Supervision of SE (PhD course)

Seminar 2 “Types of theses, Research Methods, and Scoping of theses”
2024-10-09

R. Feldt, P. Leitner, **G. Gay**, H. Heyn, B. Penzenstadtler, R. Torkar

Seminar Schedule:

1. Kickoff, Philosophy of Supervision, Topic/Project Selection & Definition, Process+Rules+Admin, (September 18th 10:00-12:00), teacher Robert & Greg & Phillip
   Pre-reading: [Grohnert2024], [Feldt2010]

2. Scoping thesis projects, Types of Theses (industry vs academia, bachelor vs masters etc) and Research Methods (October 9th 10:00-12:00), teacher Greg & Robert
   Pre-reading: [Jaakola2022], [Stol2018]

3. Experiences and Case Studies, Assessment and Giving Feedback (Nov 20th 10:00-12:00), teacher Hans-Martin & Robert
   Pre-reading: [Ralph2020]

4. Time Planning and Week-to-Week Supervision, Motivation and Crisis Management (January 15th 10:00-12:00), teacher Birgit & Robert
   Pre-reading: [SUHigherEd2024]

5. Coaching Session 1 (March 5th 10:30-12:00), teacher TBD

6. Coaching Session 2 (April 2nd 10:30-12:00), teacher TBD

7. PhD Student experiences/presentations (June 13th or August 22nd 10:30-12:00), teacher Phillip & Birgit

Reading list

[Grohnert2024] Grohnert, T., Gromotka, L., Gast, I., Delnoij, L., & Beausaert, S. (2023). Effective master's thesis supervision–A summative framework for research and practice. Educational Research Review, 100589.

[Feldt2010] R. Feldt “The BTH Master thesis process and evaluation rubrics”, 2010

[Jaakkola2022] Jaakkola, Hannu, et al. "Practices for Supervising Master’s Theses in Company Context: An Anti-Pattern Approach." 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2022.

[Stol2018] Stol, Klaas-Jan, and Brian Fitzgerald. "The ABC of software engineering research." ACM Transactions on Software Engineering and Methodology (TOSEM) 27.3 (2018): 1-51.

Pre-reading for November 20th!

[Ralph2020] Ralph et al, “Empirical Standards for Software Engineering Research”, 2020

[SUHigherEd2024] Stockholm University Center for Higher Education Questionnaire’s and Documents for Student-Supervisor Alignment, downloaded online August 2024.

Industry vs Academic theses: Qs

What are the main differences between industrial and academic theses?

Does it matter which “industry” organisation it is?

Does your supervision have to be different if the thesis is done in industry?

What are the key things you would do differently?

MIPRO 2022, May 23 - 27, 2022, Opatija, Croatia

Practices for Supervising Master's Theses in Company Context: An Anti-Pattern Approach

Hannu Jaakkola[*], Tommi Mikkonen [**], Kari Systä[*] and Jaak Henno[***]
[*] Tampere University, Computing Sciences, Tampere, Finland
[**] University of Helsinki, Department of Computer Science, Helsinki, Finland
[***] Tallinn University of Technology, Tallinn, Estonia
hannu.jaakkola@iki.fi; tommi.mikkonen@helsinki.fi; kari.systa@tuni.fi; jaak.henno@taltech.ee

***Abstract* - Software Engineering (SE) university students often work part-time during their studies. In this setup the students can reform the practices of companies by transferring what they have learned to companies and correspondingly utilize what they have experienced at work in their studies. This symbiosis often continues as the students begin to work towards their thesis. The topic of the thesis relates to the problems in the company. These topics often solve a practical problem, which are not always in a perfect match with academic expectations. On the one hand the employer has certain expectations in terms of working for the company, whereas the supervising professor needs to follow the university guidelines. In this paper, we study this tension by focusing on the problems appearing in MSc thesis process in company context. We propose ways to act so that the different stakeholders -- the student, the professor, and the company -- reach the best possible results. We have analyzed the problems and their root causes. We have also taken the first steps toward anti-patterns for analysis and salvaging of the problems. The study is based on the authors' collective supervision experience, which covers over 1000 MSc theses, with the combined supervision experience of over 100 years.**

Keywords - Software engineering education, industry-academy collaboration, thesis supervision

company employing the student provides working time for writing the thesis. However, this is rare in practice for various reasons. The norm in Finland, where our experiences is from, seems to be that the employer provides the topic and time for empirical research, but the final reporting in the form of writing is not paid for. In some cases, the results of the research provide the company with real value, and hence the time invested in reporting the technical work, eventually resulting in the thesis, is compensated. Finally, in some cases, the company provides a topic -- sometimes only a problem statement, but no supervision or other support for the work.

