# **Teaching Research Design in Software Engineering**

Jefferson Seide Molléri, Kai Petersen

Note: This chapter is part of the upcoming book titled "Teaching Empirical Research Methods in Software Engineering," to be published by Springer.

In the dynamic field of Software Engineering (SE), where practice is constantly evolving and adapting to new technologies, conducting research is a daunting quest. This poses a challenge for researchers: *how to stay relevant and effective in their studies?* Empirical Software Engineering (ESE) has emerged as a contending force aiming to critically evaluate and provide knowledge that informs practice in adopting new technologies. Empirical research requires a rigorous process of collecting and analyzing data to obtain evidence-based findings. Challenges to this process are numerous, and many researchers, novice and experienced, found difficulties due to many complexities involved in designing their research.

The core of this chapter is to teach foundational skills in research design, essential for educating software engineers and researchers in ESE. It focuses on developing a well-structured research design, which includes defining a clear area of investigation, formulating relevant research questions, and choosing appropriate methodologies. While the primary focus is on research design, this chapter also covers aspects of research scoping and selecting research methods. This approach prepares students to handle the complexities of the ever-changing technological landscape in SE, making it a critical component of their educational curriculum.

Jefferson Seide Molléri

Kristiania University College, Norway, e-mail: jefferson.mollleri@kristiania.no

Kai Petersen

Flensburg University of Applied Sciences, Germany, e-mail: kai.petersen@hs-flensburg.de Blekinge Institute of Technology, Sweden, e-mail: kai.petersen@bth.se

**Teaching Modules**. In this chapter, we incorporate active learning principles to enhance student engagement through practical activities and reflection. Following the constructive alignment framework [4], we formulate teaching modules that include Intended Learning Outcomes (ILOs), Teaching-Learning Activities (TLAs), and suggested assessments. For ease of pedagogical application, these teaching modules are clearly highlighted at the end of each relevant subsection.

# 1 Research Design at a Glance

This chapter outlines a systematic approach to designing research in SE, combining structured methodology with practical advice to help researchers effectively navigate the complexities of SE research. Additionally, we offer guidance on teaching each step of the research design process.

The approach to research design can vary significantly, generally categorized into two types, as shown in Figure 1. **Fixed designs** are highly structured and follow a systematic progression through clearly defined phases, typical for studies that utilize structured methodologies. This offers clarity and predictability, with each phase building directly on the one before [18]. In contrast, **flexible designs** are iterative and adaptive, ideal for research projects where the scope isn't initially clear and may evolve with new findings or trends [11, 18].

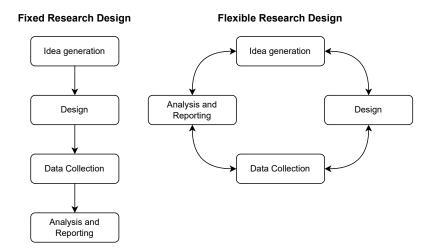


Fig. 1 Procedures for (a) fixed and (b) flexible research designs, according to [11].

The generic framework for research design depicted in Figure 2 provides the overarching structure for this chapter. It involves nine methodological steps, covering research strategy formulation (steps 1-3), tactical decisions about the overall research method (steps 4 and 5), selection of operational methods (steps 6-8), and practical considerations (step 9).

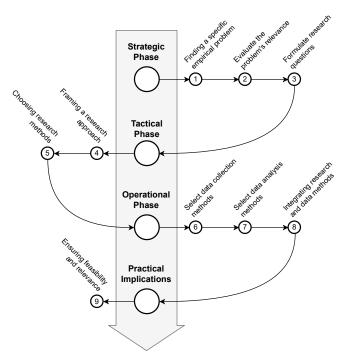


Fig. 2 Overview of the research design process.

While this framework is better suited for fixed research designs, it can also fit flexible designs with minor modifications. In flexible research design, the framework's steps are not required to be completed sequentially before data collection begins. Instead, multiple iterations may refine and adjust the research process as more data are collected and analyzed. Additionally, adjustments in the research design can also lead to changes in idea generation and other phases of the research project.

Educators should strive to teach both fixed and flexible design paradigms, providing students with the tools to choose the appropriate approach based on their research goals. Practically, researchers should consider their project's scope, the nature of the research question, and the degree of uncertainty in the data when deciding between a fixed and a flexible design.

# 2 Finding a Relevant Empirical Problem

Finding a relevant problem in SE requires understanding of the field's dynamic nature and how it shapes research directions. This section begins by exploring the evolving landscape of SE and its implications for research, then defines criteria for determining the relevance of a research problem.

# 2.1 The Current State of SE Research and Practice

Software Engineering is a broad field that encompasses both academic study and practical applications. This duality, common to applied fields, offers unique advantages and challenges when identifying a potential problem that contributes to the body of knowledge. This knowledge spans the *state of practice*, i.e., practical applications and standards known, accepted, and adopted by professionals, and the *state of the art*, i.e., the recent findings and theories that have been attested by empirical research.

Integral to SE is its broad spectrum of knowledge areas, which include diverse topics ranging from software requirements and testing to software engineering management, each offering distinct perspectives and insights into the field. Many of the established knowledge areas related to SE are outlined by the Software Engineering Body of Knowledge [5].

Currently, SE research is on cutting-edge areas such as Artificial Intelligence (AI) and machine learning, cloud computing, DevOps, and cybersecurity. Agile methodologies, while still pivotal, are evolving to emphasize scalability and adaptability. In practice, there's a strong trend towards automating development processes and incorporating data-driven decision-making. Additionally, ethical considerations and sustainability in software development are gaining traction. This evolving landscape reflects the dynamic nature of SE and also underscores the need for research that is relevant to practice.

# 2.2 Sources for Identifying Problems

Essentially, a research problem is a specific issue or gap in the existing knowledge that research aims to address or resolve. Thus, a well-defined research problem typically goes hand-in-hand with an identified knowledge gap in the existing body of knowledge.

#### 2.2.1 Analyzing Existing Literature

Analyzing the existing literature is a first step to identify obvious knowledge gaps. The body of knowledge represented by literature (i.e., state of the art) is a source for research problems of interest. Research articles serve as a primary source to identify problems the industry is currently facing, and outlining topics for further investigation. For example, a researcher interested in the challenges of integrating AI into software testing would start by examining the current literature on AI techniques within SE to determine what has been extensively studied and where gaps exist.

A thorough systematic literature review (SLR, see Chapter Y) offers a structured approach to survey the existing body of knowledge and obtain a comprehensive overview of a topic. However, researchers should consider the possible delay in academic literature reflecting the latest advancements, and thus, might also include grey literature like blogs and whitepapers for a more current perspective [9].

#### 2.2.2 Bridging Theory and Practice

Researchers should extend their focus beyond academic context and actively build relationships with companies to integrate practical insights into their research design. The diverse practices and the multi-faceted aspects of the software process adopted by companies are key sources for identifying current problems in SE. Therefore, researchers should seek a dialogue with industry for formulating relevant research problems.

