Coding Excellence

Master Digital Sciences

Stefan Bente

(Some) Do‘s & Don‘ts for Scientific Work in Software Engineering

SS 2025, 25.4.2025

Technology
Arts Sciences
TH Köln

Disclaimer

- This workshop **does not claim** to cover scientific work in its entirety!

- See e.g. **»Seminar Computer Science Research« (SCSR)**
  - https://digital-sciences.de/en/modules/seminar-computer-science-research/
  - Prof. Dr. Westenberger
  - This workshop uses some material from SCSR

What is Science?

What do I need this for?

What is Science? What is Scientific Work?

- What is Science?
  - **“Knowledge** arranged in an orderly manner, esp. knowledge obtained by observation and testing of facts” (1)
  - “The intellectual and practical activity encompassing the **systematic study of the structure and behaviour** of the physical and natural world through observation and experiment.” (2)

- Scientific Work is a **systematical** and **methodological** process which is conducted in order to
  - produce **new** ideas, concepts or insights
  - reflect existing problems from a **new perspective**
  - induce **new hypotheses**
  - **test** existing hypotheses

Sources (accentuation by the authors of this slideset):
1. Hornby, Albert Sydney (1974): Oxford Advanced Learner's Dictionary of Current English. 3. ed. Oxford [u.a.]: Cornelsen & Oxford Univ. Press.
2. http://www.oxforddictionaries.com/de/definition/englisch/science, 23.09.2015

Master Examination Regulations about Master Thesis

- The Master's thesis is a written assignment.
- It is supposed to show that the student is capable of independently working on a topic from his or her subject area within a specified period of time,
- both in its subject-specific details and in its interdisciplinary contexts,
- using **scientific** and **practical** (“fachpraktische”) methods.

MPO DigSc § 25 (1), translation by author

„But I plan a career in business, not in science …“

- **Q:** I plan my career to be in business, not in science. I also don‘t plan for a Ph.D. What for do I need principles of scientific work?
- **A:** In your future roles, you will be frequently asked for your expert opinion. You need be able to conduct a research that takes the relevant sources into account.
  - If you fail to do so, your proposal can be easily discredited by opponents.
- Your proposal is usually delivered in writing.
  - I.e. you should be able to transform your thoughts and ideas into a well-readable, well-structured document.
- This document must not have any formal flaws like plagiarism or faulty citations – see „opponents“ above.
- You need to present your ideas in a precise yet entertaining way.
  - From a 90 sec pitch to a full-fledged 45 min customer/board/ conference presentation.