With such variance, it is difficult to define a one size fits all solution for making and supervising a thesis in company context. As the companies and universities have interests and expectations of their own in the process, there can be various complications in the thesis process. Often, these complications are case-dependent, involving the student, the employer, and the university that eventually accepts the thesis. Sometimes the problems are recurring, for instance, defining a research problem that is relevant to both company and university.

The goal of this paper is to provide *practical support* for industry and academic supervisors who face the above

Support for Different Roles in Software Engineering Master's Thesis Projects

Martin Höst, *Member, IEEE*, Robert Feldt, *Member, IEEE*, and Frank Lüders, *Member, IEEE*

***Abstract*—Like many engineering programs in Europe, the final part of most Swedish software engineering programs is a longer project in which the students write a Master's thesis. These projects are often conducted in cooperation between a university and industry, and the students often have two supervisors, one at the university and one in industry. In particular, the Bologna Process that is currently underway to align different higher educational programs in Europe discusses industrial Master's theses as a major type of thesis project. However, there is a lack of knowledge on how best to support these projects and the different stakeholders involved. This paper presents a study where students and supervisors from software engineering Master's thesis projects at three different Swedish universities are interviewed. The intention of the study is to explore what the major problems of different stakeholders are during a project and to investigate what type of support is needed. Based on the interview results, a support model is defined, which outlines the different types of support that are needed for different roles in Master's thesis projects.**

there is now some consensus on which subjects and courses are crucial in software engineering education. However, fewer detailed guidelines are available concerning how Master's thesis projects should be conducted and supported by universities. Even in the Graduate Software Engineering Reference Curriculum (GSwERC), currently under development, no concrete support for so-called "capstone experiences" is given [2]. Since no guidelines are available to support the research-oriented nature of thesis projects at the Master's level, there is thus a need for information on how to conduct and support Master's thesis projects.

For many students, the thesis project is very different from their earlier studies. In earlier courses, they have been more supported by teachers, and the contents of the courses have, to a large extent, been predefined. They may have been involved in student projects, where they have cooperated with other students

Industry-theses (vs Academic)

- Definition and Characteristics
  - Collaboration between university and company
  - ”Practical problem-solving with academic rigor”
- Comparison with Academic Theses
  - Balancing practical goals, expectations, & research requirements
- Common areas for issues:
  - Topic issues
  - Resourcing issues
  - Communication and collaboration issues

Industry-theses: Topic issues

- Challenges in Selecting a Suitable Topic
  - Mismatch between company interests & academic standards
  - Lack of relevant prior research
- Examples:
  - 'Not a company's project' scenario
  - Topic selection ignorance
  - “Industry inferiority” complex

Industry-theses: Resourcing issues

- Challenges with Time and Resources
  - Student workload imbalance
  - Company prioritization over thesis work
- Examples:
  - Overloaded student (but really more common in industrial!?)
  - Company stakeholder's overloaded or distant/non-supportive (looming deadlines, loose interest, don't see value in academic parts, …)

Industry-theses: Communication/Collab issues

- Barriers to Effective Supervision
  - Misalignment between academic and industry supervisors
  - Intellectual property and confidentiality concerns
- Examples:
  - Wall between supervisors
  - (Slow) Pre-publication approval processes

Industry-theses: Best Practices

- Establish early communication and clear expectations
- Formal agreements on topic and resources
- Ensure early access to industry environment and data
- Regular collaboration between supervisors
- Anti-patterns for use/awareness to guide decision-making

Bachelor vs Master theses: Qs

What are the main differences between bachelor and master theses?

Does your supervision have to be different if it's a bachelor thesis?

What are the key things you would do differently?

How to "transition" the students expectations?