Engaging with practitioners that have extensive experience and understanding of a particular field helps identify under-researched areas and validate the most current state of practice. For example, software development teams using Agile methodologies can reveal adoption rates, challenges, benefits, and organizational impacts of these practices, offering a practical context that academic literature alone may not provide.

Suggested steps for industry engagement:

- Partner with companies through joint research and development (R&D) projects to gather data on relevant phenomena and industry needs.
- Attend industry conferences and workshops to stay updated on current practices and emerging trends, and to consult with experts for practical insights.
- Share preliminary findings with industry partners and incorporate their feedback to refine research focus and ensure it addresses real-world problems.

Building on the previous example, after identifying gaps in integrating AI into software testing from the literature, researchers consulted industry experts to understand the practical challenges faced when adopting AI in testing environments. This led them to review project post-mortems from development teams that had attempted AI solutions, uncovering common issues such as integration complexity and tool usability. These insights helped define a specific empirical problem: the difficulties in integrating AI for automated test generation in agile development cycles.

#### 2.2.3 Validating Knowledge Gaps

To ensure the integrity of identified research problems, researchers should cross-reference information from various sources, like literature reviews and expert consultations. Additionally, open-access databases and industrial publications help enhancing the trustworthiness and accuracy of the information obtained. Using this multi-faceted approach, researchers can confirm that the identified gaps are not only theoretically significant but practically relevant.

# **Teaching Module 1**

**Intended Learning Outcome:** Develop the ability to formulate a well-defined research problem in SE.

**Teaching and Learning Activities:** Have students develop a research proposal, including a clearly defined problem statement. This task should involve background research, stating the significance of the problem, and hypothesizing potential outcomes.

After developing research proposals, facilitate peer-review sessions where students critique and provide feedback on each other's problem statements, enhancing their ability to identify and articulate research problems.

**Assessment:** Students will be assessed on the quality of their feedback and their ability to critically analyze the research problems posed by peers.

# 2.3 Criteria for Relevance

Relevance refers to the degree to which research findings are applicable and beneficial to real-world problems and practices. The concept of relevance commonly adopted in SE is inherited from Information Systems (IS) research [3], in which qualities such as being of future interest of practitioners, applicable in the current state of practice, and written in a style that professionals would understand, are worthy of investigation.

Petersen et al. [14] further refine this by proposing a framework to assess industry relevance in SE research. This framework evaluates research contributions across six key aspects: What, How, Where, Who, When, and Why, each addressing different dimensions of a project. For example, when assessing the relevance of a new testing framework for continuous integration, researchers might evaluate its application in various industry settings (Where), its alignment with current testing practices (How), and its potential benefits for software quality (What). This multi-dimensional relevance assessment helps understanding the feasibility, applicability, and generalizability of research findings in SE.

# **Teaching Module 2**

**Intended Learning Outcome:** Develop the ability to critically reflect on the relevance of research problems in SE.

**Teaching and Learning Activities:** Assign tasks where students must perform relevance assessment assignments and write about the relevance of a chosen research problem (e.g. from real case studies), including the use of literature on research relevance definition.

Organize debate sessions on the relevance and impact of various research problems in SE, helping students to articulate and defend their perspectives. **Assessment:** Students must submit written assignments demonstrating their ability to articulate and defend their evaluations of research problem relevance

# 3 Formulating Research Questions

Research questions (RQs) set the direction of the investigation efforts and determine how research is conducted. They are also essential for the further steps of research design, i.e., choosing a research method and detailing a research approach. A good research question should be clear and focused, stating precisely what is intended to be studied and avoiding ambiguities. It should also addresses a significant gap in the existing body of knowledge, as stated in the previous step.

Consider the following poorly formulated RQ: "How can software development be improved?" This question is too broad and lacks specificity. It does not outline a particular aspect to be investigated, the variables or the context involved. A more effective RQ could be: "How does continuous integration and deployment (CI/CD) practices impact the software release cycle time and bug detection rate in small-sized software development organizations?" This question targets a specific topic within the SE field, outlines the variables to be studied, and also defines the context by means of a population.

Well-formulated research questions not only guide the study's direction but also shed light on the most suitable methodologies to use. For example, our example of a well-formulated RQ above suggests using a quantitative approach, specific measurements like release cycle intervals and bug detection rates, and a comparison of different practices, such as the implementation versus non-implementation of CI/CD practices.

# 3.1 Crafting Effective Research Questions

In selecting empirical methods for SE research, Easterbrook et al. [8] distinguished different types of research questions that drive the choice of research and data collection methods. Table 1 provides an overview of the question types.

Table 1 Types of research questions, according to Easterbrook et al. [8].

Type	Structure	Example
Existence	Does [phenomenon] exist in [context]?	Does continuous integration exist in small software development teams?
Description and Classification	What are the [characteristics] of [context]?	What are the <i>common challenges</i> faced by <i>teams</i> during the <i>adoption of Agile methodologies</i> ?
Descriptive- Comparative	How does the [aspect] of [group 1] compare to [group 2]?	How does the <i>code quality</i> of <i>projects</i> using test-driven development compare to those not using it?
Descriptive- Process	How is the [process] [action] in [context]?	How is the <i>deployment process</i> typically <i>executed</i> in <i>DevOps environments</i> ?
Relationship	Is there a [relationship] between [variable 1] and [variable 2] in [context]?	Is there a <i>correlation</i> between <i>team size</i> and <i>productivity</i> in <i>software development projects</i> ?
Causality	Does [event] causes/prevents the [effect] in [context]?	Does the use of automated testing reduce the number of defects in software releases?
Causality- Comparative	Does [condition 1] cause/prevent [effect] than [condition 2] in [context]?	Does pair programming lead to fewer bugs than solo programming in small-scale software development?
Causality- Comparative Interaction	Does [variable 1] on [outcome] differ across [conditions]?	Does the effect of <i>continuous deploy-</i> ment on release frequency vary with team size?
Design	What is an effective way to [action] into the [context]?	What is an effective way to integrate security practices into the Agile development lifecycle?

From the questions it already becomes clear that they require different methods to be answered. Descriptive-Process questions focusing on detailed process descriptions require to gather deep insights of how processes are executed, and hence require detailed examinations of real processes with a variety of options (such as case studies, observations, or ethnographies). Causality questions require controlling confounding factors and hence point towards experimental studies.

Wohlin and Arum [22] introduced various decision points to arrive at a research design. They considered the definition of the research outcome (basic vs. applied), research logic (inductive and deductive), research purpose (explanatory, descriptive, exploratory, and evaluation), and research underpinning (positivist, interpretivist, constructivist, critical) as the strategy phase as input to choosing the concrete ap-

proaches for obtaining the results. These are strongly linked with the above questions.

#### **Teaching Module 3**

**Intended Learning Outcome:** Develop the ability to formulate and critically evaluate research questions based on given research problem formulations and goals.

**Teaching and Learning Activities:** Provide students with examples of software engineering problems and associated goals to perform a problem formulation exercise. Ask students to critique and discuss the clarity and focus of these formulations, encouraging collaboration and critical thinking. Then, have them develop a set of research questions that would effectively address the given problem and associated goal. As a role-playing activity, students take on the roles of different stakeholders (e.g., researchers and industry professionals) to formulate research questions from various perspectives.