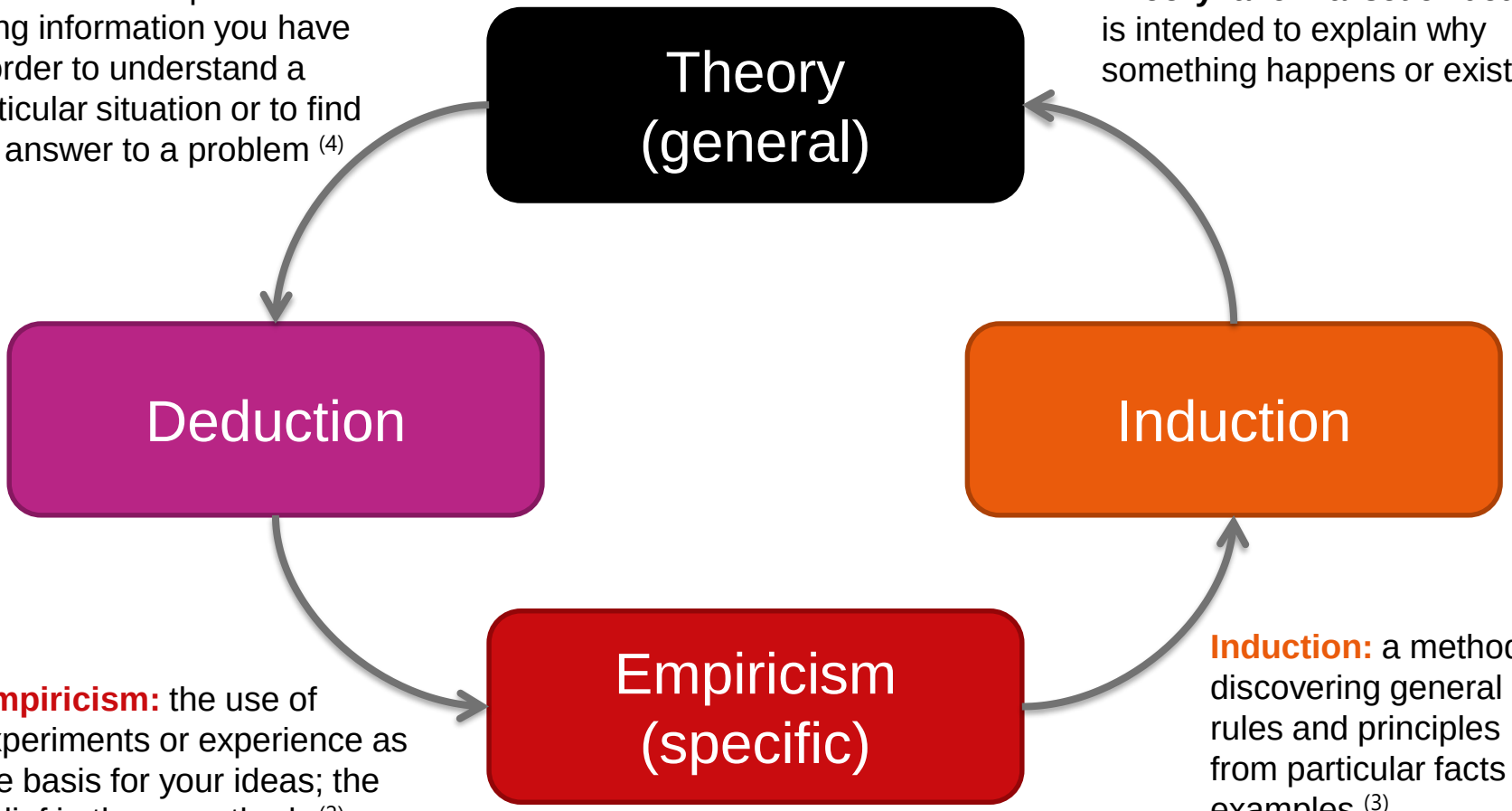
Basic Paradigms and Patterns

Pattern 1: Induction / Deduction

Quellen (alle abgerufen 23.09.2015)
1. http://www.oxforddictionaries.com/definition/learner/theory
2. http://www.oxforddictionaries.com/definition/learner/empiricism,
3. http://www.oxforddictionaries.com/definition/learner/induction
4. http://www.oxforddictionaries.com/definition/learner/deduction,

Deduction: the process of using information you have in order to understand a particular situation or to find the answer to a problem (4)

Theory: a formal set of ideas that is intended to explain why something happens or exists (1)

