Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

Magnus Tvilde

# Peer Code Review in Secondary Education: A Systematic Literature Review and an Interview Study

Master's thesis in Natural Science with Teacher Education
Supervisor: Monica Divitini
December 2020

NTNU
Norwegian University of Science and Technology

Magnus Tvilde

# Peer Code Review in Secondary Education: A Systematic Literature Review and an Interview Study

Master's thesis in Natural Science with Teacher Education
Supervisor: Monica Divitini
December 2020

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

# Abstract

Peer review is commonly used in academia, and is also being introduced into educational settings. Code review is a common practice in the software industry where colleagues review each others code. Both practices have known motivations, benefits, challenges and best practices. Combining these two into an educational setting we get peer code review, which has a lot of similarities with the two aforementioned practices. This research focuses on what is known about peer code review in secondary education, and how secondary school programming teachers could implement the method into their own practice. Especially understanding which challenges teachers face in implementing the method, and how it can be implemented in the best way possible, would be of value to the computer science education community. The research has conducted a systematic literature review to find out what the current status of peer code review in secondary education is. This included how it is used, motivations, benefits, and challenges. Further, to explore the real-world aspect eight secondary school programming teachers were interviewed. The interviews covered the status of peer code review in secondary education, the teachers thoughts about the method and how to support them in using it. From the systematic literature review it was found that the knowledge about peer code review in secondary education is very scarce, with a systematic review of the literature resulting in only four articles. The results from the interviews indicate that teachers easily relate to peer code review, but are not using it. There are many benefits and challenges with the method, but these challenges are possible to face with well thought out implementation strategies. There is need for empirical evidence of use of peer code review, and further research into its best practices.

# Sammendrag

Fagfellevurdering er ofte brukt i akademia, og blir også introdusert inn i undervisningskontekster. Kode-tilbakemelding er en vanlig praksis i IT-industrien der kollegaer gir tilbakemelding på hverandres kode. Begge praksiser har kjente motivasjoner, fordeler, utfordringer og beste praksiser. Kombinasjonen av disse to i en undervisningskontekst blir medstudentvurdering i programmering, som har mange likheter med de to nevnte praksisene. Denne studien fokuserer på hvilken kunnskap som er kjent om medstudentvurdering i programmering på ungdomsskolen og videregående skole, og hvordan programmeringslærere i disse trinnene kunne implementert metoden inn i deres egen praksis. Spesielt ved å forstå hvilke utfordringer lærer møter om de implementerer metoden, og hvordan den kan bli implementert på best mulig måte, det ville ha vært verdifullt for informatikkdidaktikk fagfeltet. Studien har gjort et systematisk litteratursøk for å finne ut den nåværende statusen for medstudentvurdering i programmering på ungdomsskolen og videregående skole. Dette inkluderte hvordan det blir brukt, motivasjoner, fordeler, og utfordringer. For å utforske aspektet i den virkelige verden har åtte lærere på ungdomsskolen og videregående skole blitt intervjuet. Intervjuene dekket statusen av medstudentvurdering i programmering på ungdomsskolen og videregående skole, lærerenes tanker om metoden og hvordan de kan bli støttet i å bruke den. Fra det systematiske litteratursøket ble det funnet ut at kunnskapen om medstudentvurdering i programmering på ungdomsskolen og videregående skole er veldig knapp, med et systematisk søk i litteraturen som endte opp med kun fire artikler. Resultatene fra intervjuene indikerer at lærere lett relaterer til medstudentvurdering i programmering, men ikke bruker det. Det finnes mange fordeler og utfordringer med metoden, men disse utfordringene er mulige å møte på gode måter med godt gjennomtenkte implementeringsstrategier. Det er behov for empirisk kunnskap om bruk av medstudentvurdering i programmering, og videre forskning på hva dets beste praksiser.

# Preface

Writing a master's thesis is hard work! Who knew? Despite this I got good support from people around me! First, I want to extend a warm thank you to my supervisor Monica Divitini. She has straightened my course whenever I needed, replied to emails very quickly and set up meetings with short time in advance. She has been tremendously helpful, thanks a lot! Secondly, I would like to thank the respondents of my interviews, who kindly participated in this research project. Thirdly, a big thanks too The Department of Computer Science at NTNU for allowing me to write my thesis about what I wanted, and for the lifesaving help I got with transcribing my interviews. Last, I have to thank my study buddies Ole-Martin, Sven and Vegard for keeping me in good spirits throughout the entire period, always supporting me. I hope this thesis can be of some value to the computer science education community.

# Contents

# Abbreviations

| | | |
|---|---|---|
| **CS** | = | Computer science |
| **PCR** | = | Peer code review |
| **SLR** | = | Systematic literature review |

# Chapter 1

# Introduction

This research project has explored both the status of peer code review (PCR) in secondary schools and attempted to find out how to support programming teachers in adopting the method. First a systematic literature review (SLR) regarding PCR in secondary education was conducted to find out what is known about PCR in secondary education and to which degree it might be used. Then I further explored how teachers relate to PCR and how to support them.

Computer science (CS) is a growing field and is becoming increasingly more important, evidently in the school system. This is true in everyday life, but maybe even more in schools. The Norwegian sector of higher education had a record number of applications in 2020 with the field of CS gaining a significant increase in applicants, all the while the Norwegian government has given more money to the field since 2016 (Kunnskapsdepartementet 2020). In the Norwegian context of secondary, and primary, schools the field of programming, which is encompassed by CS, has become much more prevalent after *Fagfornyelsen* (Utdanningsdirektoratet n.d.). Fagfornyelsen is a complete overhaul of the national curriculum, and aims to keep the school system relevant. Programming was in it being introduced into different courses such as mathematics and natural sciences. Further, the Norwegian context has in recent years had an experimental course in lower secondary schools called *Programmering*(Utdanningsdirektoratet 2016), which now has gained a more permanent place (Utdanningsdirektoratet 2020). This all goes to show that programming practices in the school system should be researched to create for better learning, and better educating and preparing up-and-coming generations in CS for the reality that will meet them when they have completed school.

There is a need for teachers to "develop appropriate teaching approaches (...) and implement them in their classroom teaching of programming" (Sabarinath and Quek 2020, p. 3554). Programming in education brings many benefits, but also challenges. Programming is considered challenging to learn for students (Sabarinath and Quek 2020, p. 3554). Students face challenges such as difficulty in reading code, writing code, writing programs, trace through codes, and grasping programming concepts (Sabarinath and Quek 2020, p. 3554). For teachers

one of the challenges is the scarcity of good resources (Sentance and Csizmadia 2017).

## 1.1 Problem definition

Peer review is a strategy widely used in academia. It has been used for a long time, and is seen as a best practice (Blum et al. 2018). Peer review is the assessment of a written piece of work, by peers of its author (Allen et al. 2019, p. 164). Peer review is possible to introduce into an educational setting, and implies that students evaluate each other. This can introduce a range of benefits to both students and the teacher, such as more timely feedback to the author and learning by the reviewer (Topping 2016, pp. 254–255).

Peer reviews done with code is in the software industry called code review. Code reviews are a common practice in the software industry (Fu et al. 2017). A code review is the process of giving ones code to a colleague, having them go through the code checking for defects and code quality. Code reviews in the software industry improves the code through finding errors, defects and bugs, improves improves project efficiency and sets the participants up for better knowledge transfer and collaboration (Rigby and Bird 2013, p. 210).

PCR is the combination of peer review and code review, when implemented into an educational setting. It has the with the potential of bringing all the benefits of peer review and code review, and add new ones. In higher education, many of these motivations, benefits, and challenges are also reported, but with slightly more emphasis on learning throughout the process. PCR has the potential of knowledge development, creating support for learning, improving code quality, improving review skills, making for a effective process for the teacher and creating social benefits (Indriasari et al. 2020, pp. 12–13). Active learning is a central aspect of how the method works in higher education (Aalberg and Lorås 2018), giving the students the ability to participate actively in their learning as opposed to being passive receivers of knowledge. Moreover, it prepares students for how things are done in the software industry, gives practice in giving and receiving constructive criticism (Anewalt 2005, p. 150).

In general, PCR and code review has many benefits, but its hard to generalize to all educational settings. Different motivations, benefits, and challenges are reported in different areas of use. PCRs and code reviews are also implemented differently in different settings, although there do exist some research on what the best practices are and recommendations when implementing the strategies.

This has motivated me to research the status of PCR in secondary education. Given that there exists a lot of research on code reviews in the software industry, and a good amount of research on PCR in higher education, I wish to find out more about PCR in secondary education. There are many aspects to find information about PCR in secondary education. It might be that the motivations, benefits, challenges and best practices are transferable from higher education, into secondary education. The things I want to find out are what literature exists about PCR

in secondary education, what is the status of PCR in secondary education, what are the motivations, benefits and challenges of using PCR in secondary education and what are the best practices to adopt when implementing the method. I wish to find out if PCR should at all be used in secondary education, and why. Moreover, if it is beneficial to use in secondary education, I wish to find out how secondary school teachers can be supported in using PCR. This includes how PCR can be introduced into secondary education.

## 1.2 Research questions

With the motivations for researching the problem as described in Section 1.1, there is a need for research questions to address the problem. First, I have created a concise problem statement:

> *What is known about PCR in secondary education, and how can we support teachers in implementing it?*

The issue at hand is divided into two main parts; the literature, and the real-world context. The research questions will mirror this division, creating a logical flow of the research. The research questions are as follows:

> **RQ1:** What is the current knowledge about PCR in secondary education?
>
> > **RQ1.1:** What is the state of the art in the research about PCR in secondary education?
> >
> > **RQ1.2:** What do secondary school programming teachers know about PCR?
>
> **RQ2:** How to support secondary school programming teachers in using PCR in their practice?