**Assessment:** Students submit the set of research questions they developed, assessing clarity, focus, and alignment with the initial problem.

# 3.2 Frameworks for Formulating Research Questions

While the above distinction of question types helps with methodological choices, the process of formulating well-defined RQs often requires additional support from specific frameworks. These frameworks are especially useful for novice researchers, as they help in thoroughly considering all aspects of the research. Key among these are the PICO framework and the Goal-Question-Metric (GQM) approach.

**PICO Framework:** Originally from healthcare research and often adopted in SLRs, the PICO framework aids in formulating precise research questions by focusing on four essential elements [12]:

- Population: Who are the subjects of the study?
- Intervention: What is the action or technology being investigated?
- Comparison: What is the intervention being compared to?
- Outcome: What are the expected effects or outcomes of the intervention?

The PICO framework is particularly useful for empirical studies looking at the impact of specific technologies or methodologies. For instance, to investigate the impact of agile methodologies on project success rates, a research question might be, "In large tech companies (population), how does the adoption of Agile methodologies (intervention) compare to the traditional Waterfall model (comparison) in affecting project delivery times and productivity (outcomes)?"

**GQM Approach:** Tailored for the SE domain by Basili et al. [2], the GQM approach is particularly useful for evaluating technologies and strategic decisions

in SE. It involves a three-step process: (1) define overall goals to be achieved; (2) establish specific questions that reflect the goal; and (3) determine the metrics needed and how data will be collected to answer the question. The stepwise approach ensures that all the elements are mapped, and a chain of evidence is established.

GQM is better suited for process improvement studies in SE. For example, if the goal is to assess the effectiveness of CI/CD practices, a suitable question might be, "What is the reduction in deployment failures observed after implementing CI/CD practices in software development teams?" To address this question, one would measure key performance indicators (metrics) such as the deployment frequency, success/failure rates, and mean time to recovery, both before and after CI/CD implementation.

**Other Frameworks:** The FINER criteria evaluates the feasibility, interest, novelty, ethics, and relevance of research questions [10]. The SPIDER Tool (standing for Sample, Phenomenon of Interest, Design, Evaluation, and Research type), offers a multifaceted approach to question formulation for mixed-methods research [6].

Frameworks such as the ones listed above may not be ideally suited for interpretivist and constructivist research initially. These researches prioritizes understanding the meanings, experiences, and complexities of human interactions rather than measurement or prediction. In these paradigms, RQs are broad and open-ended at the initial stages to allow for an unrestricted exploration of the topic. The 5W1H method (Who, What, When, Where, Why, and How) supports the development of such questions and enhance their communication with a broader audience [20].

As the research progresses, these broad RQs evolve in response to new data and insights, capturing the dynamic nature of human behavior. For instance, a RQ might begin as "What processes influence software development practices in start-up environments?" and evolve into "How do agile practices impact team dynamics in tech start-ups?" Frameworks such as PICO, GQM, FINER and SPIDER become more relevant in the later stages of research, helping to refine and structure the developed theories into more precise research questions.

#### 4 Framing a Research Approach

Since our emphasis is on empirical research, we prioritize applied research over theoretical reasoning within this context. However, before connecting the research questions to the research logic, the purpose of the research and the research approach, we first briefly introduce the concepts.

**Research logic:** *Inductive reasoning* involves making generalizations based on specific observations or experiences. This approach starts with specific data or instances and then draws broader generalizations or theories from them. *Deductive reasoning*, on the other hand, works from the general to the specific. It starts with a general statement or hypothesis and conducts studies to test the hypothesis. An example is shown in Figure 3. *Abductive reasoning*, is a third form of logical inference that seeks the simplest and most likely explanation from a set of observations. It useful for

generating hypotheses that can be subsequently tested through inductive or deductive methods. Hypotheses may also be generated from descriptive and relationship questions (induction), and later be tested with causality questions (deduction).

# Inductive vs. Deductive

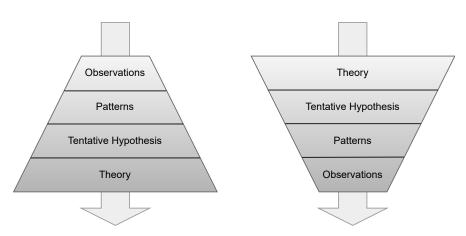


Fig. 3 Inductive versus deductive research logic.

**Research purpose:** Explanatory research aims to clarify why and how there is a relationship between two aspects of a study. It focuses on understanding the causality of relationships and often follows up on findings from exploratory or descriptive research. Descriptive research seeks to accurately and systematically describe a population, situation, or phenomenon. Exploratory research is conducted for a problem that has not been clearly defined yet. That is, exploratory research aims to gain new insights and to formulate propositions and hypotheses, or more concrete research questions. Evaluation research is concerned with, for example, assessing the effectiveness and efficiency of a software engineering intervention.

**Research underpinnings:** *Positivism* assumes a single, objective reality that can be observed and measured through scientific methods. *Interpretivism*, also known as constructivism, posits that our perception of reality and to a certain extend reality itself is shaped by human and social experience and interpretation. The *Critical approach* views reality as shaped by, for example, social, cultural, or economic values with a focus on critically evaluating phenomena in that reality. *Pragmatism* draws from the other underpinnings and often from abductive reasoning to emphasize that the validity or knowledge is determined by its practical utility in explaining or improving human practices.

The interconnection between research questions and the characteristics of research logic, purpose, and underpinnings is key for framing a research approach. We

illustrate this interplay with examples from SE research, explaining how research logic, purpose, and underpinning align to address research questions:

RQ1: What processes are emerging in DevOps teams to handle continuous integration and continuous deployment (CI/CD)?

This open-ended **descriptive-process** question is suitable for an **exploratory**-purposed study that aims to generate insights via **inductive** reasoning . It aligns with an **interpretivist** underpinning as it seeks to understand emerging practices from the perspective of practitioners.

RQ2: What are the most common types of bugs encountered in Agile software development projects?

This **description and classification** question is grounded on a **positivist** underpinning. It uses **inductive** reasoning to draw on observed data, with a **descriptive** purpose of identifying the types of bugs based on empirical evidence.

RQ3: Does implementing systematic code reviews reduce the frequency of bugs in software releases?

This aims to **explain** the relationship between code reviews and bug frequency, making it a **causality** question. It uses **deductive** reasoning, starting with a hypothesis to be tested, and aligns with a **positivist** underpinning for objective measurement.

RQ4: How effective is a given new project management tool in improving team productivity compared to the previous tool?

This **causality-comparative** question uses **deductive** reasoning to **evaluate** the impact of the tool. The **critical approach** underpinning examines the effectiveness in an organizational context.

It's important to note that research can employ a variety of combinations beyond those depicted in our examples. This inherent flexibility in research design allows researchers to effectively navigate the complexities of framing their specific research approach, facilitating the identification of suitable research question types (see Section 3). Vice versa, by formulating specific research questions, researchers can ensure that the research approach is consistent with the purpose, logic and underpinning.