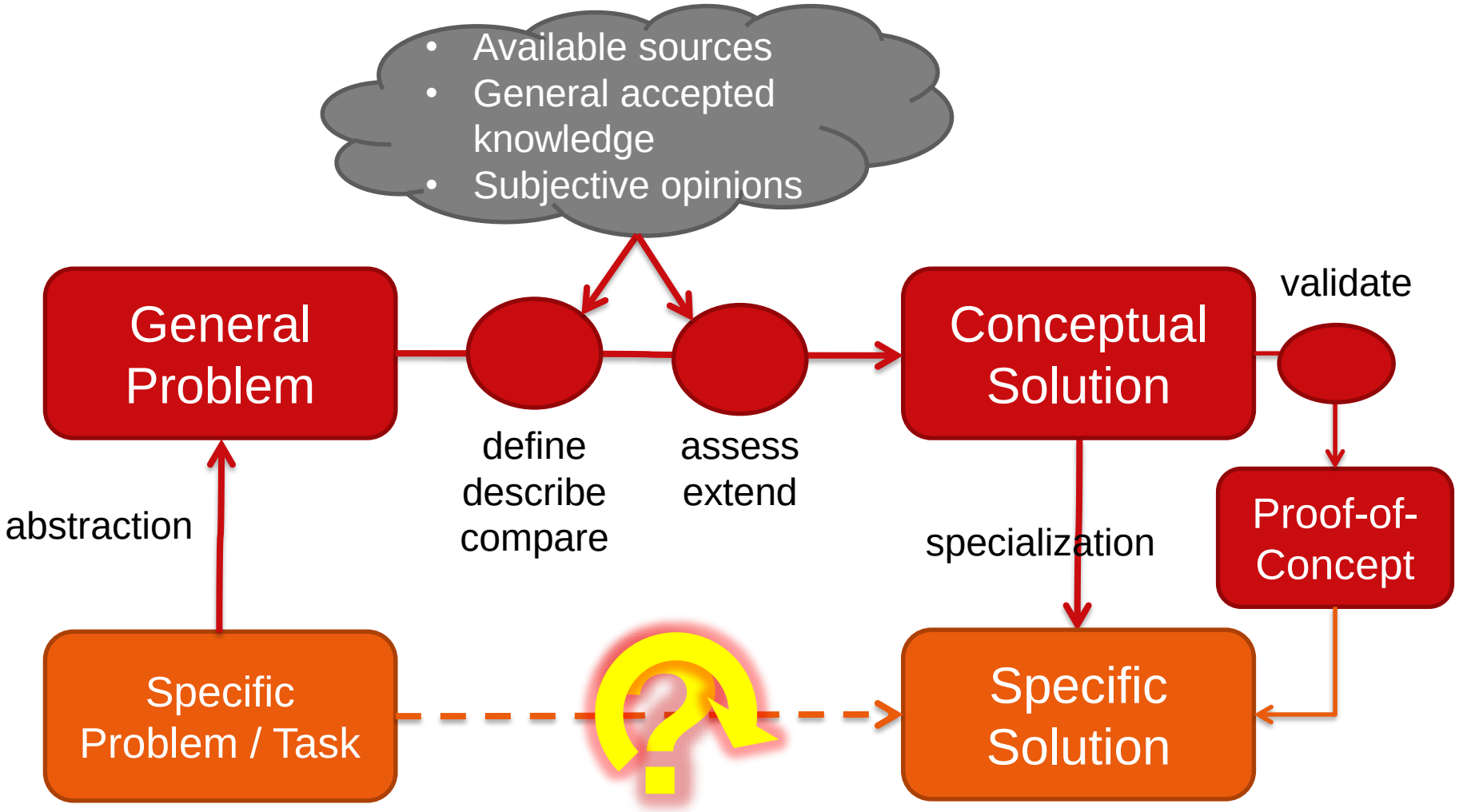


Empiricism: the use of experiments or experience as the basis for your ideas; the belief in these methods (2)

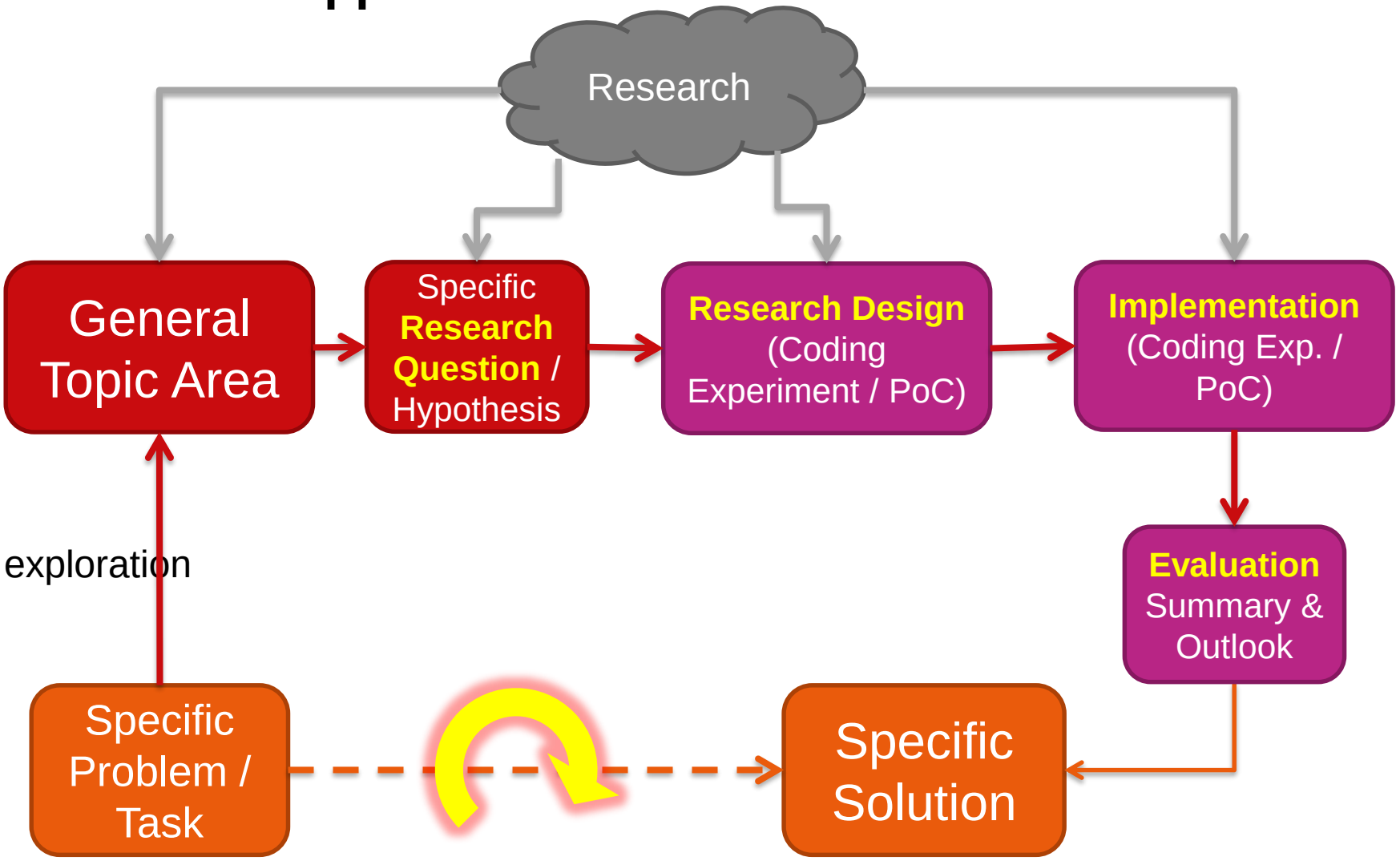
Induction: a method of discovering general rules and principles from particular facts and examples (3)