RQ1 is designed to address the current knowledge of PCR in secondary education. RQ1.1 takes the literature aspect of it and addresses the perceived knowledge gap found in the initial stages of the research project. To my knowledge, there is very limited research on the topic, which the research question is designed to further explore. RQ1.2 takes the real-world aspect of it and addresses how secondary school programming teachers view PCR. It both encompasses their knowledge of PCR, and PCR's current status in secondary schools. In this way, RQ1 addresses the first part of the problem statement above.

RQ2 is designed to create a real-world grounding for how programming teachers view PCR in secondary education. To properly address this research question, the status of PCR in secondary education in a real-world context needs to be addressed. It also needs to address the motivations, benefits, and challenges, as well as which choices to be made to implement PCR into secondary school. RQ1.2 also somewhat helps to address this issue, as they are somewhat connected. In this way, RQ2 and RQ1.2 addresses the second part of the problem statement above.

## 1.3 Methodological approach

To explore the research questions posed, a research strategy was developed for each of the research questions. The research strategy contains how the research question will be provided answers to, through the method of data collection and the analysis (Robson and McCartan 2016, p. 72). The strategies are connected in that the results from RQ1.1 will have indications of how to research RQ1.2 and RQ2.

To explore RQ1.1, about what the current knowledge of PCR in secondary education is in the literature, a SLR has been conducted. The SLR followed the guidelines of (Kitchenham 2004). This review took RQ1.1 as a base and developed multiple sub-questions to help provide a comprehensive and systematic answer. Five highly relevant databases were searched with a carefully chosen search string. Initial results of over 2021 were run through a selection process with inclusion and exclusion criteria, leaving four primary papers. Data was extracted from these four papers, summarizing the current knowledge in the literature on PCR in secondary education. These results gave further motivation to research RQ1.2 and RQ2, as knowledge is scarce.

To find answers to RQ1.2 and RQ2, qualitative research was carried out. The method was qualitative because it allows for the focus of the research to be on meaning, taking the context into account and describe the situations from the perspective of the teachers. To answer the real-world research question, the method of choice was semi-structured interviews. These interviews allows me as a researcher to understand the respondents perspective of themes and topics from their daily lives (Kvale and Brinkmann 2015, p. 46). The qualitative interviews were done due to the constraints of the project, but more so due to the way qualitative interviews allow for the in-depth exploration of the research question. Eight secondary school programming teachers were interviewed, over the timespan of just over two weeks. These interviews were transcribed and analyzed with thematic coding to create the results of the interviews. Thematic coding is the process of marking different parts of the text with different codes, and putting those codes into overarching themes (Robson and McCartan 2016, pp. 468–480). The coding of the thematic analysis was done in NVIVO.

## 1.4 Contribution the study makes

This project has two main contributions.

The first contribution is the results from the SLR, summarizing the current knowledge on PCR in secondary education and identifying a gap in the literature. From the research questions, I have found that there is a lack of literature on the subject and indications that PCR is a promising method. The perceived knowledge gap on PCR in secondary education is indeed a knowledge gap. The existing knowledge is very scarce, and mostly provides insight into the benefits of PCR in secondary education.

The second contribution is the empirical evidence that teachers show a positive attitude towards the method, and see it as promising. This contribution also discusses the way the method should be implemented, especially with concern to the challenges facing the implementation, helping with understanding the potential and challenges connected to PCR in secondary education.

They also have many thoughts on how to implement the method, providing a good starting point for implementation. It is, however, somewhat teacher dependent and requires the teacher to be confident in their own ability to conduct the method.

## 1.5 Structure of the report

The structure of the report is presented here. In this chapter one gets familiar with the contents of the report. The next chapter contains a conceptual framework to set a common understanding of the subject concepts and to clearly define what work previously has been done. The chapter afterward describes how the SLR was carried out, with its findings reported in Chapter 4. Chapter 5 describes how the data collection for RQ2 was carried out, and Chapter 6 how the data was analyzed. Chapter 7 reports the findings. In Chapter 8 the results of the interviews are put into context and discussed. The chapter after discusses the ethical aspects of the research. Finally, there is a conclusion in Chapter 10, followed by references and appendices. Throughout the report, chapters are referred to as the main parts of the report, while sub-chapters are referred to as sections, e.g. Chapter 1, Section 1.1. and Section 1.1.1.

# Chapter 2

# Conceptual framework

In this chapter a conceptual framework that will be used in the report is presented. The chapter first elaborates what peer code review is and gives a definition that is to be further used in the study, then elaborates on peer code review in an educational context and presents some main implementation strategies.

## 2.1 Peer code review

Peer code review (PCR) stems from the combination of peer review and code review put into an educational setting. It is motivated by all of their respective promised benefits, and is a term that yet seems to not have one clear definition. In the following subsections the background of the PCR process referred to in this report is explained.

The concept of peer reviewing has been around for a long time and is about reviewing others' work. Peer review is a somewhat broad term, referring to any assessment of a written piece of work, by loosely defined peers of the author (Allen et al. 2019, p. 164). It is a common practice in academia, being regarded as a best practice in scientific literature (Blum et al. 2018).

Peer reviews with focus on coding, which we refer to as code review, has a prevalent part of the software development practice in the software industry (Badampudi, Britto and Unterkalmsteiner 2019). This started as software inspection, first coined by Fagan (1976), which is the process of inspecting code looking for errors. This process has evolved since then into what we today describe as code review. Modern code review is another term found in the literature, which has more of an emphasis on "elegance, logic, complexity, etc., team awareness, finding alternative options, interactions, and knowledge sharing" (Fatima et al. 2019, p. 2).

Code review has since become a regular part of the software development practice, with goals such as identifying defects, bugs or errors, improving code quality and overall project efficiency, knowledge transfer and collaboration among participants, and seeing alternate solutions to a given problem (Rigby and Bird 2013, p. 210), thus learning from reading and reviewing code. What is described

as code review often encompasses reviewing a co-developers code and then sending it back to its author for it to be revised and improved. Code review improves software quality and knowledge sharing (dos Santos and Nunes 2017, p. 28). The knowledge sharing itself reinforces all of these other benefits, being that it "decreases defect, cognitive load, rework, waiting, etc." in software engineering (Fatima et al. 2019, p. 1).

The code review used in industry has known challenges. Some of the main challenges found by MacLeod et al. (2018, p. 37): receiving timely feedback, review size, managing time constraints and understanding the code's purpose. In general, when adopting code review into a new development context, it might not be clear to those adopting the practice what challenges to expect and how to implement the practice (MacLeod et al. 2018, p. 34).

### 2.1.1 Definition of PCR

To get a precise definition of PCR to be used in this report, we take code review as base. PCR is by definition a variation of the standard code review process conducted in the software industry. It would seem from the literature that no one-size-fits-all approach is completely viable when defining code reviews. Li (2006) and Yanqing et al. (2011) are examples of how PCR is conducted differently in different studies. Their processes involve more steps than what we will have as our definition here. Rigby and Bird (2013, p. 203) state that the process of code reviewing consists of a variation of a fixed set of steps: "planning, overview, preparation, inspection, reworking and follow-up". Each paper seems to provide its own interpretation of the process, but they all generally follow the line of:

1. participant solves programming task and submits it to system
2. system designates reviewers, system distribute submissions, participants review provided submissions and submit their review
3. review is redistributed to its respective author and read

These three steps are what we define as PCR, when it is done in an educational setting. Outside of these three generalized steps, different studies implement different practices to fit their own needs. The different implementations are often made to meet different challenges, create various benefits or for a number of other reasons. In this paper we look at anything outside of this three-step process as an extension to the foundation of the PCR process. Those three steps are considered *critical* parts of the process, with any extension or addition being defined as *non-critical*.

## 2.2 PCR in education

Training novice programmers to perform code reviews of good quality with valuable feedback is a challenge (Indriasari et al. 2020, p. 3), which PCR introduction into programming subjects in school might help encounter. This is an interesting

topic that should be researched further, however the existing literature might be an indication that PCR has enough benefits to really be considered as a good, standard practice in the classroom. Challenges are subject to the education level of its participants, meaning any challenge that is reported in the industry or in higher education might not apply, or at least in the same way or extent, to PCR in secondary education.

Peer review in an educational context is relatable to terms that differ from a traditional lecture. PCR is a collaborative learning activity, also making the learning process social. PCR also has the characteristics of active learning (Prince 2004, p. 1). Active learning is beneficial to bring into the classroom, as it improves student attitudes, writing and thinking abilities, motivation for working, student engagement and promotes deep understanding of topics (Prince 2004, pp. 3–4). However, this is not something that just happens by itself, it needs to be done in a well-thought out process. Moreover, PCR is also closely tied too the concept of Contributing Student Pedagogy: "A pedagogy that encourages students to contribute to the learning of others and to value the contributions of others" (Hamer et al. 2008, p. 195). This is an attractive attribute for many reasons, one of them being the way it allows different participants to work on different levels, another one being that it closely relates to the contemporary practices in the software industry (Hamer et al. 2008, p. 207).