Bachelor vs Master theses

- Definition and Characteristics
  - Collaboration between university and company
  - ”Practical problem-solving with academic rigor”
- Comparison with Academic Theses
  - Balancing practical goals and research requirements
- Common areas for issues:
  - Topic issues
  - Resourcing issues
  - Communication and collaboration issues

Bachelor vs Master theses

- Scope and Depth
  - BT: More **limited** in scope, focus on **demonstrating a fundamental understanding**, **apply existing knowledge** to solve a defined problem (or explore a specific question)
    - Less complex research method (fewer parts, less data collection…)
    - Clearer framework and structure provided by supervisors
  - MT: **deeper exploration and more critical** engagement of subject, more **higher level** of analysis, aim to contribute **new** insights/solutions/ methodologies
    - Greater independence, less guiding from supervisors
    - Critical reading and analysis of literature

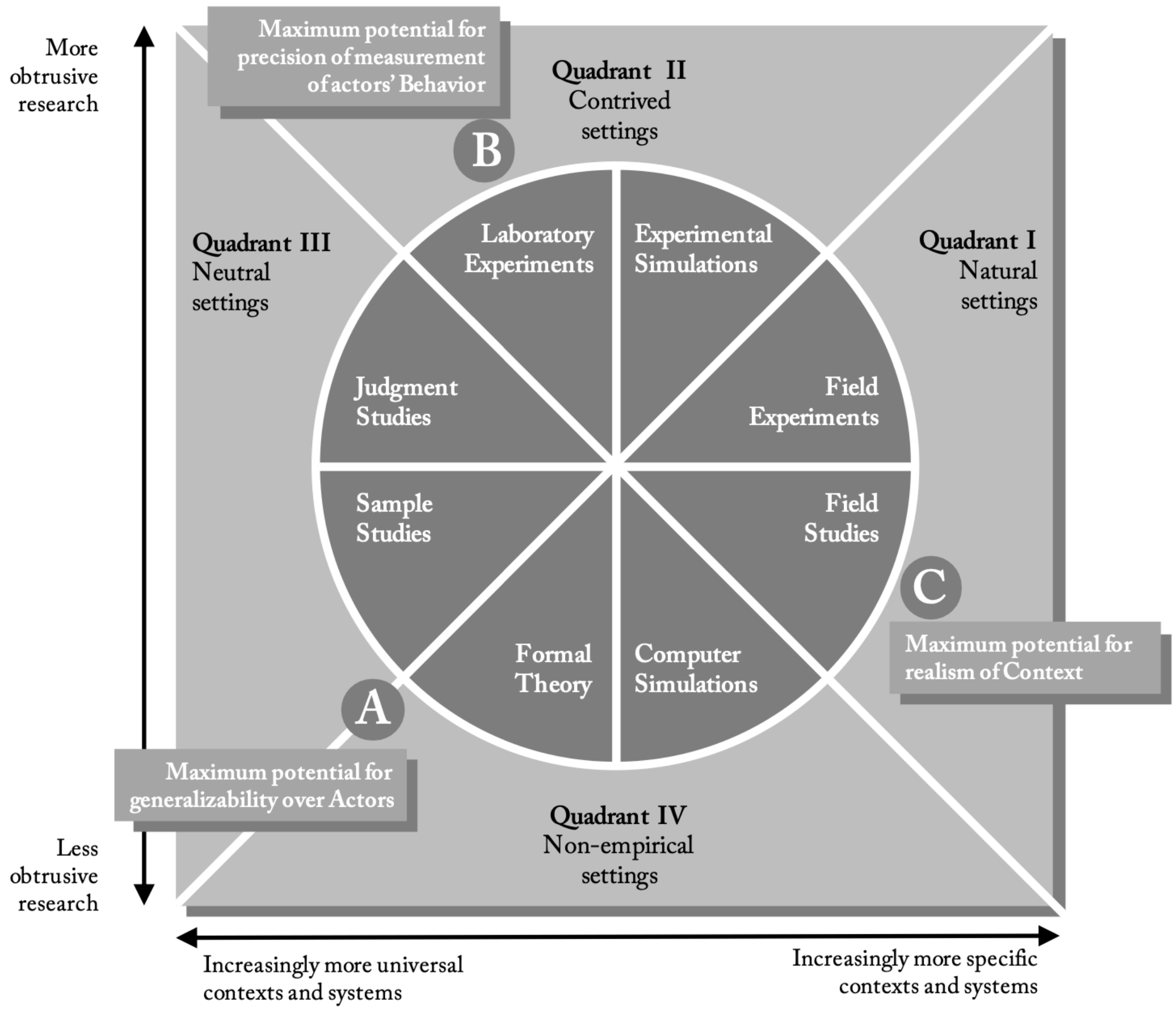

Fig. 1. The ABC framework: eight research strategies as categories of research methods for software engineering (Adapted from Runkel and McGrath [170]). Quadrants I to IV represent different research settings.