Technology Arts Sciences TH Köln

Pattern 2: Abstraction / Specialization

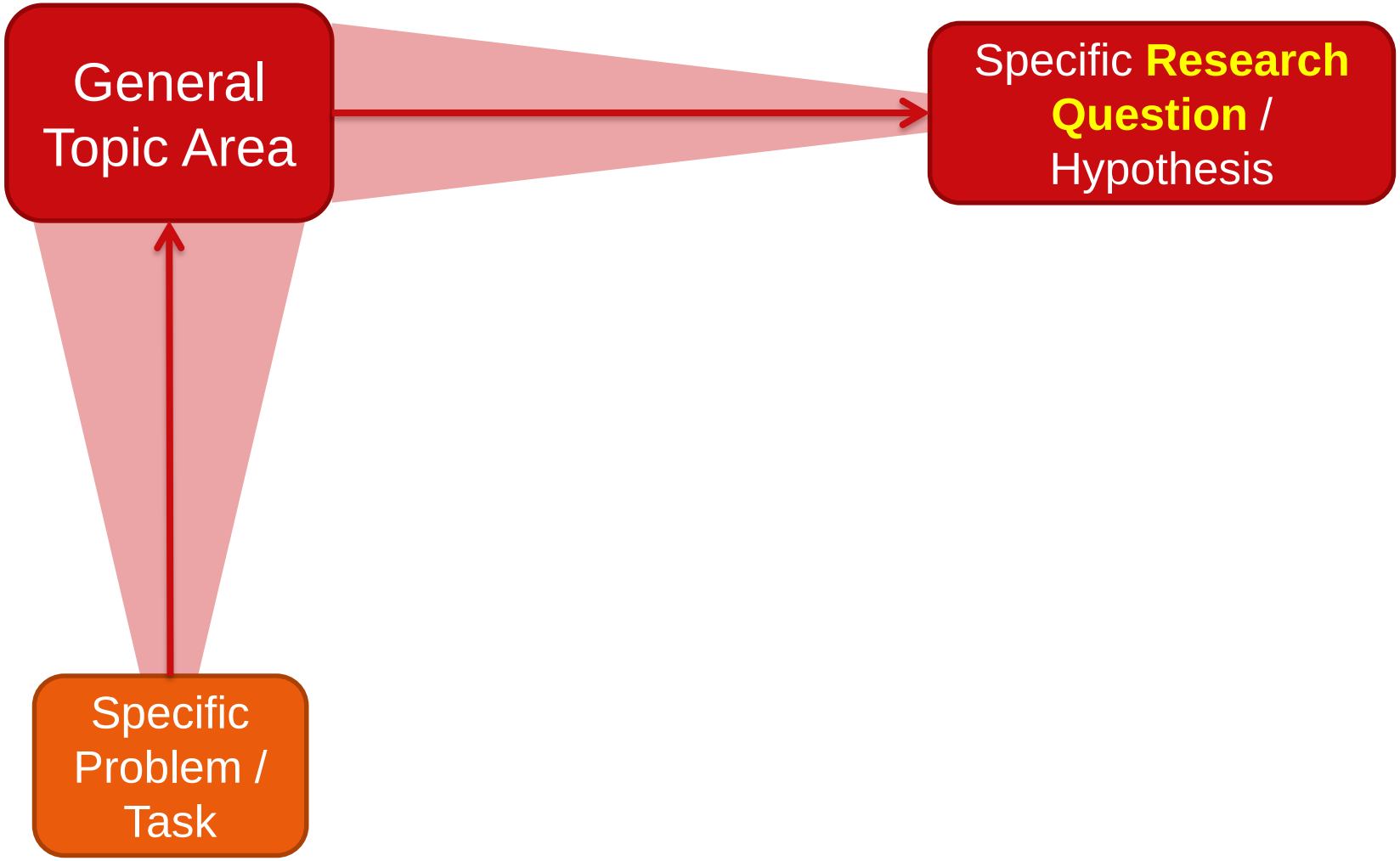


PoC-based Approach



1) Research Question / Hypothesis

1) Research Question / Hypothesis



Definition - Research Question or Hypothesis

- **Research question**
  - overall direction and focus of a research study
  - often posed as interrogative statements (questions)
  - aims to explore, describe, understand, or explain a phenomenon

- **Hypothesis**
  - provides a specific prediction or explanation for the relationship between variables being investigated
  - guiding the hypothesis testing process in scientific inquiry.
  - specific, testable statement that proposes a relationship between variables or predicts the outcome of a research study
  - Testable through empirical research methods and can be supported or refuted based on the evidence collected

What makes a good Research Question / Hypothesis?