In an educational context, PCR seem to have a large set of benefits. Søndergaard & Muller (2012, p. 346) highlight that the adoption of such a process will create diverse, and timely feedback to all students. Timely in the educational setting is translatable to the system creating an increased capacity for providing feedback to students, as they themselves partake in that process. In a case study done by Li (2006, p. 3) it is reported that PCR motivates the learning of coding standards. Further stating benefits of the process is Anewalt (Anewalt 2005, p. 150):

- Preparing students for code reviews they are likely to encounter in their career, should they become software engineers
- Higher quality submissions
- Practice in giving constructive criticism
- Practice in receiving constructive criticism

Indriasari et al. (2020, p. 13) report a wide variety of benefits in an SLR on PCR in higher education:

- Ability to identify defects in their own code
- Exchange ideas on a problem with peers
- Seeing other solutions to a problem
- Ability to give and receive criticism
- Learning from feedback
- Ability to discuss code
- Collaboration with peers
- Code review skills
- Coding skills

- Time management skills
- Soft skills
- Problem solving skills
- Knowledge of best practices in coding
- Peer review knowledge

Challenges of the PCR process are important to consider when implementing it. The participation of reviewers Dos Santos & Nunes (2017, p. 28) say is negatively affected mainly by the size of the submission to be reviewed. This indicates that the reviews get better quality if they are assigned to smaller sized code snippets.

Indriasari et al. (2020, p. 15) also report the main challenges facing the implementation of the process:

- Lack of knowledge or ability to perform reviews
- Low level of student engagement in the review process
- Low quality of produced reviews
- High administrative burden on facilitator (teacher)
- Time restrictions, too large amount of code to review for students

## 2.3 PCR strategies

In this section, different strategies for the implementation of PCR are provided. When talking of strategies in this context, any non-critical step of the PCR process is to be viewed as part of a strategy, i.e. any extension to the three steps, see Section 2.1.1. Only the imagination sets limits to what can be implemented, meaning we cannot discuss all available strategies in this section. However, a few significant strategies set themselves out in the literature: having a structured review process, having a clear strategy of reviewer designation, and the implementation of assessment in addition to the feedback of the review.

### 2.3.1 Structured review process

One strategy one can decide to implement along with the critical steps of the process, is adding some structure to the review process. This can be done e.g. with a rubric of some sort, a checklist, or other related methods. A checklist is stated by Almeida, Camargo, Basseto, & Paz (2003, p. 57) as a best practice when limited to one page in length, as long checklists are hard to use. It provides several benefits such as listing the fault types or symptoms to look for, helping maintain good coding standards and increasing the review's effectiveness. It is possible to do this in both industrial settings and educational ones.

### 2.3.2 Reviewer designation

Dos Santos & Nunes (2017) elaborate on different studies choosing different strategies of how the reviewer designation is to be done, i.e. which participant reviews which submission, this work is however somewhat irrelevant to an educational setting as it emphasizes timing and efficiency in a software industry setting. However, a relevant contribution to be used in this paper is their emphasis that appropriate reviewers are designated to appropriate tasks (dos Santos and Nunes 2017, p. 92) which is different for each context the PCR process is implemented into. This strategy is used both in industrial settings and in educational ones. In an educational setting it implies that students with the appropriate competence should be designated to appropriate tasks, i.e. that one should receive submissions not too far away from one's competence level.

### 2.3.3 Assessment

In this paper, we differentiate between an assessment and a review. A review is viewed as a broader term, thus encompassing an assessment with a score, grading or classification by numbers of some sort, while a review focuses on the total feedback to each participant of the process. This strategy is only applied, and also fits quite well, in educational settings, as the review is enough in industrial settings. Assessment is viewed as a non-critical part of the PCR process, being treated as a viable extension in the research.

# Chapter 3

# Methodological approach: Systematic literature review

This chapter regards the SLR conducted on PCR in secondary education. It covers the objective of the review and how it was conducted.

## 3.1 Objective of review

The review has an overall goal of systematically laying a foundation to build on, regarding PCR in secondary education. It seeks to facilitate the planned interviews, and most importantly to address the perceived knowledge gap that exists in the body of literature. The review will be a critical step towards better understanding the status quo of PCR in secondary education. This encompasses PCR use, PCR implementation strategy and PCR challenges and benefits, as well as why teachers would or should have the motivation to implement the method in their practice. To understand the contemporary role of PCR in secondary education, the following research question is in focus.

**RQ:** What is the current knowledge about Peer Code Review (PCR) in secondary education?

Although this question broadly covers what the SLR encompasses, the following sub-research questions complement the main RQ to make for a more structured approach to finding relevant information.

**Sub-questions to answer:**

- What definition is made of PCR in the literature connected to secondary education?
- Is PCR used in secondary education?
- How is PCR used in secondary education?
- What is the motivation behind using PCR in secondary education?
- What are the known benefits of using PCR in secondary education?
- What are the known challenges of using PCR in secondary education?

**Table 3.1:** Databases

| Databases | Search fields |
|---|---|
| ACM Digital Library | Search in any field |
| IEEE Xplore | Search in metadata |
| ScienceDirect | Search in any field |
| Scopus | Search in title, abstract and keywords |
| SpringerLink | Search in all fields |

- Are there any reported differences in using PCR in tertiary education to secondary education?

This review is the first SLR on PCR in secondary education, to the best of my knowledge. It is also only the second SLR on PCR in education, following that of Indriasari et al. (2020), to the best of my knowledge.

## 3.2 Review method

### 3.2.1 Protocol development

The protocol for performing the systematic review was developed on the basis of how Indriasari et al. (2020) performed their SLR regarding PCR in higher education, because the subject at hand is closely linked to that of their review. The review follows Kitchenhams (2004) guidelines of how to perform SLRs. The protocol was defined before the process started, but was slightly modified while conducting the review. There were changes made to the strategy for data extraction and data synthesis, as when reading through the articles themes became prevalent.

### 3.2.2 Data sources

Databases with relevant literature on the given subject are ones that have a connection to CS. As each of the databases have slightly different options for searching, we simply list how the search was conducted in each database. The chosen fields were standard for each database, but after evaluating how many more or less results a different approach would yield, the standard search strategy of each database was kept. An overview of the databases searched in is found in Table 3.1, along with their respective search fields.

### 3.2.3 Search strategy

#### Identify relevant studies by search

The keywords searched for must be prepared in a way that captures as much as possible of the relevant literature, as well as excludes as many uninteresting

**Table 3.2:** Search terms

| # | **Search term** |
|---|---|
| 1 | Peer code review |
| 2 | Peer code assessment |
| 3 | Modern code review |
| 4 | Code review |
| 5 | Peer |
| 6 | Student |
| 7 | Education |
| 8 | Learning |
| 9 | School |

articles as possible. Starting out the identification of a string to search with was simply “peer code review”. Synonyms such as “Peer code assessment” and “Modern code review” were added, and “code review” to capture a few more results. These were identified from doing initial searches, mapping out what kinds of keywords could be interesting. Different studies used somewhat different terms I found, which made me include the synomyms. To limit the results to ones relevant to education the words “Student”, “Education” and “Learning” were added. The search-string went through multiple iterations in different databases before being decided. To capture as much of the relevant literature as possible, the following keywords were used as specified. As the field is prone to new knowledge and new technology, this review limits its scope to the period 2005-2020.

How the search terms from Table 3.2 were combined: (1 OR 2 OR 3 OR (4 AND (5 OR 6)) AND (7 OR 8 OR 9)

The search was run in databases ACM Digital Library, IEEE Xplore, ScienceDirect and Scopus on October 9, 2020, and in the SpringerLink database on October 13, 2020, due to difficulties exporting the search results. This search across the yielded 443 results from ACM Digital Library, 23 results from IEEE Xplore, 603 results from ScienceDirect, 92 results from Scopus and 860 from SpringerLink, totaling 2021 results.

### Selection process

The selection process consisted of three stages. The stages were designed to go more in-depth into each paper the further into the process. Figure 3.1 illustrates the selection process and the number of articles in contention at each stage.

1. Removal of duplicates and non-English papers
   The first step consisted of removing all duplicates and excluding any paper not written in English. Starting with 2021 papers, applying the language filter gave 2014 results. The removal of duplicates was done in an external excel-spreadsheet where the metadata of the papers was exported to. The

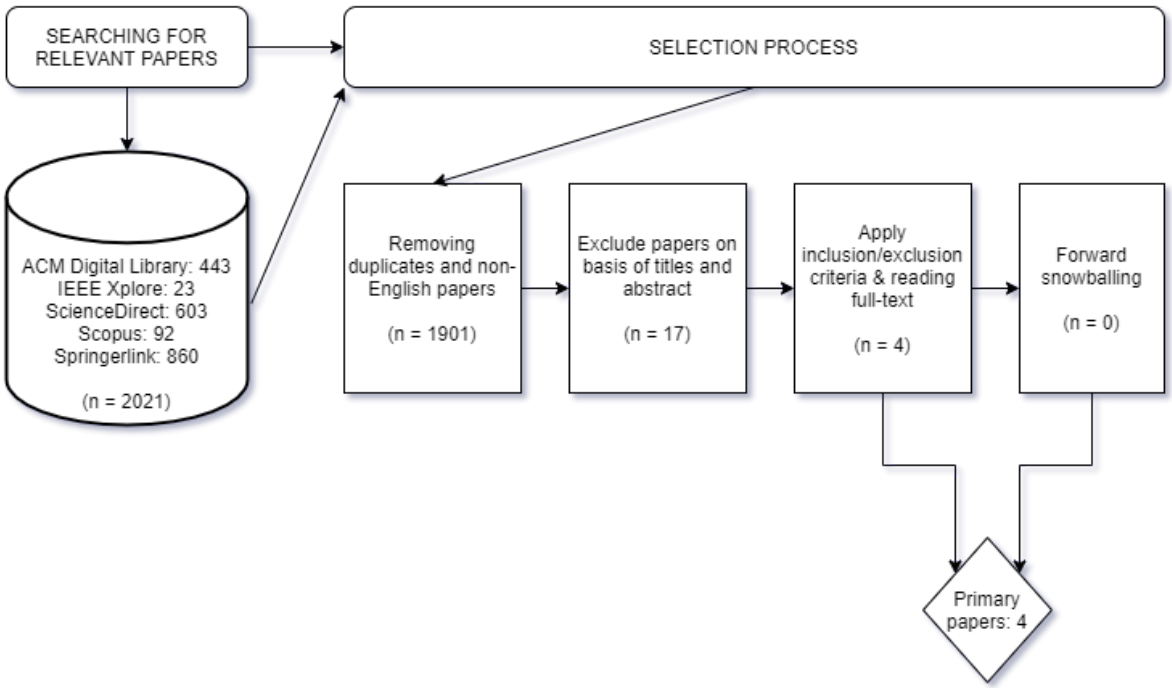


**Figure 3.1:** Selection process

metadata consisted of the information specified in the data extraction form, see Table 3.3. Excluding 113 results left 1901 papers.