# **Teaching Module 4**

**Intended Learning Outcome:** Develop the ability to characterize research regarding its logic, purpose, and underpinning.

**Teaching and Learning Activities:** Assign students articles from ESE journals for research article analysis. Have them analyze and present the research logic (inductive or deductive), purpose (exploratory, explanatory, etc.), and underpinning (positivist, interpretivist, etc.) used in these articles. Also, in groups, students role-play as researchers planning a study. Each group is assigned a different combination of research logic, purpose, and underpinning, and they must outline a research design that aligns with these parameters.

**Assessment:** Students' article analyses and/or research plans are evaluated on the understanding and application of research logic, purpose, and underpinning

# 5 Choosing Research Methods

Next step into the research design is choosing the appropriate methodologies and techniques that will be used for collecting, sampling, and analyzing the data. There are multiple decision aspects in this process [22], and they are highly influenced by the relevant empirical problem, knowledge gaps and research approach established earlier.

The research process itself can be viewed as a series of interconnected steps, each employing methods and techniques with differing scopes. These range from comprehensive processes that cover all aspects of data handling to more specific techniques focused solely on one aspect of the process [19].

In this book, we have adopted a classification that aligns with the corpus of empirical studies in SE. It's important to recognize, however, that there are varied opinions among researchers regarding the classification of specific methods. In the literature, it was pointed out that methods are often misidentified by authors [16], reflecting the complexity of method selection. See, for example, the works of Stol and Fitzgerald [19], and Molléri, Petersen and Mendes [13]; both studies employ differing approaches in classifying research methods in ESE. The former adopts a taxonomy from more mature social sciences whereas the later uses an inductive approach based solely on how SE authors label their own studies

#### 5.1 Overview of Common SE Research Methods

This section outlines common research methods in ESE, focusing on the specific contexts and scenarios where they are most effective, along with their requirements and practical considerations. It's designed to help researchers select the most suitable process for their research approach, fulfilling the methodological requirements and effectively linking them with specific research questions. Here, *requirements* refer to a short checklist that helps determine whether a study indeed meets its methodological standards; more comprehensive checklists supporting peer-review can be found in Ralph et. al [17]. The research methods are organized according to the chapters of this book where they are further detailed.

#### 5.1.1 Experiment

Experimental studies are pivotal in ESE research, particularly for evaluating the applicability of methods and technologies. These studies often involve controlled settings where participants (like practitioners, organizations, or even software systems) are exposed to different treatments to observe the effects on measurable variables. Randomized Controlled Trials (RCTs) are the gold standard for experimental studies, but other more reliant quasi-experiment designs are also common in SE research. Field experiments are conducted in real-world settings, which adds complexity and reduces control over variables compared to traditional experiments. Further information on experiments can be found in Chapter 14.

**Requirements:** The requirements on RCTs are relatively strict. The following questions help you evaluate whether your study is indeed a RCT.

- □ Does the study involve the random assignment of participants or subjects to control and experimental groups?
- ☐ Is there a control group that does not receive the experimental treatment or intervention?
- □ Does the study investigate the effects of a specific intervention or treatment?
- ☐ Are procedures, treatments, and measurements standardized and applied consistently across all participants?
- ☐ Are specific outcomes being measured to assess the effects of the intervention?
- ☐ Are participants, researchers, or both blinded to group assignments (single-blind or double-blind)?
- ☐ Is statistical analysis used to compare outcomes between control and experimental groups?

While quasi-experiments and field studies can be assessed using a similar checklist, they differ from RCTs in that random assignment and strict control over procedures (i.e., the first two questions) may not be feasible. In addition, blinding of participants and researchers (i.e, question 6) is less strict in quasi-experiments and not practical in real-world settings.

#### 5.1.2 Simulation

Simulation offers a valuable approach for exploring complex systems, processes, or behaviors in a controlled, virtual environment. This method is used when real-world experimentation is impractical or impossible. Simulations allow researchers to model and analyze software development scenarios, including system performance, user behavior, or project management strategies. They provide insights through the creation and testing of virtual models, which can replicate various aspects of SE processes. Detailed information and guidelines to simulation-based studies are provided in Chapter 16.

**Requirements:** The following questions help deciding whether it is a simulation study:

- ☐ Is a realistic and representative model of a process or system developed for the simulation?
- ☐ Are the parameters and variables clearly defined and based on empirical data or credible assumptions?
- □ Does the simulation allow for controlled manipulation and observation of key variables?
- ☐ Are the expected outcomes of the simulation aimed at drawing conclusions about real-world practices or theories?
- ☐ Is there a process for validating the simulation results against actual scenarios or data?

#### 5.1.3 Action research

Action research is a method in which researchers actively engage in real-world problem-solving and practice improvement while acquiring knowledge. This approach, characterized by iterative planning, action, observation, and reflection, is particularly suitable for exploring transformations in software processes, such as adopting new technologies or modifying existing processes or redefining organizational structures. ey to action research is the active participation of both researchers and software professionals, bridging theory and practice. Best employed for testing and refining methods and tools in live settings, action research is further detailed in Chapters 17 and 18.

**Requirements:** The following questions help deciding whether the planned study is indeed an action research:

- □ Does the study aims at solving a specific practical problem within a real-world setting?
- □ Does the research involve collaboration between researchers and participants who are directly affected by or interested in the problem?
- □ Does the study use an iterative process, involving cycles of planning, acting, observing, and reflecting?
- □ Does the research combine practical action with generating theoretical knowledge about the problem and its solutions?
- ☐ Is the research design flexible and adaptive to changes and new findings as the study progresses?

#### 5.1.4 Case Study

Case study research delves into a specific research problem within their real-life context, often focusing on organizations, projects, or processes. This method allows handling complex phenomena through triangulation, which involves integrating various data sources for deeper insights. Unlike action research or design science, the researcher in a case study observes without direct involvement in the process. Designing a case study requires careful consideration of context, as highlighted by

Petersen and Wohlin's [15] framework. Particularly effective in situations where the boundaries between the phenomenon and context are blurred, case studies are discussed more extensively in Chapter 19.

**Requirements:** The following questions help in determining whether a study is a case study, as discussed in Wohlin and Rainer [21]:

- □ Does the study describe a case? Examples for cases could be a software system, a development process, a software organization, etc.
- □ Does the study investigate a contemporary phenomena (occurs at present time)?
- ☐ Is the study conducted in a real life context (e.g., industry, open source), thus involving real developers, products and processes?
- ☐ Are multiple methods for data collection used (triangulation)?
- □ Does the researcher act passively (only collecting data) and therefore not take an active role in the investigated case?

# 5.1.5 Design Science

Design Science is ideal for creating innovative solutions to specific problems where existing theories are insufficient. It's about actively developing, testing and refining new tools, methods, or models in real-world environments. Similar to action research, this method encourages a problem-solving approach and collaboration with practice, but it focuses on creating a practical and tangible solution. Its outcomes contribute significantly to advancing SE practices and methods, bridging a gap between theory and practice. Further details are provided in Chapter 20.