- **Relevance**
  - Is this question worth exploring? Is it of interest to relevant stakeholders?
  - Is it not trivial?
- **Originality**
  - Am I the 123rd author proving / disproving the same issue?
  - (unless you write a Ph.D. thesis, you don‘t have to be the first, but there are always limits …)
- **Auditability** („Überprüfbarkeit“)
  - Can this question be actually answered / this hypothesis be validated or disproved?
  - Can **I / my team** do this (given our means – team size, equipment, …)?
  - **CEX:** can it be validated by a coding experiment or PoC?
- **Time Constraints**
  - Is the question / hypothesis focused enough, so that I can comprehensively handle it in the given time frame?

2) Literature Research

Evaluating Literature (and Other Sources)

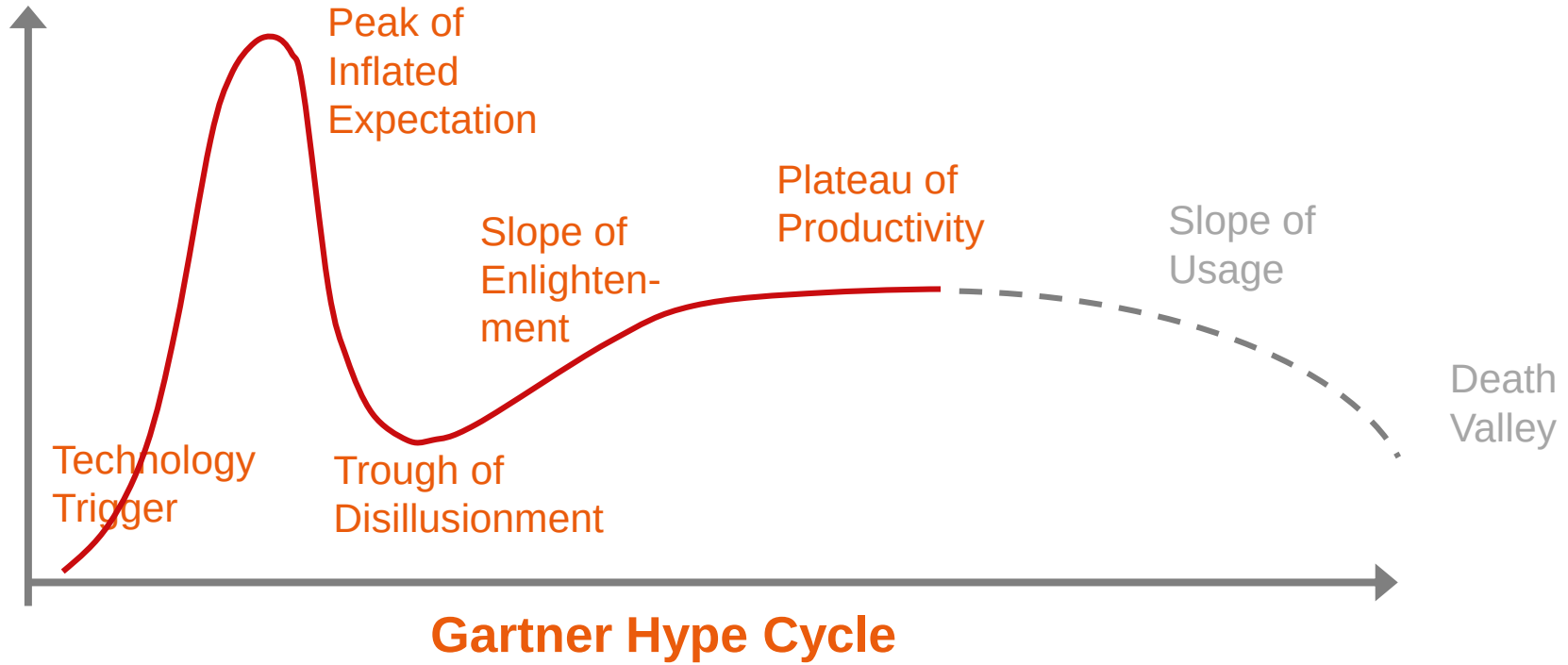
- Formulate your specific interest
  - … before reading
- Scan a paper roughly
  - … before deciding to read thoroughly – this will save time
- Describe the main statements with relation to your task
  - … **in your own words!**
- Manage your literature
  - Organize your notes
  - **Start early** with a literature reference list
    - Adding it in retrospect is nearly impossible (at least very time-consuming)
  - Keep your reference list up-to-date
    - add new references
    - remove unused ones
- Use tags for important content areas
  - Otherwise, information retrieval in 50+ sources becomes living hell