2. Exclude on basis of titles and abstracts
   The second step consisted of reading all the 1901 papers' titles and abstracts. Any paper where its title was clearly not connected to PCR or software education in some way was excluded. If the title remained relevant to the subject at hand, the abstract was read. Any paper that did not mention anything about PCR in secondary education, or did not seem very promising, was excluded. This process excluded 1884 papers, leaving 17 papers. To ensure reliability of the process, it was quality assured by having two assessors read

**Table 3.3:** Data extraction form

| Study description | | |
|---|---|---|
| 1 | Publication year | Year in YYYY format |
| 2 | Authors | Last name, first name; |
| 3 | Title | Title |
| 4 | Abstract note | Abstract of article |
| 5 | Pages | What selection of pages belong to article, if relevant |
| 6 | Type of article | Book section, conference paper, journal article, workshop paper |

**Table 3.4:** Inclusion and exclusion criteria applied

| Inclusion criteria | Exclusion criteria |
|---|---|
| Written in English | Not written in English |
| Accessible | Not accessible |
| Peer reviewed | Not peer reviewed |
| Mentions peer code review in lower education | Does not mention peer code review in lower education |
| Only one copy of title | Duplicated titles |
| Provide detailed information about peer code review | Does not provide detailed information about peer code review |

through 100 of the papers and checking whether the criteria was applied rightfully. Of the 100 papers, two were agreed to include, two were only included by one assessor and the last 96 were excluded by both assessors making the inter-rater reliability 98%. Using the Cohen Kappa statistic (Cohen 1968), it paints a somewhat bleak picture of this high percentage of agreement, giving a coefficient of 0.66 that indicates "Substantial agreement". This is due to the lack of papers that are included, making the coefficient a somewhat unreliable choice, leaving the high percentage of agreement as our best means of reliability.

3. Apply inclusion/exclusion criteria
   Any paper that does not meet all inclusion criteria, or matches any of the exclusion criteria 3.4, was excluded. This process excluded ten more papers, leaving 7 papers.
   Some of the criteria was not possible to check when only reading title and abstract, meaning the full text had to be checked. This excluded three more papers. After this forward snowballing was used to see if the papers pointed us to any relevant literature, but this part gave no extra primary papers. Using these selection stages, four papers were selected as primary papers and a total of 2017 results have been excluded. The list of the primary papers are listed in Table 3.5 with authors, publication year, title and a given ID.

### 3.2.4 Obtain papers, data extraction and synthesis of findings

#### Quality assessment

Each of the four papers that remained after the selection process were assessed by the author, according to ten criteria. These ten criteria cover the papers rigour, credibility and relevance. The assessment form of Critical Appraisal Skills Programme (CASP n.d.) for qualitative studies was used, as it fit the primary papers.

**Table 3.5:** Primary papers

| ID | Publication year | Authors (sorted by) | Title |
|---|---|---|---|
| S1 | 2018 | Kubincová, Z.; Csicsolová, I. | Code review in high school programming |
| S2 | 2020 | Kubincová, Z.; Csicsolová, I. | Code Review at High School? Yes! |
| S3 | 2020 | Kubincová, Z.; Demková, I. | Using Code Review at School and at the Programming Club |
| S4 | 2010 | Meerbaum–Salant, Orni; Hazzan, Orit | An Agile Constructionist Mentoring Methodology for Software Projects in the High School |

### Data extraction from remaining papers

Data extraction was done with a form from the four remaining papers, to find information relevant to the research questions for the review. The full-text papers were read, and relevant information to the research questions was extracted. The form was based around the seven formulated sub-questions to help answer the defined research question. The points to fill out in the data extraction form were the following bulletpoints:

- Definition of PCR (any definition that is made of PCR in the paper)
- Reported use of PCR used in secondary education? (yes/no)
- Reported level of education of participants (primary/lower, secondary/lower, or tertiary/higher education)
- Reported motivation for use of PCR
- Reported benefits of using PCR in secondary education
- Reported challenges of using PCR in secondary education
- Reported differences in using PCR in tertiary education to secondary education (if explicitly mentioned)
- Reported tools for implementing PCR
- Reported strategy for implementing PCR (steps of process, other strategy for conducting PCR)
    - Systematic or informal use of PCR in secondary education (systematic/informal)
    - Anonymity (yes/no)
    - Structured review process (existence of rubric, yes/no)
- Other relevant findings

### Synthesis of findings

There were some themes brought into the data extraction process, in the research questions. The research was open to finding new themes, but it was primarily a

deductive process. The answers to the questions from the form is how the findings are presented in this report.

### 3.2.5 Tools used

#### Zotero

Zotero (Zotero n.d.) was used as a tool to sort and order the papers. The search results were exported as RIS and BibTeX, dependent on the options given by the databases. These were opened in Zotero, in such a way that all results across the different databases were merged together. Removal of duplicates were then done in Zotero. This, however, did not remove all duplicates. This led me to export all results into a spreadsheet where duplicates were located and removed.

# Chapter 4

# Findings from systematic literature review

This chapter presents the findings of the SLR described in Chapter 3. The SLR seeks to answer the following RQ, with the sub-RQs defined to make for a better way to conduct data extraction.

**RQ1.1:** What is the state of the art in the research about PCR in secondary education?

**Sub-RQs:**

- What definition is made of PCR in the literature connected to secondary education?
- Is PCR used in secondary education?
- How is PCR used in secondary education?
- What is the motivation behind using PCR in secondary education?
- What are the known benefits of using PCR in secondary education?
- What are the known challenges of using PCR in secondary education?
- Are there any reported differences in using PCR in tertiary education to secondary education?

## 4.1 Impressions from search process

Looking through the 2021 initial search results the process of PCR seems to be closely linked to that of agile software development, which is a prevalent and somewhat new software development strategy. Many studies seemed linked to parts of the PCR process, yet they were mainly from the software industry and from tertiary education, namely CS1 courses in tertiary education.

## 4.2 General about the findings

The most important result from this systematic literature review is that there is very little relevant literature about PCR in secondary education, and only two of

the four selected primary papers report empirically on PCR implementation in secondary education. This confirms the perceived gap in the literature about PCR in secondary education, and motivates further research into the field. In one of the four papers PCR is not reported empirically at all, but only briefly mentioned the method as a part of a larger agile methodology for programming teaching in high school. The paper does not add a comprehensive and extensive contribution, however, it was added to the primary papers due to the severe lack of relevant literature on the topic. The other three papers are somewhat connected to each other, as one of the authors is a co-author on all three papers. One of these articles also has little contribution, as much of its relevant information merely points to one of the other two papers from the same author. Again, it was added anyhow because of the lack of comprehensive relevant literature.

## 4.3 Definition of peer code review

PCR is not a clearly defined term within the literature on its use in secondary education. The term is clearly motivated by the terms peer review and code review, which seem to be mentioned whenever the term is applied in secondary education. S1 and S3 suggest the definition of PCR to include a “small review” as the initial step of the method. The small review consists of a given code made by the teacher, where the pupils are to find defects, bugs and irregularities they can improve upon. This is a sort of training review, much like a code review, for the next step of the process which will contain a larger actual peer review. Two of the texts did not offer any definition of PCR, indicating that it was obvious to them what the term is defined as. PCR is termed Educational Code Review in S3.

## 4.4 Reported use of PCR in secondary education

All four articles discussed use of PCR, with their own interpretation of what PCR consisted of, in secondary education. Two of the articles report the actual implementation of the method in secondary education. The lack of literature, however, suggests that the method is not broadly known or used, at least in a systematic way. All of the papers refer to high school, suggesting that this method is not used, or at least not broady used, in education levels lower than that of high school. All four articles discuss PCR implementation into programming courses.

## 4.5 Strategy for implementation of PCR in secondary education

S1, S2 and S3 implement PCR with a “small review” first, and then a “project review”. Their definition of a project review is what correlates with PCR from the literature, while the small review serves as a way to introduce the broader concept

**Table 4.1:** Strategy for implementation

| Paper ID | Education level | Systematic or informal use | Anonymity | Structured review process | Tools for implementation |
|---|---|---|---|---|---|
| S1 | High school | Systematic | Y | N | Email |
| S2 | High school | Systematic | Y | Y | NA |
| S3 | High school | Systematic | NA | NA | NA |
| S4 | High school | NA | NA | NA | NA |

of reviewing written code. The small review consists of imperfect code written by the teacher and given to the students to comment on and improve. This step is intended to be time effective, and is to be done multiple times to train the students before moving to the project review. The project review consists of three phases, where it highly correlates with PCR in the literature as the three steps are 1) code and submit, 2) distribute, review and receive review, and 3) correct and improve reviewed code. An interesting result is that the given task was not the same for all students, meaning they most likely reviewed code of programs designed to do different things than their own. The teacher's role in the process was defined as a more passive watcher in S1, S2 and S3. The teacher was to comment on each of the three phases of the project review, as the teacher during the process monitored and gave an assessment at the end. S4 did not give any strategy for implementation.

