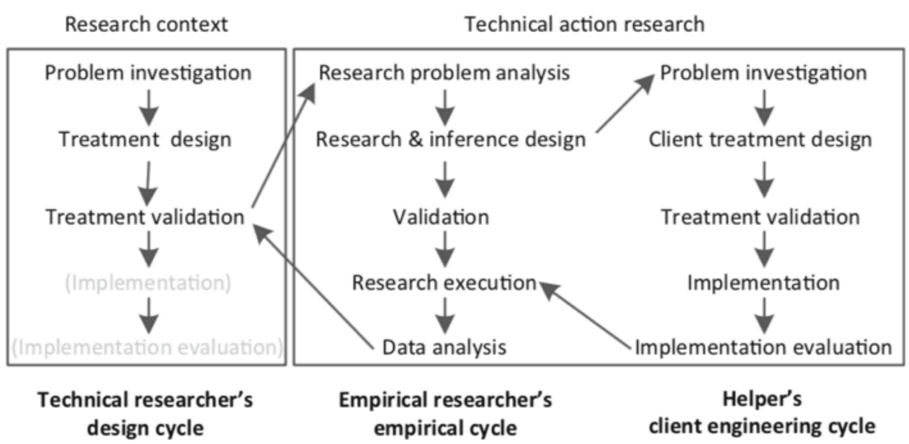
Tar are single-case studies because the artifact is studied as a case.
The difference with observational studies is that the researcher intervenes in the case to see what happens
The difference with single-case mechanism experiment is that he treatment is not merely applied to see what happens but to help the client.

TAR:

- Is artifact-driven
- Single-case studies → each use of the artifact is studied as a case
- Work with the client to solve a problem without the goal of testing a particular artifact

The researcher in TAR:

- **As Technical researcher** designs a treatment to solve a class of problems: design new technique
- **As Empirical researcher** answers some validation knowledge questions about the treatment. E.g. how accurate is the technique
- **As Helper** applies the treatment version to the client need in order to help him



⇒ *In the structure of the TAR, we start with a design cycle to design treatments for a problem -> perform an empirical cycle to validate a treatment (and therefore answer empirical knowledge) → validate the treatment by using it to help a client with the engineering client cycle.*

<u>Context</u>

- Answering the questions below
  - Knowledge goals: what do you want to know about the treatment? **Validate a treatment under conditions of practice**
  - Improvement goals: what is the goal of the treatment? **The goal of the treatment to be designed**
  - Current knowledge

<u>Research Problem</u>

- Conceptual framework: structures and architectural structures?

- Knowledge questions: open (exploratory) or closed questions?
- Population: population predicate?

Knowledge questions are validation questions:
- What effects are produced by the interaction between artifact and context?
- Do the effects satisfy requirements?

<u>Research design and validation</u>
- Acquisition of objects of study: how to acquire a clinet? How do you customize the artifact? Validity of the Oos

<u>Inference design and validation</u>
Inferences from TAR studies must be planes as carefully as inferences elsewhere: it is case based and consists of description, explanation and generalization.

Causal inferences: are **not supported** by a TAR study, because you cannot observe the alternative world

Descriptive inference: are supported, but maybe not give you all the variables you need
- Summarize all the data you collected, diary + work products
- Maybe also have a blind peer review
- Validity: support for data preparation (will prepared data represent the unprepared data?); support for data interpretation(will the interpretations be facts in your conceptual research framework? As in, will other people come to the same conclusions?); support for descriptive statistics (is the chance model of the variables of interest defined in terms of the population elements?)

Architectural inference: analyse the mechanism of the client's architectural organization;

Wohlin – Chapter 2 – Empirical Strategies

There are two types of research paradigms that have different approaches to empirical studies. They should be regarded as complementary instead than competitive.

1. ***Exploratory research*** is concerned with studying objects in their natural setting and letting the findings emerge from the observations. This implies a Flexible research design, also qualitative research. Inductive research attempts to interprets a phenomenon based on explanation that people bring forward
2. ***Explanatory research*** is concerned with quantifying a relationship or compare two or more groups with the aim to identify a cause-effect relationship. This is called fixed design study and implies that factors are fixed before the study is launched. Also called quantitative research.

<u>Overview of empirical strategies</u>

3 major types of investigations:

Strategy	Design type	Qualitative/quantitative
Survey	Fixed	Both
Case study	Flexible	Both
Experiment	Fixed	Quantitative

1. **Survey:** a system for **collection information from or about people to describe their knowledge attitudes** and behaviours
   a. *Description:* Can be seen as **snapshot of the situation to capture the current status**. Could be opinion pools or market research. A questionnaire is constructed to obtain information for the research. The information is collected and then aggregated into a from that can be quantitative or qualitative
   b. *Characteristics:* never **conducted to create an understanding of the sample**. Survey have the ability to **provide a lot of variables to evaluate**. They are often needed to **provide broad overviews →** questions in several fields
   c. *Purposes:* 1) ***Descriptive surveys*** can be conducted to enable assertions about some population This could be determining the distribution of certain characteristics or attributes. 2) ***Explanatory surveys*** aim at making explanatory claims about the population. 3) ***Explorative surveys*** are used as a pre-study to a more thorough investigation to assure that important issues are not foreseen.
   d. *Data collection:* paper in person or online. In person = more answers and less “do not know”

2. **Case study:** an empirical enquiry that draws on multiple sources of evidence to investigate one instance of a phenomenon within its real context. **It is often conducted to investigate a single entity or phenomenon**. If we would like to compare two methods the study may be defined **as a case study or experiment**. Case studies are **suitable for industrial evaluation of SE methods** because they **avoid scale-up problems**. Difference with experiments is that experiments sample over the variables that are being manipulated, case studies select from the variables representing the typical situation. Easier to plan and more realistic but difficult to generalize results. If the effect of a process change is very widespread a case study is more suitable .
   a. *Arrangements:* **three ways: 1)** comparison of the results of using the new method against a company baseline**. 2)** a sister project can be chosen as a baseline. The new project uses new method, sister uses old. **3)** Apply method at random to some components and not to other.
   b. *Confounding factors:* **necessary to minimize the effects of confounding factors** – that make it impossible to distinguish the effects of the two factors from each other. Some problems: **not full control of case study situation**, increase in scale lead to changes in the type of problems that become most indicative.

3. **Experiment:** empirical enquiry that **manipulates one factor or variable of the studied setting**. Based on randomization, different treatments are applied to or by different subjects. It is launched when we want control over the situation and want to manipulate behaviour directly, precisely and systematically. The design should be made so that the object involved represent all the methods we are interested in.
   a. They may be *human-oriented* or *technology-oriented.*
      i. *Human oriented:* humans apply different treatments to objects
      ii. *Technology oriented:* humans apply different tools to objects
   b. Carrying out an experiment involves several different steps.
      i. Scoping
      ii. Planning
      iii. Operation
      iv. Analysis and interpretation
      v. Presentation and package

4. **Quasi experiment:** is an empirical enquiry **similar to an experiment where the assignment of treatments cannot be random** *but depends on the characteristics of the subjects or object themselves.*

Empirical strategies comparison

The prerequisites for an investigation limit the choice of research strategy. A comparison of strategies can be based on a number of different factors.

- **Execution control** describes how much control the researcher has over the study. For example, in a case study, data is collected during the execution of a project.
- **Measurement control** is the degree to which the researcher can decide upon which measures to be collected, and to include or exclude during execution of the study. An example is how to collect data about requirement volatility
- **Investigation cost.** Depending on which strategy is chosen, the cost differs. This is related to, for example, the size of the investigation and the need for resources. The strategy with the lowest cost is the survey, since it does not require a large amount of resources.

Factor	Survey	Case study	Experiment
Execution control	No	No	Yes
Measurement control	No	Yes	Yes
Investigation cost	Low	Medium	High
Ease of replication	High	Low	High

Another aspect to consider is the possibility to replicate the investigation.

Replications

The replication involves repeating the investigation under similar conditions but varying the subject population. Replication may be of 2 different types:

1. **Close replication**: follow the original procedure as closely as possible – exact replication. The main advantage is building confidence in the outcome, but often it requires the same researchers to conduct the study.
2. **Differentiated replications**: study the same research questions using different experimental procedures. It may be used for more exploratory studies.

Theory in SE

"A theory provides explanations and understanding in terms of basic concepts and underlying mechanisms, which constitute an important counterpart to knowledge of passing trends and their manifestation". There are 5 different types of theories:

1. **Analysis***:* Theories of this type describe the object of study, and include, for example, taxonomies, classifications and ontologies.
2. **Explanation***:* This type of theories explains something, for example, why something happens.

3. **Prediction***:* These theories aim at predicting what will happen, for example, in terms of mathematical or probabilistic models.
4. **Explanation and prediction***:* These theories combine types 2 and 3, and is typically what is denoted an "empirically-based theory".
5. **Design and action***:* Theories that describe how to do things, typically prescriptive in the form of design science.

Constructs are the **entities in which the theories are expressed**. **Propositions** are made up from **proposed relationships between the constructs**. The explanations originate from logical reasoning or empirical observations of the propositions. The scope of the theory defines the circumstances under which the theory is assumed to be applicable.

<u>Aggregating evidence from Empirical Studies</u>

As the # of empirical studies grow aggregating evidence from multiple studies appear replication studies.

1. The research should build upon each other so to take into consideration existing knowledge
2. Empirical studies may give an answer together but not separately, the collection must meet academic standards though

Systematic literature reviews are meant to collect and synthesize empirical evidence from different sources.