Research Tools

- TH Köln Library Services
  - https://thb-koeln.digibib.net/search/katalog
    - Full text access for many platforms when in VPN, or on campus
- Query tools to find sources
  - https://scholar.google.de/
  - https://ieeexplore.ieee.org/Xplore/home.jsp
  - https://dl.acm.org/
  - https://www.scopus.com/
- Analysis tools
  - https://www.google.com/trends
  - https://www.vosviewer.com/

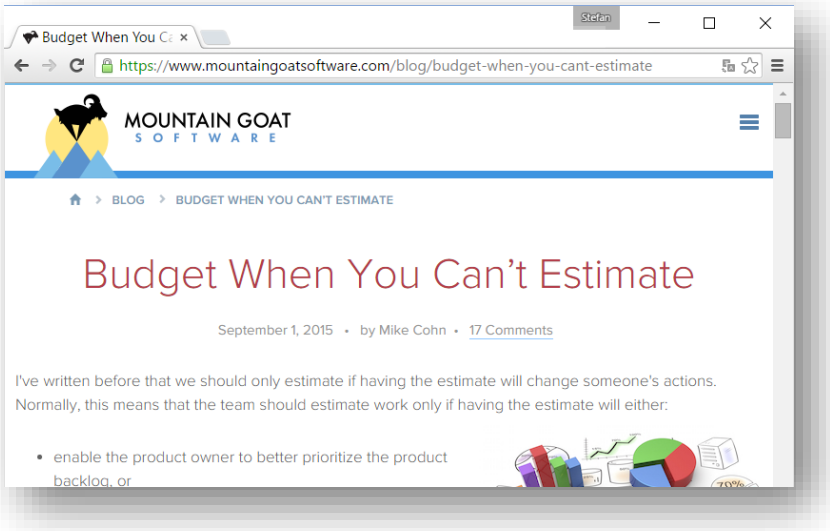
Considering Hypes: Gartner Hype Cycle

- Be aware of "hype cycles" …
  - … producing a flood of ill-researched (and potentially biased) white papers and blogs
- How to recognize a hype?



Who / What is a Suitable Source?

- Rule of thumb: peer-reviewed scientific texts can always be cited
  - text books, articles in renowned journals, conference papers
- **Software Engineering:** opinion leaders are often not scientists, but practioners/consultants/writers (example: **Agile)**
  - Therefore, many interesting, up-to-date discussions can be found **in blogs only** (before they find their way into books, eventually)



Cohn, M. (2015): Budget When You Can't Estimate. Blog post, https://www.mountaingoatsoftware.com/blog/budget-when-you-cant-estimate. Retrieved 03-Sep-15.

Selecting Suitable Sources

- How to decide if a particular source is note- (cite-)worthy?
  - In SE, innovation is frequently found outside academic world
  - this makes source selection harder, and all the more crucial

1. Research whose opinion counts in the community
2. Who is frequently quoted and referred to by other authors?
3. Bibliometrics
   - for academic source, not in scope of this workshop
4. Google hits (handle with care)

- … and last but not least:
  Form your own opinion who has something original to say!

Dealing with controversy

- What if is there is no uniform opinion amongst citeworthy authors?
  - Just say so!
  - Quote one source stating the existence of a controversy
  - Otherwise deliver quotes from the "pro" and the "con" position
- What if is there is a majority "pro" opinion, but a noteworthy "con" minority?
  - Again, one sample quote for each position would be appropriate
  - The "pro" quote is representative for the who "pro" community
    - should be as "mainstream" as you can find
  - The "con" quote should be as pointed as possible
- What if my own opinion (strongly) differs from a quoted source?
  - Try to find some source speaking for you!
  - The only way to add your own opinion could along the line of *"Own project experience also contradicts this statement."*
  - However, for such stance, you actually need to **have** own project experience

(*) Dijcks, J.P. (2013): Oracle: Big Data for the Enterprise, p. 8. Oracle White Paper, http://www.oracle.com/us/products/database/big-data-for-enterprise-519135.pdf, retrieved 03-Sep-15)

Dealing with Opinion and Context

- Differentiate between author‘s opinion and your own!

> *From an Oracle white paper (Dijcks, 2013) (*):*
> "Oracle is the first vendor to offer a complete and integrated solution to address the full spectrum of enterprise big data requirements."

- **How to quote?**
- *Oracle was the first vendor introducing a complete and integrated enterprise big data solution (Dijcks, 2013).*

  ☹ Poor citation, as it fails to express that Dijcks – as an Oracle employee - cannot be unbiased

- *In a white paper (Dijcks, 2013), Oracle claims to be the first vendor introducing a complete and integrated enterprise big data solution.*

  ☺ Good citation, as it puts the quote into its proper context

(*) Dijcks, J.P. (2013): Oracle: Big Data for the Enterprise, p. 8. Oracle White Paper, http://www.oracle.com/us/products/database/big-data-for-enterprise-519135.pdf, retrieved 03-Sep-15)

What is Plagiarism?