As shown in Table 4.1, three papers report systematic and one NA. Two papers report an anonymous designation process, where the two others do not have data on this. One paper reports not having a structured review process, one reports having a structured review process, and the two last ones do not have any data supporting any of the two alternatives. Email was reported as a means of distribution in S1, whereas there was not a description of this in the other papers.

## 4.6 Motivation for using PCR as an educational practice

Lack of research on the field compelled researchers to explore this practice, as well as the promise of benefits from research on peer review. The promises of code review are also prevalent here, especially improving code quality and overall time effectiveness of a software development project. Another reported motivation is the idea of teaching skills needed as a future programming professional, such as reading others' code. It is also motivated by the teachers enhanced ability to track student progress.

**Table 4.2:** Benefits. Describes if each article reported the specific type of benefit, Y means reported, N means not reported

| Paper ID | Promote student learning | Improved code quality | Learning to read others code | Learning to give constructive feedback | Seeing other ways of solving a problem | Code reviewing skill | Teachers' tracking of student skills | Error detection skill | Positive student view of PCR |
|---|---|---|---|---|---|---|---|---|---|
| S1 | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| S2 | Y | Y | Y | N | Y | Y | Y | Y | Y |
| S3 | Y | Y | Y | Y | Y | Y | Y | Y | N |
| S4 | Y | N | N | N | Y | N | N | N | N |

## 4.7 Benefits of using PCR in secondary education

Benefits range from improved quality of code, via learned ability to give constructive criticism, to being able to read others' code as you can see in Table 4.2. In general the findings from the papers focus a lot on learning advantages of PCR, the students ability to improve the code in the process and why it improves the code. The papers also highlight other coding skills such as seeing alternate solutions to a problem, reading code and error detection. Less prevalent, but still important, is students learning to give constructive feedback and teacher tracking of student programming skills. The results generally indicate that most students were able to comment on other students' code.

## 4.8 Challenges of using PCR in secondary education

The reported challenges of PCR are not as focused in the papers as the benefits. This might be due to the lack of literature, as there is a logical implication that one should highlight the process' benefits, as opposed to challenges, when trying to embrace it. However, there are some reported challenges, see Table 4.3. These include the fine line the teacher has to find to create tasks that are not too difficult, while still not being too easy. Too difficult tasks makes it hard for students to review, which leads to low quality reviews, which again halts the entire process. Moreover, a concerning challenge is that students initially reported the process as something they had a negative attitude towards. They saw the process as extra work, and viewed reviewing as time-consuming and complex. This was perhaps mostly because many of these papers had different tasks for each student/student group, meaning they had to give feedback to a different problem than they had solved themselves. Another challenge was the view some students had that the received reviews did not provide anything useful.

**Table 4.3:** Challenges. Describes if each article reported the specific type of challenge, Y means reported, N means not reported

| Paper ID | Difficulty of task too high makes review hard | Initial student negative attitude towards feedback process | Time demanding process | Demanding administrative process for teacher |
|---|---|---|---|---|
| S1 | Y | Y | N | Y |
| S2 | Y | Y | Y | Y |
| S3 | N | N | Y | N |
| S4 | N | N | N | N |

## 4.9 Use of PCR in secondary education compared to tertiary education

No information to be found in the papers regarding this problem, the only one that might resemble anything is S3, but it does not compare the two levels of education.

## 4.10 Other findings

Due to the lack of literature, I have chosen to include other findings not necessarily connected to any of the above research-questions. One of these findings is that the studies indicate that PCR is applicable in secondary education. S3 reports that there is reported use of PCR in higher education, but no papers describing use of PCR in lower education.

# Chapter 5

# Methodological approach: Interviews

In conducting research one must make choices on how to address the research questions. The choices are to be made both for the purpose of choosing the optimal strategy for addressing the issues, but at the same time the choices have to be realistic. That means the choices made need to be feasible to the frames of the research. This project is a master's thesis, meaning there exists a constraint on the time frame. The research question at hand poses constraints that makes it too big of a project if I were to choose a quantitative method. Another aspect is that a quantitative method, e.g. a questionnaire sent to hundreds of suitable teachers in Norway, would make for a very hard recruitment process if the teachers just chose to ignore an email sent to them inviting them to participate in the study.

This study is designed as a qualitative study because of the evident lack of prior research on the subject. From the systematic literature review results presented in Chapter 4, we know that PCR in secondary education has a severely limited amount of relevant research - and that the research actually relevant is just starting to be conducted.

In this chapter the methodology of obtaining data is explained, and why the project has a qualitative phenomenological approach, with a constructionist view of knowledge. The chosen method for collecting data was qualitative interviews, which sought to answer the RQs of the project.

## 5.1 Qualitative research

To conduct a research study, one must start with background information about a topic, creating a problem statement, find research questions to help answer that statement and figure out a methodological approach that will be one's way of finding those answers (Robson and McCartan 2016, pp. 72–73). This study started with immersing myself in literature about PCR, realising that it was very hard to find research about PCR in secondary education. This made me want to find

out more about the topic, and I first did a systematic literature review on PCR in secondary education. The results show that there exists very little relevant literature on the topic, motivating for further research. To conduct this research I chose to plan a qualitative research design, as teachers' thoughts and reflections are more easily captured in their full intent in a qualitative design. Because of the research questions posed, the design of the study is a qualitative social research study. This allows the focus to be on meanings, taking the context into account and describing the situations from the perspective of those involved (Robson and McCartan 2016, p. 20). Typical quantitative features would not have suited this study, e.g. having a strictly deductive approach would have been sub-optimal for this research, generalization of the findings would have been hard, and the quantification of the results (Robson and McCartan 2016, p. 19) would have needed a large-scale data collection which was deemed to be outside of the constraints on the study.

The study is a *phenomenological* one, as it seeks to describe the use, thoughts and reflections of PCR in secondary education from the teachers' perspectives (Kvale and Brinkmann 2015, pp. 45–46). I want to understand and describe the way teachers themselves understand PCR, if they use it, if they use similar methods, what motivations for using it they might see, what benefits and challenges PCR brings, and at last what kind of support they would need in implementing PCR into their own practice. As a researcher I then must show a certain conscious naivety, although I have read literature about the topic, to allow the thoughts and reflections of the participants to show, as I seek to understand the way the teachers view PCR in its entirety and describe it thoroughly (Kvale and Brinkmann 2015, pp. 47–48).

Further, the study is built on social *constructionism*. My research questions seek to understand PCR from a teachers' perspective, and positions at a constructivist view of empirical knowledge. In social constructionism research meaning "is constructed by human beings as they interact and engage in interpretation" (Robson and McCartan 2016, p. 24). This is done in this research by attempting to understand each individual's understanding of PCR in secondary education.

Even though the choices of research study is done to best suit the research questions at hand, the choices also have some implications for how this study is built. The research has to stick to what is within the bounds of phenomenology and constructionism. Trying to understand what is the current knowledge about PCR in secondary education, which is RQ1, is done on the premises of how the teachers understand the phenomenon. There is scaffolding added to the research to get the teachers to a quick understanding of PCR, yet, they have to describe their own understanding of the different aspects of PCR. To understand how to support teachers in using PCR in their practice, RQ2, benefits, motivations and challenges are sought after. The teachers are free to express what they view as benefits, motivations and challenges. Even though these implications have an impact on how the study is built, it was deemed appropriate as the study is exploratory in nature, due to the evident lack of research on the field, see Chapter 4. The

RQs might not be easy to answer because of the explorative nature, but we do however have Indriasari et al. (2020) as a good base. Although it does not cover PCR in secondary education, it does provide somewhat of an insight into what are possible topics of interest, benefits and challenges of PCR in secondary education. The research cannot explore every aspect of PCR in secondary education thoroughly, and therefore narrows the extracted information to if the teachers have used PCR, their thoughts on PCR, especially with focus on benefits and challenges, their thoughts on how to meet the challenges, and if thought to be a viable option what could be done to support them in using the method. This is done by going in depth into these themes, rather than quantifying their responses and generalizing them. The research will however discuss, explore and take general findings from the data extracted, by being explicit about the risk of generalizing the results to a larger context. In this way the research can identify problematic areas with PCR, gain insight in those areas and discuss what some teachers might think of as good ways to implement the method into secondary education.

In qualitative research the researcher must be taken into account. As a researcher in qualitative research I have knowledge of the subject prior to the data collection and analysis, there is a social dimension that needs to be taken into account and my reliability as a viable researcher is not to be taken for granted. I also have deficiencies and biases (Robson and McCartan 2016, p. 462). This is important to be aware of as a researcher, to try to limit the scope of these deficiencies somewhat.

## 5.2 Qualitative interviews and their objective

The chosen method for data collection was qualitative interviews. As described in Section 5.1 a qualitative approach is taken, and interviews are chosen as a viable method to gather relevant data to the RQs. Kvale and Brinkmann (2015, p. 46) state that semi-structured phenomenological interviews are used when trying to understand the respondents perspectives of themes and topics from their daily life.

The interviews seek to explore, understand and find meaning behind PCR in secondary education. This implicates mapping out if they use the method, if they relate the method to practices they are already doing, why they have used it or not. Further, the interviews seek to gain an understanding of how the teachers' attitude towards the method is, and if it would be a viable method to further explore and implement in secondary education. Knowledge connected to this are benefits the method would bring if implemented, and challenges it would pose. Motivations for why the method should at all be used is also of paramount importance.