- **Primary studies**: empirical studies searched for
- **Secondary study**: systematic literature review
  - The search can be done using database queries. Snowballing means going back to the original sources of the paper found.
  - If research question is general then there may be the need of a mapping study.
- When a set of empirical studies is collected on a topic the synthesis or aggregation takes place and it is based on Statistical methods called *meta-analysis*

<u>Evaluation of process changes</u>

Objective is to discuss the strategies in terms of a suitable way of handling technology transfer from research to industrial use. The objective is to order the studies base on how they may be conducted to enable a controlled way of transferring research into practice

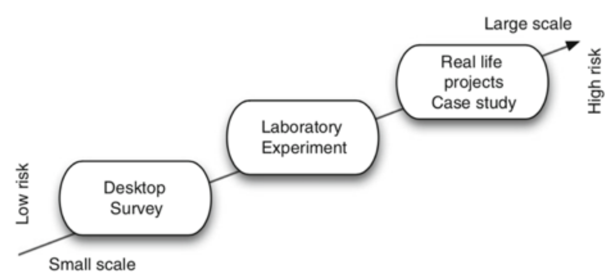


Fig. 2.1 Surveys, experiments and case studies

- Desktop: off-line evaluation without executing the changed process
- Laboratory: change evaluated in off-line laboratory setting In which an experiment is conducted
- Real life: real life evaluation of the change in a development situation

Quality Improvement Paradigm

Is a general improvement scheme for the software business. Based on 6 steps:

1. **Characterize**. Understand the environment based upon available models, data, intuition, etc. Establish baselines with the existing business processes in the organization and characterize their criticality.
2. **Set goals**. On the basis of the initial characterization and of the capabilities that have a strategic relevance to the organization, set quantifiable goals for successful project and organization performance and improvement.
3. **Choose process**. On the basis of the characterization of the environment and the goals that have been set, choose the appropriate processes for improvement, and supporting methods and tools, making sure that they are consistent with the goals that have been set.

4. **Execute**. Perform the product development and provide project feedback based upon the data on goal achievements that are being collected.
5. **Analyse**. At the end of each specific project, analyse the data and the information gathered to evaluate the current practices, determine problems, record findings, and make recommendations for future project improvements.
6. **Package**. Consolidate the experience gained in the form of new, or updated and refined, models and other forms of structured knowledge gained from this and prior projects.

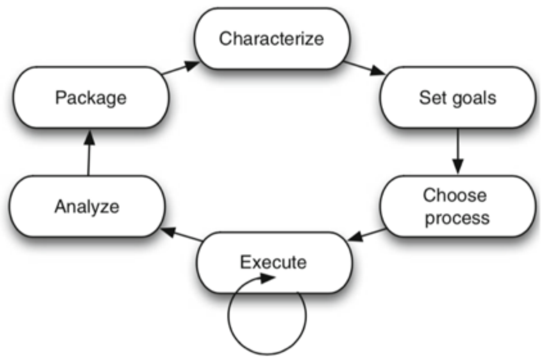
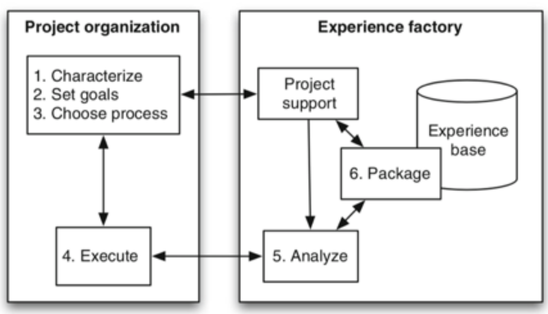


Fig. 2.2 The six steps of the Quality Improvement Paradigm [7]

There are 2 feedback cycles: project and corporate.

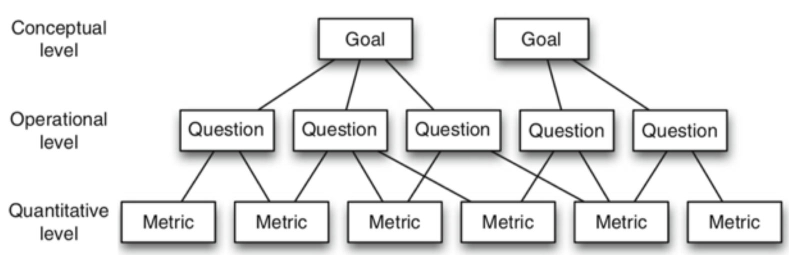
The QIP is based on the **improvement of software development** requires **continuous learning**, that must be packaged into experience models that can be effectively understood and modified. These are stored in a repository called **experience base**



QIP focuses on a logical separation of project development from the systematic learning and packaging of reusable experience. The goal of the Experience Base is to produce and maintain software. The factory processes the information and returns direct feedback to each project with goals and models tailored from similar project.

The QIP uses a **goal/question/metrics method** based on the assumption that to measure the creation in a purposeful way it must:

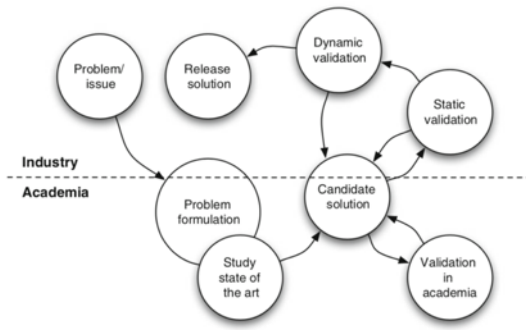
1. Specify goals
2. Trace goals to the data
3. Provide a framework to interpreting the data



The resulting measurement model has three levels:

1. **Conceptual level** (Goal). A goal is defined for an object, for a variety of reasons, with respect to various models of quality, from various points of view, relative to a particular environment.
2. **Operational level** (Question). A set of questions is used to characterize the way the assessment/achievement of a specific goal is going to be performed based on some characterization model. Questions try to characterize the objects of measurement (product, process and resource) with respect to a selected quality aspect and to determine its quality from the selected viewpoint.
3. **Quantitative level** (Metric). A set of data is associated with every question in order to answer it in a quantitative way (either objectively or subjectively).

The following model was create in order to bring a solution to a real industrial problems using empirical methods and apply it in practice.



<u>Ethics in Experimentation</u>

There are 4 key principles:

- Explicit consent to the participation
- Scientific value of the study
- Confidentiality of the data
- Beneficence must overweight the risks

Wohlin – Chapter 4 – Systematic Literature Review

Also Wholin says: Planning, conducting and reporting are the most important steps of any systematic review.
A systematic review is used to "identify analyse and interpret all available evidence related to a specific research question".

<u>Planning the review</u>

- **Identification of the need for a review**: the need originates from a researcher's aim to understand the state-of-the art in an area or from practitioners wanting to use empirical evidence in their strategic decision
- **Specify the research question**: the area of the systematic review and the specific research question set the focus for the identification of the primary studies, the extraction of the data and the analysis. Aspects to take inot account in the research question are:
    - Population in which evidence is collected
    - Intervention applied in the study
    - Comparison to which the intervention is compared
    - Outcomes of the experiment
    - Context of the study
    - Experimental designs

Develop a review protocol to define the procedure for the SLR.

<u>Conducting the review</u>

This means setting the review protocol into practice

- **Identify the research:**
    - Specify search strings and apply them to the database (snowballing is systematically searching for primary studies based on references from other studies)
    - Search is always a trade-off between finding all relevant studies and not getting too many false positives
    - Remember that positive results are more likely to be published than negative results → search for "grey" literature
- **Select primary studies:**
    - Formulates the criteria to select primary studies beforehand, otherwise you'll be biased in choosing "less" primary studies or easier ones, etc.
- **Study quality assessment:**
    - No universally agreed definition, but use checklists
    - It could help exclude some primary studies
- **Data extraction and monitoring:**
    - Extract the data from primary studies
    - Data extraction is based on the research question – log everything you do
    - If a study is published in more than one paper, use only one instance of it
- **Data synthesis:**
    - This is the most advanced form of meta-analysis – to analyse the outcome of several independent studies
        - It compares the effect size and p value to assess the outcome of multiple studies
    - It is applicable to replicated experiment
    - Involves three main steps:
        - Decide which studies to include
        - Extract the effect size
        - Combine the effect sizes
    - It should also include an analysis of the publication bias – plotting a funnel and seeing where are the gaps in the values = studies are not published

- The effect size here is an indicator that usually is the difference between the mean values of each treatment – usually is normalized
- If homogeneity is assumed: use fixed effects models
- If homogeneity is not assumed: use random effect model – allow for variability in the factors due to unknown factors
- Less formal methods are the descriptive or narrative synthesis, forest plots or mixing these methods about (thematic, comparative analysis , line of argument, case survey, scoping analysis, …)

<u>Reporting the review</u>

- Self explanatory, depends on the audience – ideally include changes in the study protocol, lists of primary studies, data etc.

<u>Mapping study</u>

- Also referred as scoping, searches a broader field for nay kind of research to get an overview of the state-of.art on a topic

Wohlin – Chapter 5 – Case studies

The term case study is used in parallel with terms such as field study and observational study – these focus on particular aspect of the research methodology.
The case study methodology is suited for many kinds of SE as the object of study are contemporary phenomena, hard to study in isolation/ controlled experiment.
Case studies do not generate the same results on causal relationship but provide deeper understanding of the phenomena under study in its real context.

Case studies in its context

There are three definitions by Robson, Tun and Benbasat
1. case study is an empirical method aimed at investigating contemporary phenomena in their context
2. Case study is a research strategy and uses multiple sources of evidence
3. Case study is an inquiry where the boundary between the phenomenon and its context may be unclear, there is lack of experimental control and there is information gathering from few entities

- Action research is closely related to case study as it aims to influence or change some aspect of the focus of the research.
- Case study is purely observation, while action research is focused and involves the change process

A case study may contain elements of other research methods:
- A survey may be conducted
- A literature search precede a case study
- Ethnographic methods are used for data collection in case studies

A case study also :
- Copes with the situation in which there will be many more variables than data points
- Relies on multiple sources
- Benefits from the prior development of theoretical propositions to guide data collection and analysis

Key characteristics:
- It is flexible – copes with the complex and dynamic characteristics of real world phenomena
- Conclusions are based on a chain of evidences – qualitative and quantitative
- Adds to existing knowledge by basing itself on a previous theory

Why case study in SE?
- Se is a multidisciplinary discipline that involves areas in which the research questions often are suitable for case study research
- Case studies also often do not need a strict boundary between the studies object and its environment

Case study research process
- There are 5 major steps:
  - Case study design: define the objectives and plan the case study
  - Preparation for data collection
  - Collection of data
  - Analysis of collected data

- o Reporting

Design and Planning

Often the research is flexible, but planning is still needed

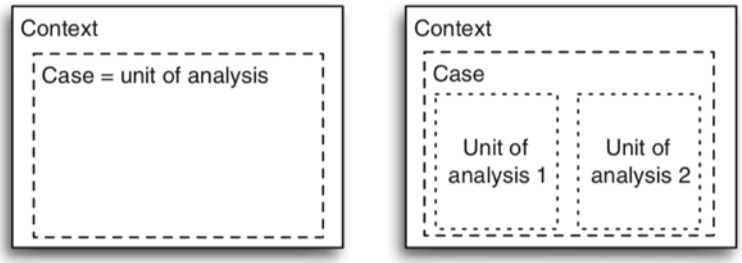


Fig. 5.1 Holistic case study (*left*) and embedded case study (*right*)

The plan should include:

- Objective: what to achieve?
  - o Is more generally formulated than in fixed design research
  - o More like a focus point
- The case: what is studied?
- Theory: frame of reference
- Research questions: what to know?
- Methods: how to collect data?