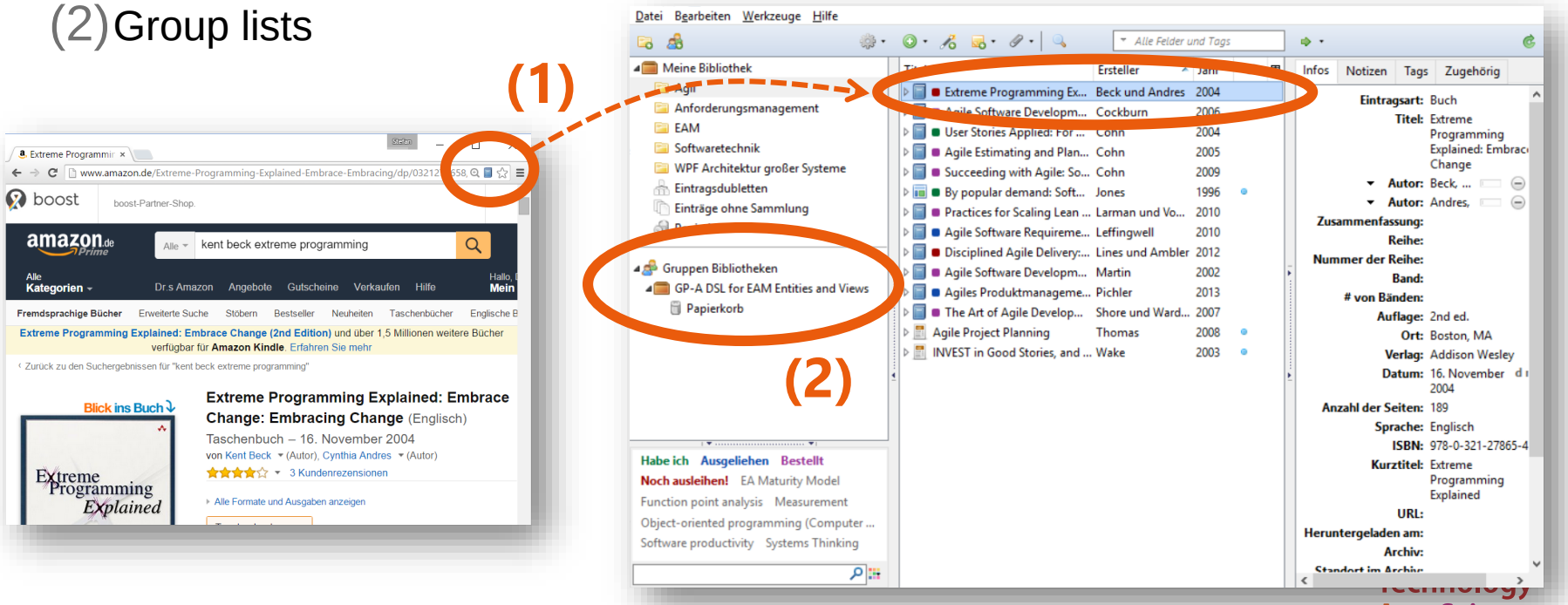
- **Plagiarism:** „The practice of taking someone else's work or ideas and passing them off as one's own.” (1)

Quelle: (1) https://en.oxforddictionaries.com/definition/plagiarism, abgerufen 02.10.2017

- Often, plagiarisms happen unintentionally:
  - During research, the student copies & pastes an internet quote into his text, and forgets about creating a reference (or finds it too much effort)
  - Certain wordings „sneak into one‘s mind“ while reading them; later, the student writes them down as part of his own text, without noticing.
- Plagiarism **can involve your own work**, too!
  - The student includes five pages from his own project report, e.g. a description of some library for the addendum, verbatim into his thesis, without marking it as reference.

Literature Management Software

- Using a tool to manage your literature references helps enormously
- TH Köln provides **Citavi**:
  https://www.th-koeln.de/hochschule/citavi-campus-lizenz_28056.php
- Personal recommendation as an alternative: **Zotero**
- Killer-Features:
  (1) Browser plugin to save directly to reference list from blogs, Amazon, …
  (2) Group lists



3) Designing
Coding Experiment / PoC

Definition: Proof of Concept (PoC) / Coding Experiment

- In software development, **a proof of concept** is a verification methodology implemented in the initial stage of the product development lifecycle.
- The purpose of the proof of concept is to test the validity of the software idea - it's all about proving that the proposed system, application or product can work in real life before you begin development.

https://www.netguru.com/blog/proof-of-concept-in-software-development, accessed 3.5.2024

- A **coding experiment** (in our context) can be
  - a PoC
  - a small sample project, which serves as a base for empiricial studies, e.g. in DevX