PCR is, as explained in Chapter 2, subject to expansion. This indicates that expansions are a large part of what would be discussed in interviews, as anything that is not encompassed by the three steps are viewed as an expansion. The objective of the interview therefore also seeks to understand which expansions would be beneficial to implement, and why. How they are to be implemented is also a

concern of the interviews. Which challenges might arise is up to the respondents, but it might be difficult for the respondents to have an understanding of which challenges might arise (MacLeod et al. 2018, p. 34). For that reason the challenges presented by Indriasari et al. (2020, p. 15) are used as scaffolding to help the respondents in their thought-process.

## 5.3 Data collection

The data collection done in the research is presented in this section. Data collection was done with personal interviews, as opposed to focus groups or other similar methods. The respondents are all teachers of IT in secondary education that I recruited through emailing them. The interviews were all conducted in person, within a span of two weeks. To conduct the interviews a interview guide was made, with questions that aimed at reflecting the RQs and the objective of the research explained in Section 5.1 and Section 5.2. The interviews were recorded and transcribed afterhand.

### 5.3.1 Personal interview

The interview conducted was a personal interview with a semi structured approach. This means that the interview guide is there to be a guidance, it is neither an open conversation or a completely rigid one (Kvale and Brinkmann 2015, p. 46). This type of interview is well suited to gather descriptions and interpretations of the topics at hand (Kvale and Brinkmann 2015, p. 46). This type of interview lets the conversation unfold naturally, and the respondents get to explain their view of it in their own way. The researcher can then stray away from the rigidity of the interview guide and seek to gain a deeper understanding of what the respondents present.

### 5.3.2 Respondents

The respondents of the interviews are CS teachers in secondary school. To conduct the personal interviews, the study needs respondents. The amount of respondents are subject to the constraints of the project, and the eight respondents for this interview is well within what is considered normal (Kvale and Brinkmann 2015, p. 148). Eight respondents were deemed to be an adequate number for the data to be somewhat complete. That number of respondents need to reach a point of saturation, meaning somewhere around the number chosen, increasing it with new respondents will add less and less new knowledge (Kvale and Brinkmann 2015, p. 148). Two of these eight respondents work in lower secondary school, and six in higher secondary school. They have anything between ten and 37 years of teaching experience. Teaching programming they are more spread, having anything from a few months up to over 30 years of experience. As a group they do not only teach programming subjects, but also teach many natural sciences subjects

and some social sciences subjects. Their formal programming competence is also varied, all respondents having between 7.5 study points and about a few years worth of competence.

Recruiting respondents was a whole process in itself. Potential respondents were mapped out from searching on the internet for both lower secondary and higher secondary schools in Trondheim and the area around. What is regarded as potential respondents were teachers in courses that had programming as one of its main parts in the curriculum. In practice, this meant recruiting teachers from the course Programmering (Utdanningsdirektoratet 2020) in lower secondary school and IT1 (Utdanningsdirektoratet 2006a) and IT2 (Utdanningsdirektoratet 2006b) in higher secondary school. An initial email with some limited information about the research with a request for participation in an interview was sent out, making initial contact. More emails were sent out than were responded to. Further contact was kept with respondents replying to the initial contact, giving them more information about the project and attempting to set a date, time and location for the interview to be conducted. The respondents were sent an information sheet and a written consent, see Section 9.1. This was not required to be read by the teachers before the interview, but was sent to them as an attempt to inform the respondents in a as good and simple way as possible about the project.

In advance of the interview the teachers received some information about PCR. I made an information sheet explaining PCR to be used as scaffolding during the interview, see Appendix B. This sheet was not by any means required, or intended, to be read before the interview, but to avoid confusion and the sense of secrecy the sheet was sent to the teachers with an explicit note saying this was not a required read before the interview.

### 5.3.3 Conducting the interviews

The interviews were conducted in a timeframe of just over two weeks. The first interview was conducted 28.09.20, and the last one was conducted 12.11.20. At most, two interviews were done in one day. Four of the interviews were carried out the same day as one other interview, and four of them were done as single-day events. This was done on purpose, both to have adequate time to arrive at the schools to carry out the interviews, and to allow for time to reflect upon a conducted interview, as well as preparing for the next.. As an interviewer I wanted to have a good understanding of the entire situation of each interview, in a way that I would remember parts of the interviews when listening through them afterwards, which is important for properly understanding the interview afterward when transcribing and reading through the transcriptions (Kvale and Brinkmann 2015, p. 161).

The interviews were recorded with a recording device from NTNU. The most common way of registering interviews is by recordings, because it allows for the researcher to focus on the contents of the interview and its dynamic (Kvale and Brinkmann 2015, p. 205). The interviews later transcribed and analysed. There

were no notes taken during the interviews, as I as a researcher found that the gains I would make from having notes would not outweigh the distraction the taking of notes would have caused (Kvale and Brinkmann 2015, p. 206). See Section 9.1.4 for information about the storage and processing of the recordings.

The interview was essentially split into three parts, with a briefing before and a short debrief after. The briefing was an informative part from me. The information given was before the recording started, and was intended to clarify the situation for the respondent, as well as inform about the process going forward. This included telling the respondents about the purpose of the interview, the recording device I was using, how the data would be used afterhand, what kinds of rights they had and my duties as a researcher. I also explained how the interview would be conducted and asked if they had any questions before I started recording. These steps are crucial to create a good atmosphere (Kvale and Brinkmann 2015, p. 160).

At the next part the interview itself started, starting the recording and asking questions about the respondent. The questions about the respondent were about defining who the respondents are in this context.

The second part of the interview, i.e. the main part of the interview, was where all of the questions about PCR were asked. In this part the focus was about PCR, and I asked questions outside of the interview guide to verify that I understood the respondents' answers to my questions. Kvale and Brinkmann (2015, p. 165) focus on the importance of having the later stages of the research in mind when conducting the interview, as well as the questions' thematic and dynamical aspects. This was a focus throughout the interviews, trying to ask more questions when I sensed there was an interesting theme being mentioned, asking the respondents to elaborate or trying to verify their answers.

The last part of the interview was about wrapping it up. In this part, it was important to check if the respondent had some last thoughts, any topics they wanted to talk about in the interview, or if they just have something else they want to say Kvale and Brinkmann (2015, p. 161). This was done to let them walk away from the interview with the feeling that they have said all they wanted to say.

After the recorder was turned off, there was a short session to end the interview in a natural way. This was done to ensure that the interviewer felt that their interests were taken care of, and to further explain the process of what happens next with the recording if needed.

### 5.3.4 Interview guide and questions

The instrument the data collection is done through is an interview guide. A tentative interview guide was also a requirement of NSD, see Section 9.1.1. An interview guide can be built up in a variety of ways. It can contain themes to be covered, an order for the questions to be asked in, topics to be talked about and suggestions for questions to be asked (Kvale and Brinkmann 2015, p. 162). The interview guide was built upon questions that sought to gain insight of PCR in secondary education and with respect to the respondent. The questions the interview guide

contains have to take into account both the thematic aspect of the interview, i.e. the creation of knowledge, and the aspect of the relation between the researcher and the respondent (Kvale and Brinkmann 2015, pp. 162–163).

The questions were designed into three parts for the interview. The parts were as described in Section 5.3.4. This means that the main part of the questions was mostly where the knowledge was intended to be created. In the first part of the interview I asked the respondents about which subjects they teach, how long they have been teachers, their formal competence in programming or CS and how long they have been teaching programming. The questions were designed to allow for both a strict run through the questions, and for the possibility of straying away from the designed interview guide. There is a fine line one has to thread when making the decision of how rigid one should follow the interview guide, with a less rigid approach allowing for more spontaneity and a more rigid approach making the structuring of the analysis easier (Kvale and Brinkmann 2015, p. 163). This was done through having short and concise questions, allowing for the guide to be followed, and making it somewhat natural to ask follow-up questions.

The main part of the interview contained most of the questions, including some scaffolding. In the main section I start with asking what their own experiences with PCR is, and I get them to elaborate somewhat upon this. In this way I get their initial thoughts about PCR, how they understand it by themselves and if they might relate it to something. After this I conduct a scaffolding process, explaining the information sheet about PCR, setting a common ground, as it can be difficult to further discuss PCR without having a common understanding of the concept. After the scaffolding, comes a general section about PCR. Here the respondents are asked about what benefits and challenges they see with PCR, with me as a researcher actively listening trying to make them elaborate upon their answers. They are also asked if they see any differences between implementing PCR in secondary education as opposed to higher education. Before the next questions I gave some more information to the respondents, informing them very briefly about the results of the SLR conducted, and telling them that we will be using some challenges reported from PCR in higher education as a base for the next part of the interview. These challenges were the ones presented in Section 2.2 by Indriasari et al. (2020). They were told to reflect upon the challenges reported, state if they thought it would also be a challenge in secondary education and if they had any thoughts about how to meet the challenges in secondary education. In the main part they were also asked if they thought the method feasible in secondary education, and lastly asked what could be done or provided to aid them in implementing PCR in their own educational practice.

The last part of the interview guide was to round it out. The last question was an open question where the respondents could elaborate on what they wanted, say some choice words or just wrap up the interview in their own words. See Appendix A for the entire interview guide that was used.