  Selection strategy: where to seek data

Holistic case studies: case is studied as a whole (2 companies, 2 project: is holistic if the context is software companies in general)

Embedded case studies: multiple units of analysis are studied within a case (example above is embedded if the context is the specific company or application domain)

Case study protocol

- Used to decide the procedures of the study, is continuously changed
- Serves as a guide for data collection
- Makes the research concrete in the planning phase
- Serves as a log or diary for data collection and analysis

Table 5.1 Outline of case study protocol according to Brereton et al. [25]

Section	Content
Background	Previous research, main and additional research questions
Design	Single or multiple case, embedded or holistic design; object of study; propositions derived from research questions
Selection	Criteria for case selection
Procedures and roles	Field procedures; Roles for research team members
Data collection	Identify data, define collection plan and data storage
Analysis	Criteria for interpretation, linking between data and research questions, alternative explanations
Plan validity	Tactics to reduce threats to validity
Study limitations	Specify remaining validity issues
Reporting	Target audience
Schedule	Estimates for the major steps
Appendices	Any detailed information

Preparation and collection of data

Several different sources of information in a case study to limit the effect of interpretation of one single data source

- Conclusions are drawn by analysing differences between data sources

Three levels, in order of cost to apply:

- **First**: direct methods (interviews, focus groups, etc) – researcher can control what data is collected
- **Second**: indirect methods (researcher directly collects raw data without interacting with the subjects during the data collection) – researcher can control what data is collected
- **Third**: independent analysis of work artifacts where already available and sometimes compiled data is used – less expensive but no control of quality

Interviews:

- can have open questions or closed (limited alternative)
- Unstructured: interview questions are formulated as general concerns and interests from the researcher
- All questions are planned in advanced
  - Similar to questionnaire
- Semi-structured: planned but not following precise order

Table 5.2 Overview of interview types

	Unstructured	Semi-structured	Fully structured
Typical foci	How individuals qualitatively experience the phenomenon	How individuals qualitatively and quantitatively experience the phenomenon	Researcher seeks to find relations between constructs
Interview questions	Interview guide with areas to focus on	Mix of open and closed questions	Closed questions
Objective	Exploratory	Descriptive and explanatory	Descriptive and explanatory

Observations:

Table 5.3 Different approaches to observations

	High awareness of being observed	Low awareness of being observed
High degree of interaction by the researcher	Category 1	Category 2
Low degree of interaction by the researcher	Category 3	Category 4

Alternatives:

- Observe the work of someone and analyse
- Ask to "think aloud"

Archival Data

- Meeting minutes
- Documents from development phases, …
- Is a third degree of data that can be collected in a case study.
- Management tool is an important source

Data Analysis

Both for quantitative and qualitative data

Quantitative: usual stuff

Qualitative data analysis

- Used to derive conclusions from the data (obvious)

- Keeps a clear chain of evidence in order to be able to follow the derivation of results

General technique for analysis:

- Hypothesis generation is to find hypotheses from the data
- Hypothesis confirmation is sued to confirm that an hypothesis Is really true (triangulation and replication are examples)
- Negative case analysis (tries to find alternative explanations that reject the hypotheses)
- Hypothesis generation may take place within one cycle of a case study or with data from one unit of analysis

The analysis of qualitative data is conducted in a series of steps:

- Data is coded
    - Code cas be assigned in hierarchy of codes and sub codes
- Coded material is combined with comments and reflections by the researcher
- Once done, the researcher can go thorugh the material to identify a set of hypotheses

<u>Level of formalism</u>

- The analysis can be conducted at different levels of formalism:
    - **Immersion approaches**: least structured, rely on intuition and interpretive skills – hard to combine with a chain of evidences
    - **Editing approaches**: few a priori codes included – based on findings of the researcher
    - **Template approaches**: more formal approaches and include more a priori based on research questions
    - **Quasi-statistical approaches**: much formalized and include calculation of frequencies of words and phrases

<u>Validity</u>

- It denotes trustworthiness of results and to what extent the results are true and not biased
- There are different ways to classify aspects of validity
    - **Construct validity**: reflect to what extent the operational measures that are studied really represent what the research has in minds. E.g. the constructs discussed in the interview questions are not interpreted in the same way by the researcher and the interviewed persons
    - **Internal validity**: concern the examination of causal relations. There is the risk that the factor that the researcher is investigating is affected by a third factor – if the researcher is not aware: threat
    - **External validity**: concerned with to what extent it is possible to generalize the findings and to what extent the findings are of interest to other people outside the research → relevance for other cases. The intention is to generalise the results, however in case study this ability is not present: the intention is to enable analytical generalization where the results are extended to cases with common characteristics
    - **Reliability**: this aspect is concerned with the dependency of data analysis on the specific researcher. The results should be the same between studies

<u>Reporting</u>

The report communicates the findings of the study and is the main source of information to judge its quality

- Characteristics: the study should say what it was about
    - Communicate a clear sense of the studied case
        - Not too detailed because data could be sensitive
    - Provide a history of the inquiry
    - Provide basic data in focused form
    - Articulate the conclusions

Table 5.4 Proposed reporting structure for case studies based on Jedlitschka and Pfahl [86] and adaptations to case study reporting according to Runeson et al. [146]

Section headings	Subsections
Title	
Authorship	
Structured abstract	
Introduction	Problem statement Research objectives Context
Related work	Earlier studies Theory
Case study design	Research questions Case and subject selection Data collection procedure(s) Analysis procedure(s) Validity procedure(s)
Results	Case and subject descriptions, covering execution, analysis and interpretation issues Subsections, which may be structured e.g. according to coding scheme, each linking observations to conclusions Evaluation of validity
Conclusions and future work	Summary of findings Relation to existing evidence Impact / implications Limitations Future work
Acknowledgements	
References	
Appendices	

Wohlin – Chapter 7 – Scoping

An experiment is a labor-intensive task. Therefore need a determined foundation in order to ensure that the experiment can be fulfilled.

Scope Experiment

The scope of the experiment is set by defining its goals. The purpose of a goal definition template is to ensure the ascpects of an experiment before the execution phase takes place.

Template:

Analyze <object> - can be products, processes, resources, models, metrics or theories
For the purpose of <purpose> - intention of the experiment “evaluate the impact of two different techniques”
With respect to their <quality focus> - primary effect under study “effectiveness, cost, reliability, ecc”
From the point of view of the <perspective> - from which the results are interpreted
In the context of <Context> - environment in which the experiment is run

Table 7.3 Goal definition framework

Object of study	Purpose	Quality focus	Perspective	Context
Product	Characterize	Effectiveness	Developer	Subjects
Process	Monitor	Cost	Modifier	Objects
Model	Evaluate	Reliability	Maintainer	
Metric	Predict	Maintainability	Project manager	
Theory	Control	Portability	Corporate manager	
	Change		Customer	
			User	
			Researcher	

Wohlin – Chapter 8 – Planning

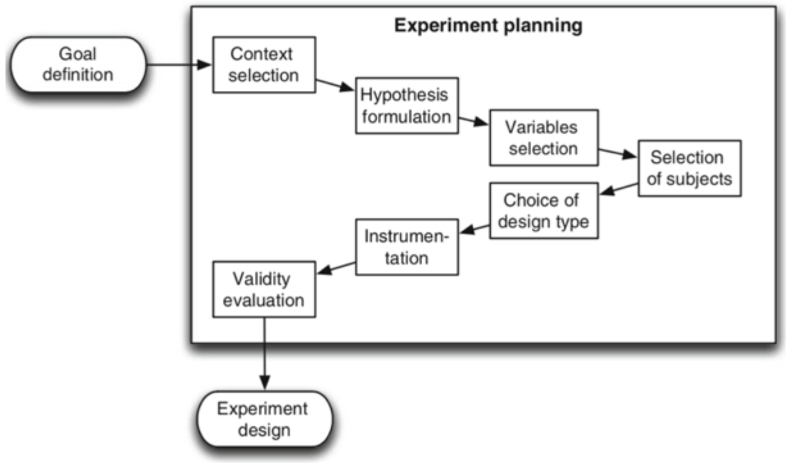
While the scoping determines the foundation of the experiment – why the experiment is conducted – the planning prepares how the experiment is conducted

The planning phase can be divided into seven steps:

1. Context selection
2. Hypothesis formulation
3. Variable selection
4. Selection of subjects
5. Experiment design type
6. Instrumentation
7. Validity evaluation

Context selection

The best option would be to execute experiment in the real world, as this is not often the case, a new method must be examined. Alternatives can be running off-line projects in parallel with the real projects, reducing the risk but costing more.



Overall the context can be characterize according to 4 dimensions:

- Off-live vs on-line
- Student vs. professional
- Toy vs. Real problems
- Specific vs. general

A common situation is that something existing is compared to something new, e.g. a method.

Hypothesis formulation

The basis for the statistical analysis of an experiment is hypothesis testing. A hypothesis **is stated formally and the data collected during the experiment is used to, if possible, reject the hypothesis.**

2 hypothesis have to fomulated:

- **Ho, states that there are no real underlying trends or patterns in the experiment settings: the only reasons for differences in our observations are coincidental**.
    - E.g. the new tracing method is not, on average, considered better than no method.
- Alternative hypothesis: in favor of which the null is rejected.
    - The method helps students be better in detecting changes

Variable selection

Independent: those we can control and change in the experiment
Dependent: often one dependent variable, not directly measurable.