**Requirements:** For assessing if a study is Design Science, use the following questions:

- □ Does the study aim to create a novel artifact (tool, method, model) to solve a specific problem?
- ☐ Is the artifact designed based on identified requirements and knowledge gaps?
- □ Does the research involve iterative cycles of development and evaluation of the artifact?
- ☐ Are the results of artifact evaluation used to refine and improve its design?
- □ Does the study contribute both to practical applications and theoretical understanding?

# **5.1.6** Survey

Survey is a research method for gathering data on practices, opinions, or behaviors from a broad audience, often practitioners. They are particularly effective in understanding trends, attitudes, and patterns within the community. This method is ideal to collect extensive data from a large population, making it particularly useful for understanding the prevailing practices in SE. It can also effectively compare different subgroups within the population. Surveys can be tailored to specific research

questions, making them versatile for various knowledge areas. Further details about surveys are given in Chapter 21.

**Requirements:** The following questions help determine if a study is a survey:

- ☐ Is the survey designed to collect data on practices, opinions, or behaviors from a wide audience?
- □ Are the survey questions aligned with the research objectives and hypotheses?
- □ Is the survey sample representative of the target population?
- ☐ Is there a clear plan for data collection and analysis of survey results appropriate for the study design?
- ☐ Does the study ensure ethical considerations, such as consent and anonymity of respondents?

#### 5.1.7 Literature Review

Systematic Literature Review (SLR) is a comprehensive overview of existing literature on a specific topic, ensuring a thorough understanding of the current state of the art. SLRs follow a structured approach to minimizing bias in literature selection and analysis. They are crucial for identifying research gaps, summarizing existing evidence, and setting the foundation for future research. Systematic Mapping Studies (SMS) are broader in scope but less focused than SLRs, providing a broader overview of a research field and categorizing existing literature. Multivocal Literature Review (MLR) includes both academic and grey literature, offering a perspective from both research and practice. These methods are invaluable for consolidating knowledge in SE. More on this research method is detailed in Chapters 22 and 23.

**Requirements:** The following questions help deciding whether the planned study is indeed a SLR:

- ☐ Is the review guided by a clear set of research questions or objectives?
- □ Does it follow a predefined and transparent methodology for searching, selecting, and appraising literature?
- ☐ Are inclusion and exclusion criteria for literature selection clearly defined and applied consistently?
- ☐ Are the credibility and relevance of literature sources critically assessed?
- □ Does the review synthesize findings from selected literature to answer the research questions?
- ☐ Is the review's methodology documented in a manner that allows for future replication and validation?

SMS can be assessed using a similar checklist, but the synthesis approach (see question 4) in SMS is about mapping the field's landscape rather than providing detailed answers to research questions. For MLR, we suggest adding the additional requirements:

□ Does the review include a mix of academic and grey literature (such as industry reports, blogs, and whitepapers)?

☐ Is there a strategy to integrate and compare insights from both types of sources?

# 5.1.8 Ethnography

In ESE, ethnography studies focus on understanding the cultural and social dynamics within live environments, such as a software project or a development team. It is a qualitative research method involving immersive observation and interaction with participants in their natural settings, gathering insights into the behaviors, practices, and interactions. It aims at uncovering implicit knowledge, workflows, and challenges that might not be evident through other research methods. Ethnography is useful when exploring the human and organizational aspects of software development, offering rich, detailed insights into the lived experiences of practitioners. Chapter 24 offers more details on ethnographic studies.

**Requirements:** The following questions help in designing an ethnography:

- □ Does the study involve immersive, long-term observation within a realistic setting?
- □ Are qualitative data collection methods like participant observation and interviews used?
- □ Does the researcher engage deeply with the cultural and social context of the studied environment?
- ☐ Is there a focus on investigating human or organizational factors, such as behaviors, practices, and interactions of participants?

#### 5.1.9 Grounded Theory

Grounded Theory (GT) is a qualitative research method aimed at generating theories from empirical data. Primarily used in exploratory studies, it involves systematic data collection and analysis to construct theoretical frameworks that explain the observed phenomena. GT involves systematic data collection and analysis, often through techniques such as interviews, observations, and document analysis. It is particularly useful for understanding complex human and organizational dynamics in software development, offering insights into processes, behaviors, and interactions within this field.

**Requirements:** For assessing if a study is grounded theory, use the following questions:

- □ Does the study aim to develop a theory grounded in empirical data?
- ☐ Is there an iterative process of data collection and analysis?
- ☐ Are concepts and categories developed inductively from the data?
- ☐ Is there a constant comparison of data to refine and develop theoretical insights?
- □ Does the study emphasize theoretical sampling to explore emerging themes and concepts?

#### **Teaching Module 5**

**Intended Learning Outcome:** Develop the ability to assess whether studies meet specified methodological requirements.

**Teaching and Learning Activities:** Provide students with a set of criteria for different research methods. Assign a selection of research papers labeled as a specific type of study, and ask them to analyze each paper to determine if it meets the methodological requirements of the given study type.

Also, students role-play as authors and peer-reviewers in a research method debate. The 'authors' defend their methodology choice, while 'peer reviewers' critically assess the alignment with method requirements.

**Assessment:** Students' analyzes are evaluated on the accurate assessment of methodological adherence and argumentation on methodological validity.

#### **Teaching Module 6**

**Intended Learning Outcome:** Develop the ability to argue for and against the choice of research methods based on a formulated research strategy.

**Teaching and Learning Activities:** Students are divided into groups, each advocating for a different research method as the best fit for a given research approach (including problem formulation, study goals and questions/hypotheses). A debate is held to discuss the merits and drawbacks of each method in relation to the strategy.

**Assessment:** Students are assessed based on their participation in debates and their arguments in support of or against the methods discussed.

#### 5.2 Data Collection

This section serves as a toolkit to various data collection techniques, outlining the applications, advantages, and the types of data that each technique yields. These methods address different research challenges in ESE, offering both breadth and depth in data collection strategies.

## 5.2.1 Document Analysis

This method involves examining and interpreting documents to gather data. It is essential in SLRs, case studies, and sometimes in action research. In literature reviews, researchers systematically search and select academic papers and other scholarly articles to understand the existing body of knowledge on a topic. In field studies, researchers gather a variety of documents such as technical specifications, project reports, user manuals, and development logs. This method offers insights

into the historical and contextual background of a research topic, contributing to a more comprehensive understanding of it.

#### 5.2.2 Data Mining

Data mining or Mining Software Repositories (MSR) involves analyzing extensive data sets to uncover patterns, trends, and anomalies. It's particularly aligned with experimental studies and design science research. Data mining can reveal insights into development practices, code quality, and team productivity. By applying statistical and machine learning techniques, researchers can extract meaningful information from large volumes of data, offering a data-driven understanding of processes and outcomes. Further details on MSR can be found in Chapter 15.

#### 5.2.3 Observations

Observations as a data collection method involve directly watching processes, practices, and behaviors as they naturally occur in real-world settings. This method is useful in case studies, action research, and ethnography, where understanding the workflow and interactions in software development environments is vital. Through observations, researchers capture nuances and dynamics that might not be revealed through direct inquiry. It allows for a deeper understanding of phenomena such as team collaboration, decision-making processes, and the actual use of tools and methods in practice, providing rich, contextual insights into the everyday realities.