Table 5. Research strategies, their metaphors, purpose, methods and inherent limitations

Study Type	Metaphor and setting	Purpose	Typical methods & data	Inherent limitations
Field Study	*Jungle*: Natural setting that exists before the researcher enters it. Minimal intrusion of the setting so as not to disturb realism, only to facilitate data collection.	Facilitates study of phenomena and actors and their behavior in natural contexts. Exploratory, to understand *'what's going on,' 'how things work,'* or to generate hypotheses.	Case study, ethnography, observational study; qualitative data incl. interviews, field notes, archival documents, may include quantitative data.	• No statistical generalizability • No control over events • Low precision of measurement
Field Experiment	*Nature reserve*: Natural, pre-existing setting (in vivo), but some level of intrusion due to the deliberate manipulation of aspects of the setting; study affected by confounding factors.	To investigate, evaluate, or compare techniques, practices, processes, or approaches within a real-world and pre-existing setting.	Evaluative case study, quasi-experiment, Action Research; studies may use either quantitative data or qualitative data.	• No statistical generalizability • Precision of measurement affected by confounding contextual factors
Experimental Simulation	*Greenhouse, Flight simulator*: Contrived setting (in virtuo) created specifically for a study to represent a concrete type of setting. Environment is created by the researcher to study behavior of actors.	To study behavior of participants or systems in a controlled setting that resembles a real-world, concrete class of settings as close as possible.	Simulation games, management games, instrumented multi-player games; quantitative or qualitative data, depending on the simulation instrument.	• Generalizability reduced as setting is designed to mirror a specific type of setting • Realism reduced due to artificial setting

Continued on next page

continued from previous page

Study Type	Metaphor and setting	Purpose	Typical methods & data	Inherent limitations
Laboratory Experiment	*Cleanroom, Test tube*: Contrived setting (in vitro) created specifically for a study, with high degree of control of all measured variables.	To study with a high degree of precision relationships between variables, or comparisons between techniques; may allow establishment of causality between variables.	Randomized controlled experiments and quasi-experiments, comparative evaluations benchmark studies; usually quantitative data exclusively.	• Abstract or unrealistic context due to highly artificial setting • Typically scope of problem reduced to study the 'essence,' optimizing internal validity at cost of external validity
Judgment Study	*Courtroom*: Neutral setting; may be actively designed to nullify the context, so that 'responses' are in relation to some stimulus (question or instructions), independent of setting.	To elicit information from subjects for purposes of evaluation or study of an object. To seek generalizability of responses to stimuli, not generalizability to a population.	Delphi studies, interview studies, focus group, evaluation studies; use of qualitative and/or quantitative data.	• Responses not related to any specific or realistic context • Less generalizability than sample studies due to lack of *representative* sampling • Less control and precision of measurement than a lab. exp.
Sample Studies	*Referendum*: Neutral setting. Limited level of precision of measurement; no variables are manipulated. The researcher must deal with whatever data is collected.	To study the distribution of a particular characteristic in a population (of people or systems), or the correlation between two or more characteristics in a population. Information is sought of the subjects.	Software repository mining, questionnaires, interviews; analysis includes correlational methods e.g. regression. Typically, quantitative data (e.g. Likert scales) but can include qualitative data.	• Reductionist—depth of and number of data points per participant limited • Data collection not 'interactive': no option to clarify questions; repository data comes *as-is*, no opportunity to *manipulate* variables, only to *correlate* them
Formal Theory	*Jigsaw puzzle*: Non-empirical setting; typically a research office or library.	To develop a conceptualization, framework or theory on a topic. Focus is on formulating relations among concepts, or explanations that hold for a wide range of contexts.	Conceptual reasoning, concept development, development of propositions and/or hypotheses; framework development.	• Low on realism: does not consider a specific context but rather abstract concepts • No manipulation of variables or measurement (no empirical information is gathered)
Computer Simulation	*Forecasting system*: Non-empirical setting (in silico); no recording of observations in the real world. There are no actors (people, real-world systems) or real-world behavior: everything is specified in the simulation.	To model a particular system or phenomenon that facilitates evaluation of a large number or complex scenarios that are captured in the pre-programmed model.	Development of software programs that contain symbolic representations of all variables a researcher considers important; usually these variables are derived and calibrated based on prior empirical studies.	• No manipulation of variables or precision of measurement (no empirical data is gathered) • Results will be as good as the accuracy of the model representing the simulated system • Low generalizability as it attempts to model a specific class of real-world systems