Selection of subjects

- Simple random sampling
- Systematic sampling – *the first subject is selected form the list of the population*
- Stratified random sampling: population divided into a number of groups – *random sampling is applied within the strata.*

Non-probability sampling

- Convenience sampling: the nearest and most convenient persons are selected as subjects
- Quota sampling: used to get subjects from various elements of a population

Experiment design

An experiment consists of a series of tests of the treatments, to get the most out of the experiment, this series of tests must be planned and designed. A design of an experiment describes how the tests are organized and run

An experiment is a set of tests

General design principles are:

- **Randomization:** All statistical methods used for analyzing the data require that the observations be from independent random variables. To meet this requirement, randomization is used.
- **Blocking:** often we have to block out the effect of a factor which does not interest us in the analysis. This technique increases the precision of the experiment.
- **Balancing:** assigning the treatment so that each treatment has an equal number of subjects. This simplifies and strengthens the statistical analysis of the data, but it is **not** necessary.

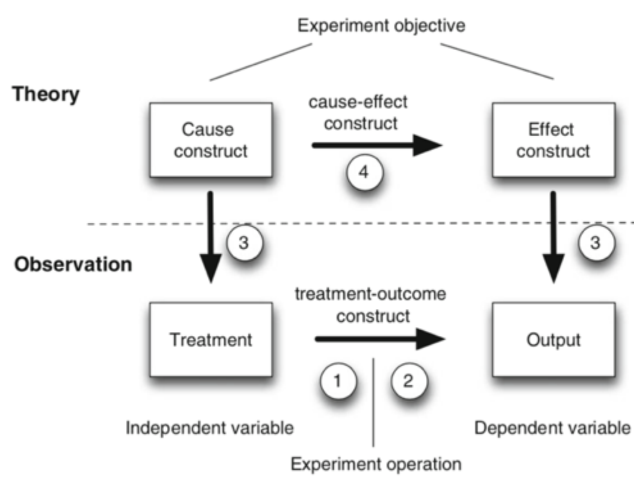
Standard Design types

- **One factor with two treatments**
  - *Completely randomized design:* we want to compare the two treatments against each other. The most common is to compare the mean of the dependent variable for each treatment.
    Example of hypothesis is: $\boldsymbol{H_0}: \mu_1 = \mu_2\ vs. \boldsymbol{H_1}: \mu_1 \neq \mu_2, \mu_1 < \mu_2\ or\ \mu_1 > \mu_2,$
  - *Paired comparison design:* we can make comparison within matched pairs of experiment material – here each subject uses both treatments on the same object. This is called also crossover design. There is random order to each subject [PAPER ?] and the comparison for the experiment can be to see if the difference between the paired measures is zero
    Example of hypothesis is: $\delta_j = y_{\{1j\}} - y_2\ and\ \mu_d is\ the\ mean\ of\ the\ difference$
    $\boldsymbol{H_0}: \mu_d = 0\ vs. \boldsymbol{H}_1: \mu_d \neq 0, \mu_d < 0\ or\ \mu_d > 0$
- **One factor with more than two treatments:**
- **Two factors with two treatments:**
- **More than two factors each with two treatments:**

<u>Validity</u>

How valid are the results? Adequate validity means that the result should be valid for the population of interest. At least they should be valid for the population from which the sample used in the experiment was drawn. Then we could generalize to a broader population. Valid results are generalizable results

Validity threats



- **1. Conclusions**: concerned with the relationship between the treatment and outome: at a given significance we want a statistical relationship
  - E.g. Low statistical power, violated assumption of statistical tests; fihing and error rate; reliability o fmeasure and treatment implementation; random irrelevancies in settings
- **2. Internal**: relationship between the treatment and the outcome, we want a causal relationship and not a random confounder relationship. We want the treatment to cause the outcome
  - **E**.g. History, maturation (effect of passing of time), testing instrumentation (badly designed), selection, mortality, ambiguity of direction of causal influence; multiple groups (different groups and interaction), social threats
- **3. Construct**: validity of the relation between theory and observation –if the relationship is causal we must ensure 2 things: the treatment reflect the construct of the cause and that the outcome reflects the construct of the effect well

    - **E**.g. Design threats; inadequate explication of constructs (not defined, theory not clear, etc); mono-operation bias (if single independent variable – experiment may under-represent the construct); confounding construct and levels, interaction of testing and treatment
- **4. External**: concerned with the generalization (theory) if all above is ok, can we generalize the study outside the scope of our study?
    - **E**.g. Interaction of selection and treatment; interaction of setting and treatment;

Wohlin – Chapter 9 – Operation

There are three steps:
- Preparation
- Execution
- Data validation

<u>Preparation</u>
- Commit participants
    - Obtain their consent – otherwise data is invalid
    - If Sensitive results – assure that the experiment will be kept confidential
    - Inducements: offer some kind of reward, with low value otherwise people will be motivated only because ofit
    - Disclosure: reveal all details of the experiment as openly as possible
- Instrumentation concerns:
    - Stay as anonymous as possible in the forms one gives out to participants
    - Do not have unclear instrumentation – risk on the experiment's results
    - If randomized design – create multiple set of ijnstruments

<u>Execution</u>
- The simpler the better and easier to collect data
- Data collection:
    - Manually or by participants – a drawback if the researcher does not collect data is that there is no possibility for the experimenter to directly reveal inconsistencies, uncertainties and flaws in the form
- Experimental environment:
    - The experiment should not affect the project in which is being doing, if it is a development project. Otherwise the experiment effect could modify the environment of the development project
    - Reduce the interaction between development project and experiment
- Data validation:
    - Experimenter must chelck that the data has been collected correctly
    - Review that the experiment has actually been conducted in the way that was intended
    - People may not have participated seriously in the experiemnt

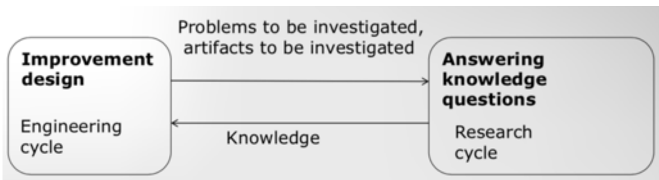
<u>Operation</u>
- Preparation: identify all the subjects. Follow the preparation steps
- Execution: ensure that all the subject can do the experiment at the same time
- Data validation

Wohlin – Chapter 11, deals with presentation of the paper

Lecture – Part I, L2 – Introduction to Design Science

- Explore the characteristics of the design science framework
- Specify research goals and research questions
- Explore basic characteristics of experimental design

Design Science
- Is the design and investigation of artifacts in contex
- In the context of the ARM course we are going to conduct a knowledge-oriented research project, motivated by a design problem

The design cycle
- 

- Treatment are designed and the design (deciding what to do) is specified (documenting the decision)
- Implementation = introducing an artifact in the intended problem context
- **Validation:** justify that it would contribute to the goals of the stakeholders before transferring it to practice (mainly is in a simulation context- to predict effects)
- **Evaluation (of implementation):** investigating whether an implementation has contributed to the stakeholder goals (mainly in real-world research)

Main points of the design cycle:
- Engineering cycle is a rational decision cycle:
  - Problem/ evaluation -> look where you are and what you want to do
  - Design possible treatments
  - Validate treatments without executing them
  - Choose one and implement -> evaluate
- The design cycle is the preparation for action

(controlled) Experiment
- Numerical evaluation that manipulates one factor or variable of the setting
- Based in randomization, different treatments are applied → keep other variables contstnt and measure the effect on outcome variables
- Done in laboratory environment

Quasi experiment
- Assignment of treatments to subject is not based on randomization but emerges from the characteristics of the subject or object themselves
- Statistical inference methods are used to show with significance that one method is better than the other

Strategy	Design type	Qualitative/Quantitative
Survey	Fixed	Both
Case Study	Flexible	Both
Experiments	Fixed	Mainly quantitative

<u>Experiment</u>

Advantages:

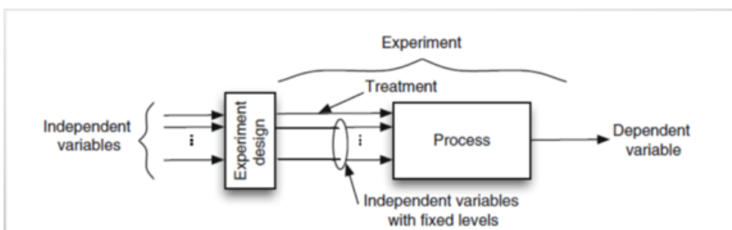
- Control of subject, object and tools (instruments)
- Possible to draw general conclusions
- Possible to perform statistical tests (hypothesis testing)
- Possible to replicate the study in different settings

Process of an experiment

1. Experiment idea
2. Scoping
3. Planning
4. Operation
5. Analysis and interpretation
6. Presentation & package
7. Report

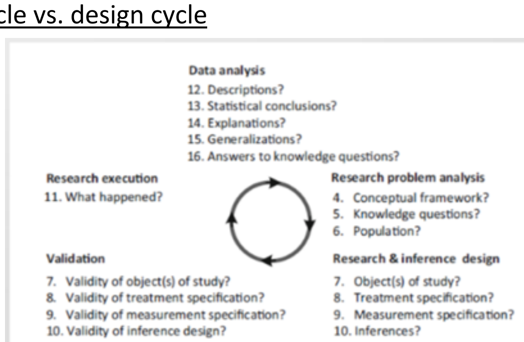
<u>Variables</u>

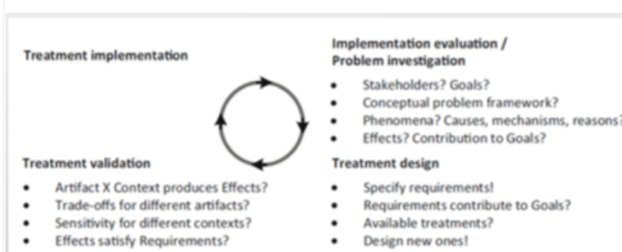
- **Dependent**: variables that we want to study and see the effect of the changes in the independent variable (y)
- **Independent variables**: manipulated variables and controlled (x, x_1, …)
    - They are also called **factors**, in this context a treatment is only **one value** of a factor



- From picture above it is visible that the treatment is only 1 independent variable that will be modifies to see its effect on the dependent variable -> this treatment can take multiple values and those will be tested in their influence in modifying the dependent variable

<u>Empirical cycle vs. design cycle</u>





Design problems: call for a chage in the world (solution is design, many solutions, evaluated by utilty)

Knowledge questions: ask for knowledge about the world

Goal structure: of the research should be composed in part by motivation of the researcher and in part by the design of the research goal – so, part "social context" and part "design research"

Template to design any problem

Improve <problem context>
by <treating it with a (re)designed artifact> such that <artifact requirements>
in order to <stakeholder goals>

Knowledge questions may be analyitics or empirical (descriptive or explanatory, open or closed, effect-realted or requirement related)

What is a model?