Do‘s and Don‘ts for PoC / Coding Experiment

Dos

- Focus on the research question / hypothesis
- Prioritize the non-functional aspects, depending on your research question
  - maintainability
  - architectural quality
  - DevX
  - performance
  - …
- Plan time-boxed
- Be ambitious

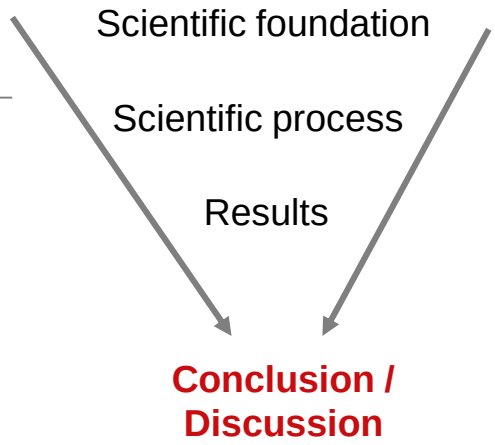
Don’ts

- Include more features than necessary
- Get too attached to an initial idea
- Overcomplicate the technology stack
- Forget about documentation
- Don’t be over-ambitious

4) Structuring Your Paper

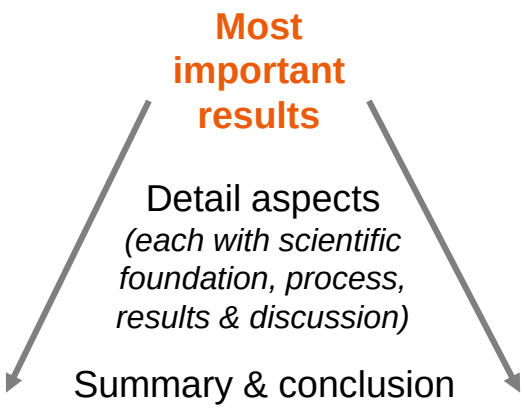
Paradigms for Structuring a Scientific Document

- A scientific document can be structured according to two main paradigms



- **Scientific Funnel**
  - Bottom-up style: foundations first, conclusions last
  - „traditional“ scientific structure
  - Suitable when focus is on the scientific process (i.e. how the results are obtained)

- **Minto (or Pyramidal) Principle**
  - Top-down (newspaper-like) style: „Most important message first“
  - (Greatly) improves readability
  - Less suitable in certain traditional contexts (e.g. research papers for peer-reviewed publications)



Quelle: Minto, B. (2008). *The Pyramid Principle: Logic in Writing and Thinking: Logical Writing, Thinking and Problem Solving* (3rd Revised edition.). Harlow, England ; New York: Financial Times Prent.

Preparation

- There are many suitable sources on scientific writing
- **My recommendation** (both wise & witty ☺): Day, R. (1977). How to Write a Scientific Paper. *IEEE Communications on Professional Communication*, *PC-20*(1), 32–37.
  - https://irkmlab.soe.ucsc.edu/sites/default/files/Day%20-%20How%20to%20Write%20A%20Scientific%20Paper.pdf

- Define a collaboration-supporting document format
  - Latex / Overleaf
  - Google Docs
  - Microsoft Word using Teams

Recommendations for Document Structure (Checklist)

- Title
  - Catchy, not too long
- Abstract
  - Summary of the thesis on ½ page max.
- Glossary
  - Definitions of terms that may be unclear
- Introduction
  - Motivation, context of the paper
  - Nature & scope of the problem analyzed
  - Research question / hypothesis
- Methodoloy / Setup of Coding Experiment
  - incl. reasoning why this will properly address the research question

Recommendations for Document Structure (Checklist), ctd.

- Results
  - Findings described in adequate details
- Summary and Conclusion
  - Conclusions drawn by results
  - Recommendations for further research
- Bibliography
- Appendix
  - if needed, use it for „specially interested readers“
  - no bloating, please

Source: Day, R. (1977), and own experience

Checklist: Quality Criteria for Scientific Work

Quality	*(in German)*	Meaning
Honesty	Ehrlichkeit	Have I properly declared all content taken over from someone else? No plagiarism?
Objectivity	Objektivität	Have I summed up my sources in an unbiased way?
Reliability	Reliabilität	Have I properly evaluated and summed up my sources?
Validity	Validität	Are my conclusions valid?
Comprehensibility	Verständlichkeit	Can the reader understand me?
Logical reasoning	Logische Argumentation	Are my conclusions correct?
Traceability	Nachvollziehbarkeit	Has the reader all information to double-check my conclusions?
Fairness	Fairness	Am I unbiased, and open to conclusions not in line with my current thinking?

Source: Adapted from: Wissenschaftliches Arbeiten - Wissenschaft, Quellen, Artefakte,Organisation, Präsentation Helmut Balzert, Christian Schäfer, Marion Schröder und Uwe Kern:. W3L-Verlag, 2008.