#### 5.2.4 Questionnaires

Questionnaires are a tool for systematically collecting data from a large number of respondents. It is particularly effective in survey research, where understanding broad trends, opinions, or behaviors across a wide audience is essential. Questionnaires are used to gather quantitative and qualitative data on various aspects, such as developer experiences, user satisfaction, or tool effectiveness. Their structured format allows for efficient data analysis, and they can be distributed widely, making them ideal for large-scale studies. The design of questionnaires, including question clarity and response format, is crucial for ensuring accurate and useful responses.

#### 5.2.5 Interviews

Through direct conversation with individuals, researchers can gather detailed information about their experiences, opinions, and practices. Interviews are especially beneficial in case studies and action research, where understanding personal perspectives is crucial. It allows researchers to explore complex topics like decision-

making processes, experiences with specific technologies, or perceptions of industry trends. This method facilitates a deeper, more nuanced understanding of the human elements, often uncovering information that isn't accessible through quantitative methods. With regard to interviews, the type of interview to be conducted needs to be decided:

- Structured interviews with a standardized set of questions that are strictly followed.
- Unstructured interviews which are informal and do not follow a set of predetermined questions.
- Semi-structured interviews mixing predetermined questions and those emerging during the conversation.

#### 5.2.6 Other data collection methods

Beyond the commonly mentioned methods, there are other data collection techniques worth noting. Focus groups and workshops involve gathering a group of individuals to discuss and provide feedback on a particular topic or issue. Delphi studies consist of rounds of inquiry to a panel of experts, with each round using feedback to refine opinions and reach consensus on specific topics. Artifact analysis involves examining the software products, such as code quality, architecture, and usability. These methods provide additional lenses through which researchers can gather valuable data, complementing more traditional data collection approaches.

#### **Teaching Module 7**

**Intended Learning Outcome:** Develop the ability to evaluate and select appropriate data collection strategies for various SE research contexts.

**Teaching and Learning Activities:** Organize a data collection workshop where students must analyze different data collection methods, discussing their advantages, limitations, and applicability to various SE research contexts.

In addition, assign students to select a data collection strategy for a given research approach, requiring them to justify their choice based on the research objectives and context.

**Assessment:** Students' justifications are evaluated on the understanding and application of data collection methods.

# 5.3 Data analysis

This section overviews various data analysis methods used in ESE research, including both qualitative and quantitative approaches. This summary helps in understanding the strengths of each approach and how they can be applied to analyze data effectively.

#### 5.3.1 Thematic Analysis

Thematic analysis is another qualitative qualitative approach, used for identifying and interpreting patterns or themes in data. It's particularly useful in organizing and analyzing qualitative data like interview transcripts and observation notes. This method helps in understanding the underlying processes and dynamics of the phenomenon under study. ESE often employs thematic analysis to gain insights into areas like developer experiences, team interactions, or user feedback.

# 5.3.2 Content Analysis

Content Analysis, offers a more structured approach to analyzing textual data. It involves quantifying and analyzing the presence, meanings, and relationships of certain words, themes, or concepts. Researchers use content analysis to systematically evaluate communication patterns, document contents, or online discussions. This method is particularly valuable for assessing trends, attitudes, and cultural norms within the SE community, providing a quantitative perspective on qualitative data.

#### 5.3.3 Statistical analysis

This is a fundamental tool for interpreting quantitative data and turning it into actionable insights. The analysis involves using statistical techniques to test hypotheses, understand relationships, and draw conclusions from data. This can include descriptive statistics to summarize data, inferential statistics to make predictions or inferences from a sample to a population, correlational analysis to assess relationships between variables, and advanced techniques like regression analysis to understand the relationship between variables. Used in experimental studies, surveys, and sometimes in case studies.

#### 5.3.4 Machine Learning

Machine learning (ML) as a data analysis approach involves using algorithms to analyze large datasets and extract insights. It's effective in identifying patterns, predicting outcomes, and automating decision-making processes. ML can be applied to areas such as defect prediction, code analysis, and user behavior modeling. This

approach transforms large, complex datasets into meaningful information, aiding in the development of more efficient and effective software engineering practices.

#### 5.3.5 Meta analysis

Meta-analysis is a quantitative method used to statistically combine and analyze results from multiple studies on a similar topic. It provides a comprehensive view of the existing evidence, identifying overall trends and drawing more robust conclusions than individual studies alone. Meta-analysis is particularly valuable for synthesizing findings across different research projects, thereby enhancing the understanding of the effectiveness and impact of various SE practices and tools. This method offers a systematic approach to assess the generalizability of findings and to identify areas where research is consistent or where there are discrepancies.

#### 5.3.6 Other data analysis methods

Beyond the most frequently mentioned analysis techniques, there are others worth mentioning. **Cross-Case Analysis** is employed in case study research to compare findings across multiple cases, enhancing the generalizability by drawing comparisons and contrasts across several scenarios. **Iterative testing and Prototype analysis** involve evaluating and refining a product or process through repeated cycles, allowing for gradual improvements in proposed solutions.

#### **Teaching Module 8**

Intended Learning Outcome: Develop the ability to select appropriate data analysis methods aligned with the data type and research needs. Teaching and Learning Activities: In a workshop, students analyze various data types (e.g., survey results, interview transcripts) to choose the most suitable analysis method and justify their selections. They also compare methods which are similar or target the same type of data (e.g., content vs. thematic analysis) and discuss their suitability in specific research contexts. Assessment: Students are assessed on the rationale behind their choices, with emphasis on their ability to align data analysis methods with data types.

#### 5.4 Selecting Research, Data Collection and Analysis Methods

Having briefly presented the different research methods, for the given questions, we have a first pointer which methods may be suitable for the questions asked. Table

2 illustrates how different research question types (see Section 3) can be matched with potential research methods, and provides a brief rationale for each method's selection. Please note that this table is not comprehensive; it serves as an example, and many other valid method-question pairings exist.

Table 2 Mapping of question types to research methods with rationale.

Question Type	Potential Research Methods	Argumentation for Usage
Existence	Surveys; Observations	Surveys are effective for quantifying prevalence, and observations allow for firsthand witnessing of the phenomenon in its natural context.
Description and Classifica- tion	Document Analysis; Case Studies	Document analysis is suitable for understanding documented properties and categorizations, while case studies provide an in-depth look at specific instances to understand their nature and classify them.
Descriptive- Comparative	Surveys; Case Studies	Surveys can gather comparative data across a large sample, and case studies allow for a detailed comparison of specific instances.
Descriptive- Process	Observations; Case Studies	Observations give a real-time view of processes as they occur, while case studies can provide a comprehensive understanding of the process within a particular context.
Relationship	Surveys; Interviews	Surveys can quantify and identify relationships across a broad sample, and interviews allow for in-depth exploration of perceptions and experiences regarding these relation- ships.
Causality	Controlled Experiments; Data Mining	Controlled experiments are ideal for testing causal relationships under controlled conditions, while data mining can uncover patterns and potential causal links in large datasets.
Causality- Comparative	Controlled Experiments; Surveys	Controlled experiments accurately compare different conditions for causal effect, while surveys can compare perceptions and experiences across different groups to infer comparative causality.
Causality- Comparative Interaction Design	Controlled Experiments; Data Mining Design Science; Action Research	Controlled experiments can analyze variable interactions under specific conditions, and data mining helps in identifying complex patterns of interaction in large-scale data. Design Science is suitable for developing and evaluating artifacts to meet desired goals, while Action Research allows for iterative testing and refinement in real-world settings.