- An analogic model is an entity that represents entities of interest → called its targets in order to answer questions by studying the model

Lecture – Part 1 - L7 – Sample Based experiments

Sample-based experiments vs case-based

- Sample based:
  - Typical in domains with high variability
  - Also used in engineering when validating artefacts with a concrete sample
- Case-based:
  - Typical in engineering domains
  - High reproducibility
  - Also used in social sciences to study many variables in one case

Comparative experiment

- Evaluate benefits and drawbacks of using A and B.
- Experimental tasks: used to compare the subjects' performance and perception when they use the discovery process and the guidelines

(1) Object of study

- Can be selected and/or constructed
- It is part of the world that the researcher interacts with in order to learn something about population elements
- It is an entity where the phenomena occur from which measurements are taken

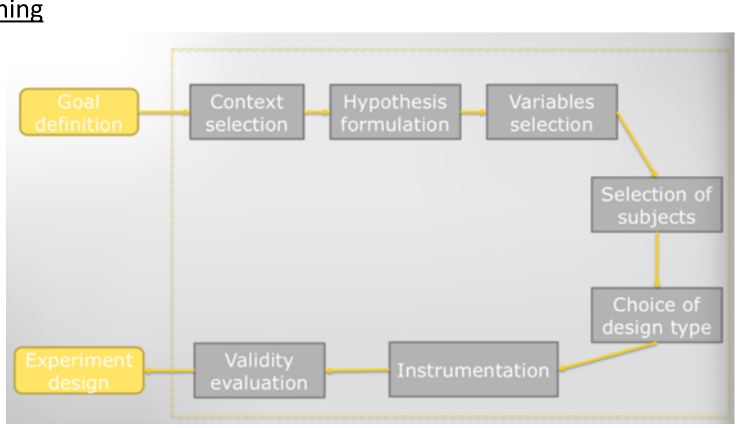
(2) Treatment specification

- An experimental treatment is a treatment of an object of study performed with the goal of finding out what the effects of the treatment are
- It is the exposure of an object of study to a treatment in order to answer a knowledge question of the researcher

(3) Measurement specification

- It is the assignment of a value to the phenomenon denoted by the variable
- One has to define a measurement scale with values and manipulation rules that have real-world meaning

(4) Planning

- 

- **Context selection**: Where do we want our experiment to run?
- **Hypothesis**: H_0 : no underlying trends or patterns in the experiment settings -> you want to reject it
  - H_1:
- **Selection of subjects:** probability (random sampling, systematic sampling, stratified sampling); non-probability (convenience sampling, quota sampling)
- **Choice of design type:** focus on independent variables.
  - One factor with two treatments: compare two treatment against each other.

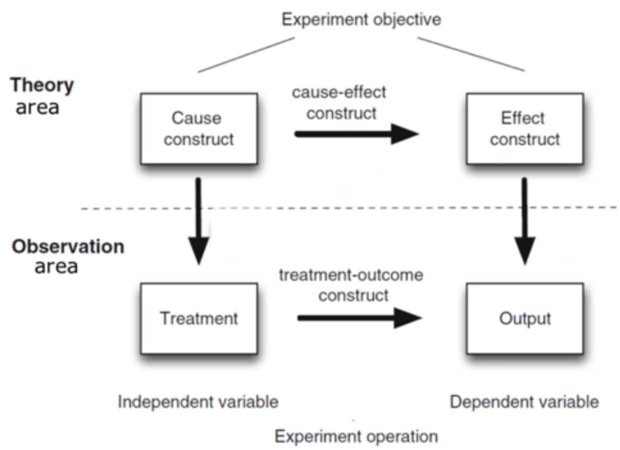
- One factor with more than two treatments: compare the treatments against each other;
- Two factors: compare the treatments against each other and the factors against each other, e.g. 2*2
- More than 2 factors: 2k factorial design

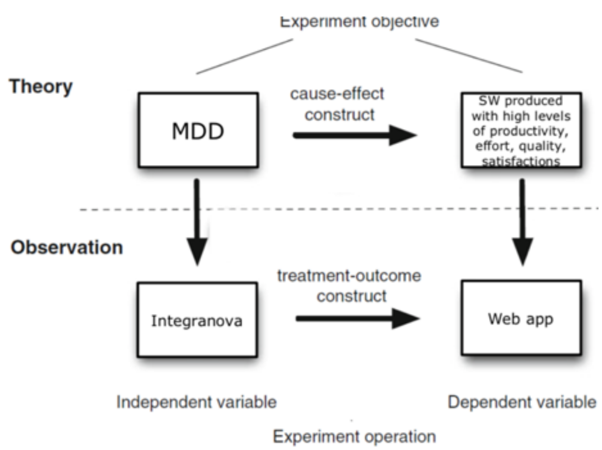
<u>(5) Instrumentation</u>

- Three types of instruments:
  - Experimental Objects: objects used in the investigation (models, code documents, descriptions)
  - Guidelines (guidelines to guide the subjects in the experiment)
  - Measurement instruments (objects to collect data during the experimental tasks)

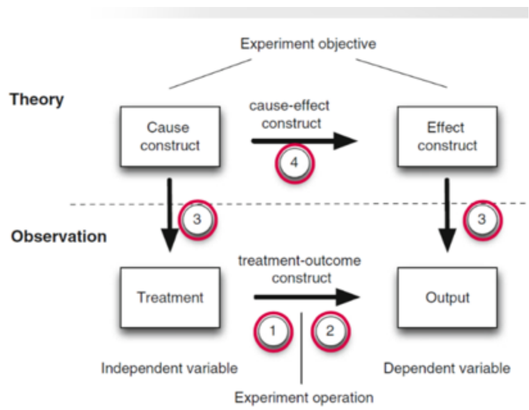
Lecture – Part 1 - L8 – Validity Evaluation

Should answer to the following questions: what is the population of your experiment, the results should be valid just for your population (yes)? What happen if I want to generalize the results?





Validities:
1. Conclusion validity
2. Internal validity
3. Construct validity
4. External validity



<u>1. Internal validity</u>
- Multiple group treaths
- Interaction with selection
- Socail treath
- Diffusion of imitation of treatmetns
- Compensatory equalization
- Compensatory rivalry
- Resentful demoralization

<u>2. Construct validity</u>
- Design of the experiment
- Mono-operation bias
- Mono-method bias
- Interaction of different treatments
- Interaction of testing and treatment
- Hypothesis guessing
- Evaluation appranshion

<u>(3) External validity</u>
- Interaction of selection and treatment
- Interaction of setting and treatment
- Interaction of history and treatment

<u>(4) Conclusion validity</u>
- Low statistical power
- Violated assumption
- Reliability of measure and treamtne
- Random things

Lecture – Part 1 - L9

There are three types of instruments:

- **Objects**: models, code documents, textual descriptions
- **Guidelines**: includes process descriptions and checklist, training in the method, etc.
- **Measurement instruments**: objects to collect data during the experimental tasks, questionnaires, interviews, results of tasks, time used when performing a certain task.

<u>MEM: method evaluation model</u>

- Used to evaluate subjects perceptions and satisfaction
  - Perceived usefulness
  - Perceived ease of sue
  - Intention to use
- Measure subject's performance:
  - Follow a set of quality goals to inspect conceptual models
  - Quality model framework:
    - Semantic :
      - **Validity**: elements of the model are correct and relevant to the problem domain – invalid statements are those that do not pertain to the problem or express something incorrectly. The validity is the degree to which all the model elements contained in the model should actually appear in the model in the right way (e.g. an element representing phenomena that does not occur in the domain is invalid)
      - **Feasible Completeness**: elements of the model include the relevant and correct statements about the problem domain. Completeness is the degree to which all the model that should appear in the model are actually contained in the model.
    - **Syntactic correctness**: is the degree to which all the model elements specified in the model use the correct syntax. (e.g. The reviewers identified candidate invalid elements based on a reference model)
    - **Efficiency**: is the degree of success during the application of a treatment according to the time consumed (model completeness/time consumed)

Lecture – Part 1 - L9 part 2

Systematic Literature Review (SLR)

How to start:

- What has been researched?
- How has research been executed?
- Where it has been executed?

SLR: answer questions like these ones! (is it worth researching this problem? Do I know enough things? Is the methodology the best way?

In a nutshell SLR is:

- Identify, analyze and interpret ALL available evidence related to a specific research question
- Give complete, comprehensive and valid picture of existing evidence

The process is in three steps

1. **Plan**
    a. *Why? Identify the need for a review*
        i. State-of-the art area needs to be identified
        ii. Practitioners want empirical evidence for decision making
        iii. Other SLR available? Are them sufficient for our investigation?
    b. *What? Specify the research question*
        i. Research question must be the focus of our research
        ii. Scope must be clear and limited
    c. *How? Develop review protocol*
        i. Follow procedures
        ii. Use a review-log document
2. **Conduct**
    a. *Which studies? Identify research studies of interest*
        i. Scientific literature and grey literature
        ii. Use keywords and reference management systems (Zotero is the best and OS)
    b. *Filer. Select primary studies*
        i. Set inclusion and exclusion criteria – beforehand to avoid bias
    c. *Extract. Collect information and data from primary studies*
        i. Collect data Based on research questions
    d. Synthesize: aggregate and synthesize information
        i. Compare studies if they are homogeneous
        ii. Otherwise, analysis of the themes and narrative
3. **Report**
    a. For whom? Identify target audience

Lecture – Part 1 - L10

Canonical action research, involves:

- Solving organizational problems
- Improving organizational circumstances for stakeholders – by means of researcher intervention
- Contributing to knowledge about both academic theory and organizational practice

Criticisms towards Action research

- Lack of methodology
- No clear difference between AR and consulting
- Either Research with little action or action with little research
- Not scientific or not research!