### 5.3.5 Transcription

The recordings of the interviews were transcribed. Transcribing is the process of translating the recordings, i.e. speech, into text (Kvale and Brinkmann 2015, p. 205). In doing this process, bits and parts get lost. The feelings and thoughts one would connect to the recordings might not be presentable in the written text, and the text is not written as a formal text. Voice tones and high and low pitches are also lost. This creates a hybrid between the two worlds, but it is nevertheless a crucial step to do to be able to analyse the data (Kvale and Brinkmann 2015, p. 206).

The data has been transformed several times. The data started as live conversations, where hand gestures, body language and other factors were at play. It was then recorded, losing the visual aspect of the conversations. Afterwards, it was again transformed into transcriptions, with the limitations that come with that form as described in the prior paragraph. Afterwards, it was analysed, with multiple iterations, and written out as Chapter 7 in this report. The analysis was done in Norwegian, and the quotes were initially put into the results in Norwegian, but was later translated to English, to make for a better reading experience as the report is in English. The language translation might lose or confuse some of the meaning that existed in Norwegian, however, to mitigate the impact of translation this was an active concern thought of during the translation.

There was a strategy for the transcription of the data. Transcription is a procedure that will create different results if no common strategy exists. There is no universal strategy to follow, but it is important that the strategy is described (Kvale and Brinkmann 2015, p. 207). The strategy as a whole includes the choice of what kinds of sounds get transcribed, what abbreviations are used, if pauses are to be included and so forth (Kvale and Brinkmann 2015, p. 208). The choices made for this report included the use of ".." when a short brake is taken and "..." when a longer brake is taken. "###" was used when the recording was inaudible, with a timestamp after. To indicate that either the respondent or the interviewer got interrupted by the other in the middle of a sentence, "../" and "/.." were used to indicate where they got interrupted and where they continued. Filler-words by the interviewer which clearly were used to create a flow in the conversation, and to keep a good dynamic, showing interest, were not included in the transcription. To anonymize the transcriptions "@@" was used. To mark conversations irrelevant to the interview "[*timestamp* irrelevant conversation]" was used. To only include relevant data to the analysis, the parts where the interviewer informed on the PCR information sheet and told about the SLR on PCR in secondary education were marked as "[*timestamp* interviewer talks about PCR/SLR on PCR]". Further, other descriptions or relevant notes to clarify something was done by writing "[*notes*]". Other relevant descriptions would include the use of e.g. irony, sarcasm, laughter. These choices were made specific to this situation, as it is more so a content analysis than e.g. a textual one I deemed this sufficient.

Transcription can be helpful to do by oneself, as it is a good first step of the ana-

lysis by familiarizing with the data material (Kvale and Brinkmann 2015, p. 207). It is also normal to receive help in the transcription (Kvale and Brinkmann 2015, p. 207). I transcribed three of the interviews by myself, and got very kind help in transcribing five of them by the Department of Computer Science at NTNU (NTNU n.d.[c]). I sent out a written instruction of how I wished that they would transcribe the interviews, to receive as equal transcription choices in all of them as possible.

The conducting of the transcribing was done with multiple iterations. The first iteration was to write word for word what was said, taking many brakes and going back ten seconds at a time. The help I got was in this iteration. The interviews were run at half speed, to less than half speed depending on the difficulties I had with hearing what was said. The second iteration was to listen through the interviews while reading the transcriptions, correcting errors, checking if the correct strategy as described above was used. The third iteration was again to listen through the interviews, this time making sure I had gotten them all right.

# Chapter 6

# Analysis

This Chapter concerns how the data was analyzed. The analysis of the transcriptions begins early, and should not start after the data is fully collected and transcribed.

Thematic coding was the method of choice, using codes and themes collected in a codebook, and it was done in NVIVO 12 Pro while following a set of guidelines. NVIVO is used as a tool to organize, store and analyze data (Alfasoft n.d.). The tool is especially useful for its compatibility with codes. The analysis was done following Robson & McCartan's guidelines conducting a thematic coding analysis (2016, pp. 468–480), further explained in Section 6.2. The guidelines divide the analysis into five phases and are intended to create a firm structure to aid in the effort to minimize human deficiency in the process.

## 6.1 Thematic coding

Thematic coding was used for the analysis. *Codes* are the smallest segment of raw data in this process (Robson and McCartan 2016, p. 471), containing small snippets of the transcriptions. This can be anything from a couple of words to a couple of short paragraphs. Codes are not mutually exclusive, meaning the same segment of text can be coded to different codes overlapping fully or partially. Codes are then grouped together to the overarching *themes*, which is a term that captures something to help me as a researcher answer my research questions (Robson and McCartan 2016, p. 468). This is something that can be defined before coding the transcriptions, or after. The development of these themes is done through the process of seeing what kinds of codes fit naturally together. The codes are imperative to understanding the data, and should lay the foundation for the analysis and interpretation (Robson and McCartan 2016, p. 468). The collection of all the codes and themes is referred to as a codebook.

## 6.2 Conducting the analysis

The analysis has been an ongoing process. The codes were continuously altered and there a memo was used to keep track of the progress and thoughts of the process, noting down ideas for codes and themes. The keeping of memos is useful for capturing ideas, views and intuitions through the analysis (Robson and McCartan 2016, p. 467). It started when the first codebook was created, before any of the data was collected. Then the data was collected, see Chapter 5, where the codebook was prone to further alteration. The codes in the codebook, and the themes, were finally edited as needed throughout the coding of the material and the identification of themes. See Appendix E for an example of how the transcriptions were coded.

The guidelines followed consisted of five steps. Robson and McCartan (2016, p. 469) name them accordingly:

- Phase 1: Famliliarizing myself with the data
- Phase 2: Generating initial codes
- Phase 3: Identifying themes
- Phase 4: Constructing thematic networks and making comparisons
- Phase 5: Integration and interpretation

These five phases are not done in a linear order, as one might think from first seeing them. However, they do generally start at Phase 1 and continue through to Phase 5, with the possibility of jumping back and forth. They are intended to create a firm structure around the analysis and to help overcome possible human deficiencies (Robson and McCartan 2016, p. 462). How these five steps were carried out in this project is described in the coming sub sections.

### 6.2.1 Familiarizing myself with the data

Getting to know the data is a crucial part of this process, as this process could be argued is as much a test of the researcher as it is of the dataset in use (Robson and McCartan 2016, p. 462). The familiarization with the data started early, as it is important to give this process time. I got familiar with the data through having the analysis in the back of my head during the interviews, during the transcription and during the coding itself. One could even argue I started familiarizing myself as soon as I started to create codes and themes, as this serves as a frame for the data. The familiarizing with the data is important to do as a researcher as one can remember aspects of the interview that might not appear to an outside viewer of the data, e.g. if there was pointing to a certain area of the provided material for the interview. This is naturally an iterative process, where I got more and more familiar with the data as I saw it represented in different ways, making for a more wholesome view of the data after some time. I both listened through all of the data, read through it, and re-read it. See Section 5.3.6 for more on this. During this phase, and during the timespan the interviews were being conducted, I kept a memo writing down ideas and thoughts about potential themes and codes. This is

an important part of the familiarization, as it can be easy to forget things (Robson and McCartan 2016, pp. 470–471). The memo was continually updated and used to keep a good structure to the potential themes and codes. As described in Section 5.3.6, I got help with the transcriptions, but did two of the iterations on all of the transcriptions myself. This part helped me double-check my initial codes, and add those I immediately saw the need for. The entire transcription process is very beneficial for familiarization with the data (Robson and McCartan 2016, p. 471).

### 6.2.2 Generating initial codes

The codebook was initially drafted based on the existing literature, see Chapter 2, before the interviews were conducted. This was done in an effort to carry out a deductive process, as opposed to an inductive one. However, the themes and codes have been altered both before, during, and after the data was collected making the process a hybrid solution. I coded for as many themes as I could come up with, trying to be as open as possible in this part of the process. However, the codes created before conducting the interviews showed themselves to be quite steady, making the changes during and after the collection as minor adjustments and identification of a few new ones. The codebook was mostly "theory-driven", with parts being adjusted to the data making it also "data-driven". The themes and codes, and analysis as a whole, do not aim at placing the findings within the current literature, as there is a strictly limited amount of this, see Chapter 4, but rather aims to explore the option of implementing PCR into secondary education. The coding process is done to organize the data into meaningful categories. The deductive part of the creation of the codebook puts the analysis at risk for being biased towards some aspects of the data, which is listed as a possible human deficiency by Robson and McCartan (2016, p. 462). This, however, can also be seen as an advantage, as it might better help me as a researcher identify features of the data that I otherwise would have missed. During this step, I naturally went back to the familiarization with the data, checking if the codes and themes were sufficient. The codes used were also being evaluated to fit in with the themes, see the next section for more about the themes. I continued to read literature throughout the analysis, to check if my codes were good. I worked systematically through the entire dataset and generated new codes where I saw the need, and used this process to get a good idea of the dataset as a whole. The coding of the dataset aimed at coding as much of the dataset as possible, however it excluded data that did not help in any way towards providing a better understanding of PCR in secondary education. For instance a conversation about how a teacher teaches a specific other method in mathematics, with no connection to PCR, was regarded as an irrelevant conversation. The codes created were mainly focused on strategies, practices or tactics, meanings, conditions or constraints, consequences, and settings (Robson and McCartan 2016, pp. 472–473). Some of the codes were used as broader codes, for instances that did not fit into any other codes, and would create for far too many code to create a new one for each of them. This

included codes like “other skills”, “other benefits” and “other challenges”.

Interestingly, there were some contradictions in my dataset. More specifically, this relates to the aspect of time. The reported findings are to be found in Sections 7.3.5 and 7.4.8. It is perfectly fine to have contradictions like this (Robson and McCartan 2016, p. 474).