To illustrate an example where a mismatch between research question types and research methods can lead to ineffective combinations, consider using surveys for answering causality questions. While effective for gathering a broad range of data, survey has limited control of confounding factors, which makes it unsuitable for establishing causal relationships. Still, surveys can be quite useful for Causality-Comparative Questions, where comparing perceptions or experiences across different groups can infer comparative causality.

The second step in research involves aligning research methods with suitable data collection and analysis techniques. This strategic pairing is essential for en-

suring the effectiveness and validity of the research. As Wohlin and Aurum [22] suggest, research methods should be viewed as frameworks incorporating various data handling techniques. Table 3 provides typical pairings for ESE, but researchers are encouraged to explore alternative combinations tailored to their specific research needs and contexts. This adaptable approach ensures a more targeted and effective research methodology.

**Table 3** Typical combinations of research methods and data collection and data analysis approaches. The data techniques are ordered by their significance to the research method.

Research Method	Data Collection	Data Analysis
Action Research	Observations; Interviews; Document analysis; Focus groups	Thematic analysis; Statistical analysis (descriptive, inferential)
Survey	Questionnaires; Interviews	Statistical analysis (descriptive, inferential, correlational)
Literature Reviews	Document analysis	Content analysis; Meta-analysis
Experiment	Data mining; Questionnaires; Artifact analysis	Statistical analysis (inferential, correlational, regression)
Simulation	Data mining; Observation; Artifact analysis	Statistical analysis (descriptive, inferential)
Ethnography	Observation; Interviews; Artifact analysis	Thematic analysis
Case Study	Interviews; Document analysis; Observations, Questionnaires; Focus groups	Thematic analysis; Content analysis; Cross-case analysis; Statistical analysis (descriptive)
Design Science	Data mining; Observations; Document analysis; Delphi study; Artifact analysis	Iterative testing; Prototype analysis
Grounded theory	Observation; Interviews; Document analysis	Thematic analysis; Content analysis

Moreover, the application of triangulation in research, which involves using multiple data sources, methods, or theoretical perspectives, further enhances the credibility and depth of findings. In ESE, triangulation allows researchers to validate results through different lenses, cross-verify findings, identify inconsistencies, and capture various dimensions of the research problem. For instance, when studying software development practices, combining quantitative surveys to gather broad data on team productivity with qualitative interviews to gain in-depth insights into individual experiences can offer a richer and more nuanced understanding. This approach ensures a more robust analysis that better reflects the complexities of real-world contexts.

# **Teaching Module 9**

**Intended Learning Outcome:** Develop the ability to synergize data collection and analysis methods to align with research questions, methods, and data types.

**Teaching and Learning Activities:** Students role-play as researchers where they must choose an overall research method, and suitable data collection and analysis techniques for a given research problem, explaining how these choices complement each other. Utilize collaborative tools like Miro or Trello to allow students to visually map out research designs.

**Assessment:** Students receive real-time feedback from peers and instructors on their collaborative tools. They are assessed on how well the methodological choices align with the research questions, methods and data types.

# 5.5 Data Collection and Analysis in Flexible Research Designs

Flexible research designs, characterized by their iterative nature, present unique challenges and opportunities in the context of data collection and analysis. Data collection methods are not static but evolve based on insights gained during the research process. Initial methods are chosen to address preliminary questions, but as new patterns or themes emerge from the data, researchers may expand their data collection techniques. This iterative approach includes emergent sampling, i.e. including additional participants or sources as needed; and triangulation, i.e. employing multiple data collection methods to uncover new dimensions of the research problem.

Data analysis in flexible designs is an ongoing process that begins early and evolves throughout the study. Researchers continuously reflect on data, refining research questions and adjusting methods as new insights emerge. This iterative analysis, knows as progressive focusing, involves broad initial data examination that gradually narrows to focus on significant emerging themes. Effective implementation of flexible data collection and analysis allows researchers to align their analytical focus with the evolving research context and findings.

# 6 Ensuring Feasibility and Relevance

While the questions we wish to answer should be the main driver for the method selected, it may not always be feasible. This section explores additional factors influencing method selection, such as practical constraints and the study's relevance to current SE problems. Emphasis is given to trade-offs inherent in these choices and how they impact the decision-making process, ensuring researchers can balance theoretical aims with practical realities effectively.

Here, we would like to encourage researchers and students to be aware of such factors and make them explicit when arguing and discussing the choice of research methods, as there may be practical reasons for adjusting the choice of methods besides

the research questions. They, however, should be the main driver and implications of deviations need to be well founded.

# 6.1 Methodological implications

Research methods vary in their focus. Surveys, for example, provide breadth, capturing wide-ranging data from large samples, ideal for general trends. In contrast, case studies offer depth, providing detailed insights into specific scenarios or organizations. Researchers must decide between a broad overview of a topic and an in-depth understanding of a particular aspect.

Research methods also differ in how obtrusive they are to participants and their applicability in specific contexts. For instance, ethnographic studies are less obtrusive, allowing observation of natural behaviors, but are highly context-specific. Surveys, while more obtrusive, are less context-dependent and can be applied more universally [19].

Context includes environmental factors, organizational settings, and specific conditions under which the software engineering practices are carried out. Understanding the context helps in designing studies that are not only theoretically sound but also practically applicable. This involves considering specific nuances of the software development environment and tailoring research methods to suit these unique characteristics. Contextual awareness ensures that findings are reflective of, and applicable to, the real-world scenarios they intend to address.

Finally, relevance is a critical factor in ESE research. It pertains to the significance and applicability of the research to real-world problems. A relevant study addresses current challenges or gaps in SE practice, thereby contributing meaningful insights. Researchers must continually assess the relevance of their work, ensuring it aligns with current industry trends, technological advancements, and practical needs.

#### 6.2 Practical implications

Wohlin and Aurum [22] pointed out that factors beyond methodological ones may also influence the method choice. As an example, they present a case where a student chooses a method based on the competence of using the method. Effective research design requires acknowledging practical limitations, such as time and resources. Researchers are encouraged to realistically evaluate their capabilities against the demands of the chosen methods.

Data availability also plays a crucial role; for instance, case studies demand detailed data from specific organizations, while extensive surveys need access to large audiences. Obviously, conducting a case study requires good connections to industry. Also, we may not always be able to obtain the sample in surveys we ideally would like to have. Practically, random sampling would be desirable, but in

software engineering the most common approaches for sampling are purposive and convenience sampling [1].