Canonical AR Projects

- Is **interpretive**
    - The CAR researcher-practitioner needs to engagae in many acts thinking about the organizational situation and representing the world as experienced by the clients
    - Disciplined subjectivity and imagination are critical skills
- Is **iterative**: one or more cyles of iterventions
- **Rigorous**: use correct methods in the task context:
    - Plan iteration to develop detailed problems
    - Continious rediagnose the problem
    - Apply flexibly the method
- **Collaborative**
    - Researchers and client must work together
    - Client must participate in the project in an active manner

There are guidelines which are not very useful or used

Principles of CAR

- Things to remember
    - Clear and systematic presentation of findings
    - Justification of courses of action
    - Explicit contributions to knowledge
    - Assessment of the results
- Principles and criteria:
    - Researcher-client agreement
        - Guiding foundation for the project
        - Client must understand what CAR involves
        - Mutual guarantees
        - Trust and shared inquiry
    - Cyclical process model (CPM)
        - Sequential progress thorugh the cyclical process model to ensure rigor
        - Between-stage iteration
    - Role of Theory
        - Two types of theory:
            - Instrumental (focused on the as-is situation)
                - used for diagnosis and planning
                - Used to mediate role between client and researcher
                - Explain and analyse organizational activities related to the project
                - Include tools, models and processes that theorise how work is done and outcome achieved
            - Focal (focused on the to-be situation)
                - Provides the basis for action-oriented change

                - May not remedy an organizational problem alone
        - The client needs to approve the theory chosen
    - Change through action
        - Action and change are indivisible
            - Both researcher and client must be motivated to deisgn and implement change (which must be culturally and contextually appropriate)
    - Specification of learning/ learning through reflection:
        - What should come out of CAR is:
            - Practical solutions for the client
            - Knowledge for the scholarly community of the researcher

Scorecard links performance measures:

- Financial perspective -> internal business perspective → innovation and learning perspective → customer perspective. All connected apart from the *Financial* to the *Innovation and learning*

Balanced Scorecard (BSC)

- Designed to measure whether an organization smaller scale operational activities are aligned with objectives

Balanced knowledge scorecard

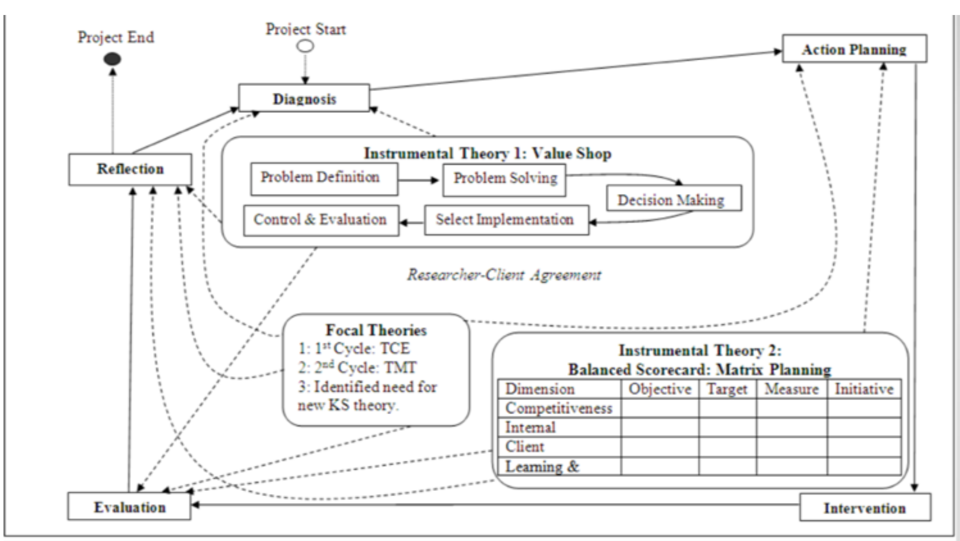
- *Knowledge Outcome Success*: are we satisfying management/ adding value?
- *Internal process perspective*: are we doing the right things?
- *Future readiness and reflection perspective*: are we ready for and reflecting on the emerging knowledge technologies and practices?
- *Client perspective*: are we delighting our clients?

Action Plan scorecard – Internal work processes

- Objectives
- Targets
- Measures
- Initiatives

Last things:

- The planned action should be designed to improve the organizational situation, including shorter and longer term actions at operational and strategic levels
- Try to follow theory and apply models
- Implement and evaluate intervention even if complicated



Lecture – Part 1 - L11

Introduction

- sell the research to a reviewer
- motivation intrinsically interesting
- Accessible to non expert

Experimental design

- Research question
- Methods
    - Justification is very important – why is this the best method for the study
- Theory selection
    - Does the theory offer a new lens to view the phenomenon? Does it enable, empower or enslave?
    - May have multiple theories
- Source of data
    - Where is it from?
    - Present data
- Data analysis
    - Which techniques?
    - Which constraints?
    - Which software?
- Contribution hoping to make?

Presenting and reviewing

- Presenting: you have to ensure it is well written and argued, no jargons
- Conferences: harder, less time (10-20 min)
    - Title, authors, affiliation
    - Introduction & background
    - Literature
    - Method
    - Discussion of results
    - Future research
    - No conclusion or references

Poster

- ¼ is text vs. visual
- Ideal for speculative and late-breaking results, for new and innovative work, or for position statements

- Graphs and not numbers
- Remove details
- Use colors
- Title for figures
- Integrate text and figures

Design and layout

Some recommendations

- Make it obvious to the viewer how to progressively view the poster: from left to right, and up to bottom
- Create individual panels to arrange the information that you want to present
- Number the panels or connect them with arrows
- Leave some open space in the design
- Use a minimum font size of 18 points, the text should be readable from five feet away
- Use all capital letters for the tittle, font size of 70 points

Lecture - Statistics Part 1

How to conduct an experiment:

1. Find Research question
2. Formulate Hypothesis
3. Operationalize abstract concepts
4. Choose Design experiment/ experimental design
5. Define pariticpant's tasks & Conduct experiment
6. Analyze data & interpret results
7. Infer from results
8. Share results

Variable: can change, we can manipulate it, we can measure it
Independent variable: we can manipulate in experiments
Dependent variable: we can measure
Empiricism: enables us to derive truth from evidences – determine relationships between observe variables

Types of studies

- **Observational study:** manipulation is outside the control of the researches
  - When? When, e.g. randomized experiment would be unethical (relationship between drug use and academic studies)
  - Researchers lacks the necessary influence to actually make something happen
  - Symptoms are rare
- **Action research:** researcher participates in the problem-solving process
  - The loop is: input-Transformation-Output
- **Contextual Inquiry:** one-on-one interaction with participants performing the target activity
- **Think aloud:** participants speak their mind without interferences
- **Focus group:** interactive group where questions are asked

Randomized controlled trials

- Randomly allocated of participants to test conditions in order to:
  - Minimize allocation bias
  - Minimize selection bias
  - Maximize statistical power

1. The research question in usability research

1. New technology *has been develop*ed: does it meet its design goals
2. New technology is *to be developed*: which abilities and limitations do humans possess that should be taken into account in the new design?

2. Hypothesis:

- Is a preliminary answer to the research question
- **Postuated a relationship** between at least two variables

- Can be generalizable, empirically tested and is can be put as a conditional sentence
- Is falifiable

3. Operationalize abstract concepts

- Allows for measurement of a fuzzy question; “healthy baby” → body weight/ size at birth
- How to do it? Check literature and previous research
    - Operational definition (in the rpesent study we define X as Y, clear approach
    - Inter-rater consenus: for difficult concepts (aggression, user-acceptance) this can be by counting the times an event happened, filming the reactions or asking multiple experts for their opinions

4. Experimental design

- Subgroup of research deisgns that fulfills certain criteria, in order of causation:
    - Experimental designs
    - Quasi experimental design
    - Ex-post-facto design
    - Correlative designs
- Correlation is not causation. Before going to causal relationship → coincidence, X subset of Y, moderator – confounder variable effect, opposite causation, THEN causal relationship
- Confounding factors are factors whose effects on a DV cannot be distinguished from each other

5. Between and within- subjects design

- Between: one participants only completes the taks in one condition
- Within: participants complete tasks in multiple conditions

People get tired and bored easily and this can create problems in the experiments – no obvious answer to do that

Deep knowledge

- Z-test: compare proportion: test difference in population proportions
    - Point estimate: difference in samples proportion

- P value meaning: *“if the null were true, then what would be the probability to observe a test statistic so far (or further) from the hypothesized value?”* P = 0.6%, if null is true, there is **very small** probability of observing this outcome -> outcome is very unlikely → strong evidence against the null
- Critical value -> decide a value at which we are willing to reject H_0 (e.g. 15 heads); decide a significance level (which is the risk of rejecting a null hypothesis by accident)
- Two sample t-test: compares two means across two populations.
    - **ONE-SIDED**
    - T-statistic: $\frac{point\ estimate - hypothesizedvalue}{standard\ error}$
    - To test :H_0: $\mu_1 - \mu_2 = \Delta_0 = 0$; alternative hypothesis: H_1: $\mu_1 - \mu_2 > 0$
        - If $H_0$is true the t distribution is centered around 0
    - P-value approach
        - Search on table with result of t-statstics and df (as right)
        - That is probability of seeing H_0
    - Critical value approach
        - $t > t_\alpha$, where α = significance level = type I error risk
        - $if\ t > t_\alpha$, reject H_0 at that significance level

$$\text{with } df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

- **TWO-SIDED**
- To test :H_0: $\mu_1 - \mu_2 = X$; alternative hypothesis: H_1: $\mu_1 - \mu_2 \neq X$
  - If H_0 is true, distribution is centered around 0
- (P-value approach)
  - Search on table; if table is one-sided, then multiply probability by two
- Critical value approach
  - $|t| > t_{\alpha/2}$, where α = significance level = type I error risk
- Confidence level: $[point\ estimate \pm critical\ value\ (\alpha/2) * standard\ error]$

Lecture - Statistics Part 2

MC : ch. 5,6,8
Hornbaek: crash course

Probability: the extent to which an event is likely to occur, measured by the ration of the favorable cases to the whole number of cases possible

Data distribution
Knowing how data is distributed is the first step towards understanding how one variable could predict another one. Models allow us to organize our data meaningfully, even while we are collecting them

Binomial distribution
Is the discrete probability distribution of the number of successes in a sequence of N independet experiments with probability p

$$P(x) = \frac{N!}{x!\,(N-x)!}\pi^x(1-\pi)^{N-x}$$

N = number of flips
x = number of desired outcome (cannot be lower than N)
\pi = probability of the desired outcome

Estimating the central tendency from data
Mode: most frequent observation, ideal for discrete data
Median: middle observation of distribution of data
Mean of sample: mean

Skewness: asymmetry of the probability distribution
- If mean = median, symmetric
- If mean < median, negative skew
- If mean > median, positive skew

Skew = $\sum \frac{(X-\mu)^3}{\sigma^3}$
Variability of observed data: sum of squared errors (SS)
Variance of the sample: $\frac{SS}{N}$ it does not grow with increasing sample size – allows us to compare the error across samples of different size
Standard deviation of the sample: $\sum\sqrt{\frac{(x_i - x_{mean})^2}{N}}$ → average error that we can expect from our current measurements, relative to the sample's mean
Variance of the population $\sigma^2 = \sum\frac{(x_i-\mu)^2}{N}$
Standard deviation of the population $\sigma = \sqrt{}\sigma^2$

Terms
μ is the population mean; x is the sample mean
σ^2 is the population variance and σ is the SD; $s^2$is the corrected sample variance, s is the corrected SD
z score = $\frac{x-\mu}{\sigma}$ → useful to measure in units of sigma around the mean as the center; the corresponding value represent the percentage of population your value is larger than. If 0.5557 → your height is larger than 55.57% of the german population
- 68.2% of population should fall within $\pm 1\sigma$ around μ ; z score = 1.645

- 95.4% of population should fall within $\pm\, 2\sigma$ around μ ; z score = 1.960
- 99.7% of population should fall within $\pm\, 3\sigma$ around μ ; z score = 2.576

Deep knowledge

- One-way analysis of Variance. We have K group and K independent samples
- $SST = \sum^k n_i(\bar{y}_i - \bar{y})^2$ this is the **variation between groups:** sum of squares due to treatment
- $SSE = \sum^k (n_i - 1)\, s^2$ this is the **variation withing groups**: sum of squares due to errors
    - **Intuition:** is SST large relative to SSE?
- **F-test:** $\frac{MST}{MSE} = \frac{\boldsymbol{SST\,/(k-1)}}{\boldsymbol{SSE/(N-k)}}$, where k is groups and N is sample size
    - Hypotheses: $H_0: \mu_1 = \mu_2 = \mu_3$ vs. $H_A: at\ least\ two\ differ$
        - With P-value the F distribution only produces value above 0, the test statistics if H_0 is true should be close to 0

Lecture - Statistics Part 3

Inferential Statistics

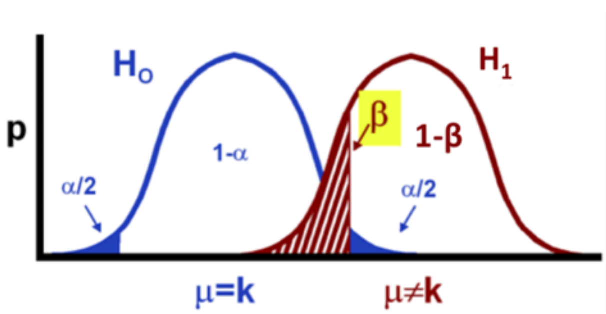
- *Binomial distribution* = $P(x) = \frac{N!}{x!(N-x)!}\pi^x(1-\pi)^{N-x}$
- *Normal distribution* = $f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
- ***Sampling distribution* = CLT the central limit theorem suggests that the sampling distribution is the distribution of a statistic over many test samples**
- **Central Limit Theorem, for large n (>30) the population is normal and the mean is approximately normal as well.**
- Confidence intervals: 95% is : $\mu \pm \frac{1.96\sigma}{\sqrt{n}}$, where 1.96 is 2 SD (95.4%). Which means that in 95 of 100 sample their mean will contain the real μx.
- $s_{xmean} = \frac{s_x}{\sqrt{n}}$ → this score is used as the standard deviation in case of one example – given that we don't have the population standard deviation
- *Calculating the z-score*: $\frac{x-\mu}{\sigma}$

Remember that hypothesis testing is educated guessing.

Possible errors are outlined in the table below

Possibilities		Actual Situation	
		H0 is True	H0 is False
Educated Guess	Reject H0	*Type 1 error*	*Correct*
	Do not reject H0	*Correct but so what?*	*Type 2 error*

- Alpha level determines the risk of rejecting the null-hypothesis when H0 is true
- Beta level determines the risk of accepting the null-hypothesis when H1 is true



- The test sensitivity (1-Beta) is influenced by:
  - Alpha
  - Sample size (N)
  - Type of test (parametric or non-parametric)
  - Variability
  - Test-directionality
  - Robustness of the effects

One-tailed vs. Two-tailed

- 2 tailed: will reject the H0 as long as the data average is different from the population
- 1 tailed: will reject the H0 as long as the data average is smaller/larger than the population
  - Perform inference only if the sign of the z-score agrees with the test hypothesis (H1)

Confirmatory testing: choose alpha-level and stick to it.

Example: Research question: Do different types of computer games result in different levels of stress?

- **IV = Game typ**
- **DV = Stress score**
- **H0 = Stress of game type 1 = type 2 → T1-T2 = it is from sampling distribution with mean 0**
- **H1 = Stress of fame type 1 noteq. type 2 -> T1-T2 = is not from sampling distribution with mean 0**
- **Degrees of freedom: (N1-1)(N2-1)**
- **Standard Error of the mean (of T1-T2) =** $\sqrt{(\boldsymbol{SE\ first\ quantity})^2} + (\boldsymbol{se\ second\ quantity})^2$
- **T-score =** $\frac{\boldsymbol{observed\ difference - expected\ difference}}{\boldsymbol{SE\ for\ difference}}$

Effect sizes (Cohen's d)

- d = (diff. Of means)/(pooled s.d.)
- d = $\frac{\bar{x}_1 - \bar{x}_2}{s}$
- s = $\frac{\sqrt{\sqrt{(n_1-1)s_1^2 + (n_2-1)s_2^2}}}{n_1+n_2-2}$
- 0.2 small effect, 0.5 medium effect, 0.8 large effect

One-way analysis of **Variance**

- One-way ANOVA, determines differences between means of two or more levels of an independent variable
- Example: *You investigate if game genres have an influence on subjective anxiety. Your independent variable is "GameType", with 3 levels. Your dependent variable is a self-reported score on a questionnaire for anxiety. You have three experimental conditions and run 10 participants per condition. This is a between group experiment design. Your null hypothesis (H0) is that the data of the three conditions come from the SAME sampling distribution.*
- **One-way ANOVA is also called F-test : F =** $\frac{\boldsymbol{betweenGroupVariance}}{\boldsymbol{withinGroupVariance}} = \frac{\boldsymbol{treatment\ variance + errorvariance}}{\boldsymbol{error\ variance}}$
    - SST (total variance) is $SS(T) = \sum x^2 - \frac{(\sum x)^2}{N}$ has two parts:
        - Effect of the independent variable – difference between treatment and effect
        - Error due to chance
    - F = $\frac{S^2 between}{s^2 within}; F = \frac{MS_{between}}{MS_{within}}$

Source of Variation	Sum of Squares	df	Mean Square	F	*p*
Between groups	$SS_{between}$	$df_{between}$	$SS_{between}/df_{between}$	$MS_{between}/MS_{within}$	0.05
Within groups	SS_{within}	df_{within}	SS_{within}/df_{within}		
Total	SS_{total}	df_{total}			

Structural model

- Each observation is a sum of mean, treatment and individual error: $X_{ij} = \mu + \tau_j + \epsilon_{ij}$
- Assumptions:
    - ***variance is the same for all treatments samples***
    - ***each sample is normally distributed;***
    - ***observations are independent from each other***

Effect size of ANOVA

- $\eta^2 = \frac{SS_{between}}{SS_{total}}$ → how much of the observed variance is accounted for by the treatment

- $f = \sqrt{\frac{\eta^2}{1-\eta^2}}$ → effect size of the F-ratio

Summary

One-way ANOVA & t-tests

- F-test and t-test are related; $F=t^2$
  (see Excel example for 2-level tests)
- Like a t-value, an F ratio is likely to be significant if:
  - the sample size (i.e., n_j) is relatively large
  - the variability within each condition is relatively small
  - the difference(s) between conditions are relatively large
- Like a t-test, an F test assumes that:
  - the samples come from populations that have equal variances
  - the dependent measure is normally distributed
  - the data were randomly sampled
  - the dependent measure is an interval or ratio scale, which allows for mean calculation

- Why not running multiple t-tests? Comparisons problems, with k levels one has *k(k-1)/2* comparisons, and with alpha 0.05 one has 1 false discovery each 20 tries

Post-hoc tests

- Bonferroni correction
  - $\alpha_{desired}$= for wrongly rejecting H0
  - m = number of hypotheses
  - $\alpha_{adjusted} = \frac{\alpha_{desired}}{m}$
- Tukey HSD test, for honest significant difference

Lecture - Statistics Part 4 – Statistical testing II

<u>Overview</u>

One-way ANOVA

- Determines differences between means of 2 or more levels of independent variable
- Effect size is eta^2 and f
- Post-hoc comparison
- Contrast analysis

Two-way ANOVA

- Determines differences between means for the independent effects and interactions of two variables