Choosing a method that demands extensive time and resources may not be feasible for a small team with limited funding. For instance, longitudinal studies evaluating the long-term effects of implementing new practices, require significant investment in time and personnel. Similarly, opting for depth, such as conducting detailed ethnographic studies to understand developer interactions within teams, may limit the generalizability due to their focus on specific environments.

The trade-offs between these elements significantly impact the research design decision-making process. Flexible designs, in particular, allow for adjustments in response to changing circumstances, making the research more responsive to evolving data and insights. This flexibility is key for research topics where the scope is likely to shift, ensuring that studies reflect current trends. Researchers must carefully weigh these trade-offs to ensure their chosen methodology aligns with their research goals and practical constraints.

# 7 Summary

This chapter is a foundational guide for SE researchers on designing research proposals, setting the stage for further chapters in this book. It focuses on identifying relevant empirical problems, formulating targeted research questions, and choosing appropriate methodologies. It addresses challenges in method selection and classification within ESE and provides insights into various methods for data collection, and analysis techniques. The chapter aims to help researchers make informed methodology choices and understand their research's practical implications. It also includes interactive and engaging learning activities for a high-education setting. The following chapters will expand onto research methods and strategies, provide in-depth insights and hands-on guidance for teaching empirical research methods in SE.

#### 7.1 Further Reading

Other chapters of the book "Teaching Empirical Research Methods in Software Engineering" mentioned in this document are:

- Chapter 14 A Course on Experimentation in Software Engineering: Focusing on Doing, by Sira Vegas, Natalia Juristo
- Chapter 15 **Teaching Mining Software Repositories**, by Zadia Codabux. Fatemeh Fard, Roberto Verdecchia, Fabio Palomba, Dario Di Nucci, Gilberto Recupito
- Chapter 16 **Teaching Simulation as a Research Method in Empirical Software Engineering**, by Breno Bernard Nicolau de França, Nauman bin Ali, Dietmar Pfahl, Valdemar Vicente, Graciano Neto

- Chapter 17 **Teaching Action Research** by Mirosiaw Staron
- Chapter 18 Action Research with Industrial Software Engineering An Educational Perspective by Yvonne Dittrich, Johan Bolmsten, Cathrine Seidelin
- Chapter 19 Teaching Case Study Research by Stefan Wagner
- Chapter 20 **Teaching Design Science as a Method for Effective Research Development** by Oscar Pastor, Mmatshuene Anna Segooa, Jose Ignacio Panach
- Chapter 21 **Teaching Survey Research in Software Engineering** by Marcos Kalinowski, Allysson Allex Araújo, Daniel Mendez
- Chapter 22 **Teaching Literature Reviews in Software Engineering Research** by Sebastian Baltes, Paul Ralph
- Chapter 23 **Teaching Systematic Literature Reviews: Strategies and Best Practices** by Marcela Genero, Mario Piattini
- Chapter 24 **Teaching and Learning Ethnography for Software Engineering Contexts** by Yvonne Dittrich, Helen Sharp, Cleidson R. B. de Souza

In addition to this book's complementary chapters, readers are encouraged to explore Eastbrook et al. [8] for selecting empirical methods in SE research. Wohlin and Aurum's [22] decision-making framework, extensively referenced in this chapter, is a valuable resource we built upon. For comprehensive guidelines on various methods, Molléri, Petersen, and Mendes [13] offer an extensive compilation of methodological literature. Creswell's book [7], while more general, provides insightful guidance for research design across various fields, aiding in formulating research proposals.

#### References

- 1. S. Baltes and P. Ralph. Sampling in software engineering research: A critical review and guidelines. *Empirical Software Engineering*, 27(4):94, 2022.
- V. R. Basili, G. Caldiera, and H. D. Rombach. The goal question metric approach. *Encyclopedia of software engineering*, pages 528–532, 1994.
- 3. I. Benbasat and R. W. Zmud. Empirical research in information systems: The practice of relevance. *MIS quarterly*, pages 3–16, 1999.
- J. Biggs, C. Tang, and G. Kennedy. Teaching for quality learning at university 5e. McGraw-hill education (UK), 2022.
- P. Bourque and R. E. Fairley. Swebok v3. 0: Guide to the software engineering body of knowledge. *IEEE Computer Society*, pages 1–335, 2014.
- A. Cooke, D. Smith, and A. Booth. Beyond pico: the spider tool for qualitative evidence synthesis. Qualitative health research, 22(10):1435–1443, 2012.
- J. W. Creswell and J. D. Creswell. Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications, 2017.
- S. Easterbrook, J. Singer, M.-A. Storey, and D. Damian. Selecting empirical methods for software engineering research. *Guide to advanced empirical software engineering*, pages 285–311, 2008.
- 9. V. Garousi, M. Felderer, and M. V. Mäntylä. Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and software technology*, 106:101–121, 2019.
- S. B. Hulley, S. R. Cummings, W. S. Browner, D. G. Grady, N. Hearst, and T. Newman. Conceiving the research question. *Designing clinical research*, 335, 2001.

- 11. V. B. Kampenes, B. Anda, and T. Dybå. Flexibility in research designs in empirical software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering (EASE)*. BCS Learning & Development, 2008.
- 12. B. A. Kitchenham, D. Budgen, and P. Brereton. *Evidence-based software engineering and systematic reviews*, volume 4. CRC press, 2015.
- J. S. Molléri, K. Petersen, and E. Mendes. Cerse-catalog for empirical research in software engineering: A systematic mapping study. *Information and Software Technology*, 105:117–149, 2019.
- 14. K. Petersen, J. Börstler, N. bin Ali, and E. Engström. Revisiting the construct and assessment of industrial relevance in software engineering research. In 1st International Workshop on Methodological Issues with Empirical Studies in Software Engineering (WSESE'2024), 2024. accepted.
- K. Petersen and C. Wohlin. Context in industrial software engineering research. In 2009 3rd international symposium on empirical software engineering and measurement, pages 401

  404. IEEE, 2009.
- A. Rainer and C. Wohlin. Case study identification: A trivial indicator outperforms human classifiers. *Information and Software Technology*, page 107252, 2023.
- P. Ralph, N. b. Ali, S. Baltes, D. Bianculli, J. Diaz, Y. Dittrich, N. Ernst, M. Felderer, R. Feldt, A. Filieri, et al. Empirical standards for software engineering research. arXiv preprint arXiv:2010.03525, 2020.
- 18. C. Robson and K. McCartan. Real World Research. Wiley, 2016.
- 19. K.-J. Stol and B. Fitzgerald. The abc of software engineering research. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 27(3):1–51, 2018.
- A. Tattersall. Who, what, where, when, why: Using the 5 ws to communicate your research. Impact of Social Sciences Blog, 2015.
- 21. C. Wohlin. Case study research in software engineering—it is a case, and it is a study, but is it a case study? *Information and Software Technology*, 133:106514, 2021.
- C. Wohlin and A. Aurum. Towards a decision-making structure for selecting a research design in empirical software engineering. *Empirical Software Engineering*, 20:1427–1455, 2015.