<u>DF Repeated measures</u>

k = groups, n = participants number, N = total participants observations

Df_total = (n*k)-1 OR N-1;

df_between = k-1;

Df_subjects = n-1;

df_within = N-k;

df_err = n-1*k-1;

because -> DF|wit = DF|subjects + DF|error

<u>DF Two-way Analysis of Variance</u>

Df_error = df_within = max(i) * max(j) * n-1

Df_total = N-1

Df_between = df_1 *df_2

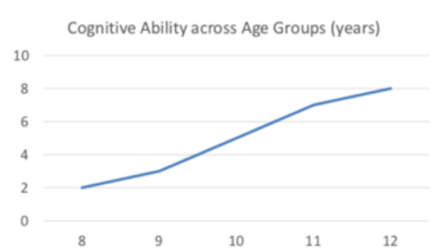
DF|wit = DF|error

<u>Questions</u>

- Factor is another term for independent or treatment variable
- If an experiment has one independent variable with five levels and a one-wat ANOVA revels a significant F ratio, one cannot tell which level is different from which other level.
- A post hoc test cannot be undertaken when an ANOVA f ratio is not significant

<u>Contrast analysis</u>

- **Used when an apparent independent variable has no real effect:**
  - k = number of conditions
  - n = number of observations
  - T = nX
  - $\lambda = \text{contrast weights linked to a hypthesis, sum must be 0}$
    - $\mathrm{MS}_{\mathrm{contrast}} = \mathrm{SS}_{\mathrm{constrast}} = \frac{L^2}{n\sum\lambda^2}$
    - $L = \sum(T\lambda) = T_1\lambda_1 + \cdots. + T_k\lambda_k$
    - $F_{contrast} = \frac{MS_{contrast}}{MS_{within}}$
    - $df = 1$



Age	8	9	10	11	12	∑
Cog. Ability	2	3	5	7	8	25
T(nX)	20	30	50	70	80	250
λ	-2	-1	0	+1	+2	0
Tλ	-40	-30	0	70	160	160

In this case, a planned contrast where performance was predicted to increase with age was significant, indicating that cognitive ability improved with age.

Two-way analysis of variance

Determines differences between means of two or more independent variables

You investigate if game types have an influence on subjective anxiety. Your independent variables are game-type and age, with 2 levels each. Your dependent variable is a self-reported score on a questionnaire for anxiety. You have 4 experimental conditions and run 10 participants per condition. ***This is a between-group 2 x 2 factorial*** *design for Age and Game Type.*

Two-way Analysis of Variance

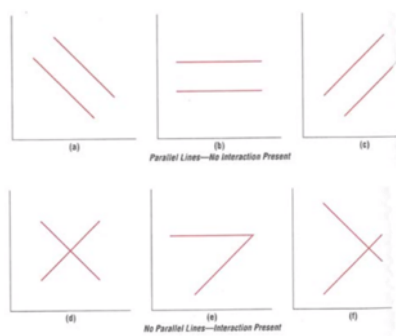
Source of Variation	SS	df	MS	F
Age	3.19	max(i)-1	SS_{Age}/df_{Age}	MS_{Age}/MS_{Error}
Game	6.16	max(j)-1	SS_{Game}/df_{Game}	MS_{Game}/MS_{Error}
Age x Game	1.72	df_{Age} x df_{Game}	$SS_{AgeGame}/df_{AgeGame}$	$MS_{AgeGame}/MS_{Error}$
Error	33.90	max(i) x max(j) x (n-1)	SS_{Error}/df_{Error}	
Total	44.98	N-1		

If you add a game type → it only becomes a between-group 2 x 3 factorial design

- ANOVA can analyse both interval or ratio scale data
- Data within each sample are either randomly or independently sampled
- Parent population from which data samples are drawn are normal and have equal variances

Interpreting interactions (X axis = A, lines = B)

- No interaction = parallel lines
  - (A) Main effect A and B
  - (B) Main effect of B, no main effect of A
  - (C) Main effect A and B
- Significant interaction = non parallel lines
  - (A) No main effect of A and B
  - (B) Main effect of B not A
  - (C) Main effect of A, not B



IF THERE IS A SIGNIFICANT INTERACTION, INTERPRET THE INTERACTION AND IGNORE THE MAIN EFFECTS

Repeated measures ANOVA

- C**an be performed where all the levels of an IV is experienced by all participants**
  - One-way only one factor
  - Multi-factorial: all factors are repeated measures
  - Mixed-design; some factors are repeated measures
- Help:
  - Fewer participants required
  - Reduce inter-subjects variability
  - Allow design with carry-over
- SS_total = SS_between + SS_subjects + SS_within

Repeated measures one-way ANOVA

H0: $\mu_1 = \mu_2 = \mu_{...} = \mu_j$;
H_{alt}: At least one mean is different from another mean

$SS_{total} = SS_{between} + SS_{subjects} + SS_{within}$

$$SS_{total} = \sum X_{ij}^2 - \frac{(\sum X_{ij})^2}{N}$$

- N=number of observations
- i=participant-index
- j=IV-index
- df = N-1

$$SS_{between} = \frac{(\sum X_j)^2}{n} - \frac{(\sum X_{ij})^2}{N}$$

- n = number of participants
- df = k-1
- k = number of IV levels

$$SS_{subjects} = \frac{(\sum X_i)^2}{k} - \frac{(\sum X_{ij})^2}{N}$$

- df =n-1

$$SS_{within} = \sum X_{ij}^2 - \frac{(\sum X_i)^2}{k} - \frac{(\sum X_j)^2}{n} + \frac{(\sum X_{ij})^2}{N}$$

(see Excel worksheet for example)

ANCOVA

- Account for a-priori factors
- Covariate – An IV with known effect

- Must be continiuous
- “predictor, risk factor”

- Use to account for obvious things
- Example: inflammation (DV) ~ protein consumed (IV) + body-weight (IV/ Covariate)

Relationship of correlation

- As X increases Y increases = positive relationship
- As X increases Y decreases = negative relationship
- No pattern = no correlation
- Finally, Correlation implies and assumes that:
    - Data is from bivariate normal population
    - Data is based on interval or ratio scale

Pearson product-moment correlation coefficient

- **Quantifies the extent to which 2 variables are associated – and if the direction is positive, negative or zero**
- Terms of the association between two variables
- From 40-59 is moderate
- T-test = $\frac{observed - expected}{SE}$; calculate t-test, df (n-2) and find critical t-value 2-talied, reject H0 if t>t_critical
    - SE = $\frac{\sqrt{1-r^2}}{n-2}$; df = n-2;

○

MacKenzie – Chapter 5

VS.NU – Netherlands Code of Conduct for Research Integrity – 2018

Ethics

- Systematizing, defending and recommending concepts or right and wrong conduct
- Are expressed in **principles**
- Main issues are: fabrication of data, plagiarism and ghost writing
    - Falsification is manipulating research processes or changing or omitting data

Four Ethical principles

- Scientific value: study should have some value
- Beneficence: should benefit others
- Informed consent of the participants
- Confidentiality of the data

Principles

Research in sciences is a process governed by standards – normativity is partly methodological and partly ethical in nature and can be expressed in terms of a number of guiding principles:

- **Honesty:** reporting the research process accurately, taking alternative opinions and counterarguments seriously, being open about margins of uncertainty, etc.
- **Scrupulousness:** means using methods that are scientific or scholarly and exercising the best possible care in designing, undertaking and reporting research
- **Transparency:** ensuring that what data the research was based on is clear, how the data were obtained, how results were achieved and what role was played by external stakeholders. If part of the data is not public, there should be an account of why this happens.
- **Independence:** it means not allowing the choice of method, assessment of data, weight on alternative statements or assessment of others research or research proposal to be guided by non-scientific or non—scholarly consideration.
- **Responsibility:** it means acknowledging that a researcher does not operate in isolation and hence tales into consideration the legitimate interests of human and animal test subjects.

Standards for good research practices

These are the principles above elaborated with more specificity to set out what researchers must take into consideration.

- **Design**
- Consider interests of science when determining the subject and structure of research
- Conduct scientific research of some relevance
- Do not make unsubstantiated claims
- Take into account lasts articles
- Make sure to answer with the research design the research question
- Be open about possible conflicts of interest
- Describe how the collected research data are organized
- Make research findings and research data public as fare as possible
- Enter into a joint research with a partner not affiliated with the research institute only if there is an agreement of this code of conduct
- **Conduct**
- Research has to be accurate and precise
- Research methods have to be scientific and scholarly
- Do not fabricate data or research results – do not remove results or data
- Describe data collected
- Manage the collected data carefully

- Contribute in making data findable accessible and reusable
- Take into consideration the interests of any humans or animals involved
- Keep expertise up to date and research only areas that coincide with your expertise (?)
- **Reporting results**
- Fair allocation of authorship
- All authors must have contributed
- All authors must have approved the final version of the project
- Be transparent about the methods and working procedure followed
- Be clear about results and conclusions
- Be explicit about serious alternative insights that could be relevant
- Cite text and research
- **Assessment and peer review**
- Be honest
- Do not use information acquired in the context without consent
- Do not use the system of peer review to generate additional citations
- Do not make assessment if not independent
- **Communication**
- Be open and honest about your role
- Be open and honest about potential conflicts of interest
- **Standard that are applicable to all phases of research**
- Inclusive culture
- Do not delay the work of other
- Call attention to other researchers' non-compliance
- Make no accusation
- Do not make improper use of research funds

Lecture 2 – Ethics in research and Experimentation

Ethics are always expressed in *principles*. Ethics can further be subdivided into issues in

- Research
- Experimentation
- University-Industry collaboration – (always ask for informed consent)

The most wide issues in research deals with

- Fabrication of data (social psychology, antropology)
- Plagiarism (data minig, computer vision, circuit design)= appropriation of another persons' idea, research results or words without giving credit
- Ghost-writing
- Improper dealing with infringement of integrity

Problem with lack of ethics standard is ruining careers of all the researchers related to the person who breaks the ethic's code.
There are two ways of committing ethics issues when dealing with data:

- **Falsification:** removing or moving data
- **Fabrication**: creating new data points

It is important to restate that research institutes have the duty to promote good research management