### 6.2.3 Identifying themes

This part of the process is about selecting which codes to put into which potential themes. The themes remained largely the same throughout the process. Theory -related material was the basis for the themes. Other than this, focus has been on repetitions and similarities and differences as techniques for identifying themes (Robson and McCartan 2016, pp. 474–475). As described in the previous section, the codes were largely “theory-driven”, and accordingly the themes were largely based on theory-related material. Repetitions in the term of themes means that the contents of the theme keeps on occurring. Similarities and differences were also a focus for the themes, checking if statements made by the respondents had similarities or differences.

The themes are the main part of the thematic analysis. Themes are what are important, the codes are only a tool to get there (Robson and McCartan 2016, p. 474). After all of the data was coded and all codes were put into themes, it was important to solidify and quality assure each theme (Robson and McCartan 2016, p. 475). The themes were checked for some emerging pattern, altering themes that did not provide any logical patterns. This is an iterative process, meaning that it is important to go over the data set after this process is done, to check if the themes match the perceived contents of the data, and then go back to the themes afterwards to change them if needed. However, it is important to realize that this process should not be expected to be perfected (Robson and McCartan 2016, p. 475). At the end of this process the themes of the analysis really started to solidify, their inter-relationships started to show and the overall takeaways became somewhat clear.

### 6.2.4 Constructing thematic networks and making comparisons

In this part of the analysis the themes are actually mapped together, creating solid inter-relationships between themes and connecting them to the overall picture. This step was already started during the coding of the transcriptions, but only as a writing down ideas and thoughts during that phase. This phase was done not separately, but was kept in mind during the entirety of the three prior steps.

### 6.2.5 Integration and interpretation

Finally, the data was derived to make some key takeaways. Robson and McCartan (2016, pp. 476–477) list a set of tactics to derive some meaning behind what story the data is telling. This goes beyond codes, themes, and schematic representations

of the data. When summarizing data such as this the line between qualitative and quantitative data might become somewhat blurry (Robson and McCartan 2016, p. 477). It is then important to interpret and describe it not as quantitative data, but as the qualitative data it is. The results of the analysis are presented in Chapter 7.

# Chapter 7

# Results

In this chapter the results from the analysis of the collected data is presented. For information regarding how the data was collected and how it was analysed, see Chapters 5 and 6 respectively. The chapter is set up to first look at the current status found in the analysis, then it covers motivations for using PCR, continuing with benefits of PCR, challenges with PCR and lastly elaborates on the implementation strategies discussed in the interviews. The results are based on the final codebook, see Appendix D. The sections in this chapter generally follow the themes, and sub sections generally follow codes covered by each theme. Some of the codes have, however, been slotted into the same result, as some of the codes are intertwined. Less descriptive codes like "Other benefits" and "Other motivations" have in the results been incorporated into other fitting places, with more descriptive formulations. The theme "Education level" was in itself almost always connected to other codes, making the theme not have its own results section but being incorporated into the other sections here.

In this chapter the terms respondent and teacher refer to the interviewees, the terms participants, students and peers are all used to describe the participants of PCR, and lastly the terms extension, addition or strategy are non-critical parts of PCR.

## 7.1 Status quo

The first result presented is the status quo of PCR in secondary education. The status of PCR in secondary education is divided into how teachers understand PCR intuitively, their attitude towards PCR and the usage of PCR or similar methods.

### 7.1.1 Programming teachers easily relate to PCR

Most respondents felt they could answer the initial question of “Have you used PCR while teaching?” indicating they intuitively understood the contents of PCR, but none of them knew of the method beforehand, apart from the direct communication with me.

All respondents report conducting the somewhat loosely defined *informal* PCR. Informal PCR in this context means there is no systematic approach to any of it, there is no formal submission, there is no guarantee of getting one's work reviewed and there is no need to participate as a student. The process is encouraged by the teachers somewhat randomly, perhaps when the students encounter difficulties and the teacher is not available, and the teacher has often beforehand, or when it happens, encouraged them to seek help from one their peers:

> Yes, I do encourage them to sit together (...) working in that way so that they can help each other. (...) That is an encouragement we have, right. Because it is a type of subject where you, if you end up with a problem you can get stuck with it for a long time before you find out what is wrong. (...) They can't get further. Because there is some kind of mistake. In the code, right. And then that can get stuck for a really long time. At that point you are totally dependent on that they actually are each others' teachers, you could say. In a much larger degree in programming than in other subjects. I think. (R3)

### 7.1.2 Positive attitude towards PCR

The attitude towards PCR was largely positive among the respondents. One very positive response, jump-the-gun style, was "*This is a great idea, now I know what I will do (...) Then we will do this. And we will do it in IT1 as well, it is completely brilliant. Nice.*" (R1). On the other end of the spectrum there was more scepticism to systematizing an already happening phenomenon in the classroom:

> No, there is not that much required before one could go and do this more systematically. Because it only means that you say that you two are to work together and evaluate. It is a good idea, but it happens nevertheless. So I have not needed to think about it. (R2)

All respondents lay somewhere in between these two, indicating that the general attitude towards PCR in secondary education was quite positive. Further cementing the positive attitude towards PCR is the result that all respondents viewed PCR as a feasible method, not seeing any major issues that would hinder them from using it.

### 7.1.3 No reported usage of PCR

While the concept was easy to grasp for most of the respondents, none of them had any experience implementing PCR as the process that was laid out to them. While all of the respondents said that they had not used the method in "this way", many of the respondents say they have used methods that they themselves would categorize as similar methods. One respondent linked PCR to formative assessment and the new national curriculum (LK2020) in Norwegian schools, and how it would be a good way to implement peer assessment, as peer assessment was

to gain a more significant place in the school system according to this respondent. Multiple of the respondents related the method, both before and after the scaffolding process in the interview, to methods they had themselves used. This was anything from larger scale projects down to asking the person sitting next to them to help. One of the different methods they found similarities to PCR was a code review the entire class was to partake in. Respondent R7 elaborates "*It has happened that if a student asks for help, and then I sit and look at it and find nothing, then I say «can you show it on the canvas in a way that the whole class can help you?» Yes, its okay, everyone says then.*" (R7)

None of them knew that it was possible to systematize this practice that they all already were doing, possibly explained quite well by R4:

> But programming has, at least for my part, been a pretty new and fresh subject that has taken a bit of time to figure out how to teach in, for me that has never done it before. And then, even though I have in a way done a bunch of things that I do in outher subjects, it seems like it can take a bit of time before I understand that I can do the same in programming too. The same way when I saw this system, why have I not done this. I have done this in other subjects, I could have just copied it and almost done it in the same way in programming too. So I don't have any other good answer than that, I just need hints and time tips on that one can do this. (R4)

The line between what is seen as a similar method and what is seen as informal PCR is somewhat blurry and up to interpretation. In this report there is an imaginary line just to paint the picture of a method as a planned out strategy of using something that is somewhat similar to PCR, and informal PCR as PCR that is not planned out.

## 7.2 Motivations

In this section the underlying motivations if one was to use PCR in secondary education is explained, as presented by the respondents. What is to be evaluated as motivation behind using PCR, and what is to be regarded as a benefit of using PCR, is somewhat hard to define. This report tries to identify motivations as general reasons for why it would be considered used by teachers.

### 7.2.1 Learning

Most respondents heavily prioritized that learning is essential. The school is about learning, and the methods have to have a solid grounding in learning. Respondent R8 elaborates how PCR is connected to how things are done in the industry, and points out that this is a significant advantage of the method: "*It will look more like how things are done in businesses and in the industry. The way I teach I want to do*

*things as closely as possible as it is done in the industry when it comes to programming and development.*" (R8).

Respondent R7 explains there is a possibility that some of his students might continue with coding, and that the higher up in the education system one is when coding, the more likely it is that the person will continue with it. The respondent further explains that students studying code will be writing more code, and will need to learn to e.g. read others' codes:

> Well there are students here that might never code again. But those who are students in these courses, they will be doing it. And then it is even more central to be able to understand other peoples' code. (R7)

### 7.2.2 Motivating as a method

One of the motivations that was discussed by a few of the respondents was the increased motivation for students while using PCR. The students might feel like they are contributing, in a way that students might not always feel after a class:

> One gets the feeling that one has value, because you give back something to another student, and it is not always that a student sits inn a class and feels that what one does is valuable. So there are many aspects with that [PCR] that touches with a person or a student or, that I think is positive. (R1)

The possibility that certain aspects of PCR might be motivating, such as setting a score on a peer. R4 says "*So it can be a bit motivating for the students sometimes to put a kind of score or grade or something like that on tasks they evaluate.*" (R4) Having peers review your submission might be an added motivation for actually using the feedback one gets:

> Both the teacher will use and the student will use it to get further in the learning process. And when it is this evaluating of each other where you get an evaluation from a peer then the motivation to look at the evaluation and to do something with it might become bigger. (R6)

One respondent directly reported the method as motivating in itself for the teacher, seeing as the method seemed intriguing for the teacher to adapt. Many other teachers, as reported in Section 7.1.2, have a generally quite positive attitude towards PCR - indicating that they too saw the method as somewhat interesting to try.

## 7.3 Benefits

The benefits as seen by the respondents are reported in this section, and there was largely a focus on learning connected to the different sub-sections presented. The

benefits include improved code quality and readability, increased student engagement partly due to social implications, learning from reviewing, learning from receiving reviews, viewing PCR as time effective both for teachers and students. Of these benefits perhaps the most important one, or at least most frequently discussed, was the learning from reviewing.

### 7.3.1 Code quality

Improved code quality is not a major takeaway from the results, as the teachers did not focus particularly much on this aspect of PCR. It was briefly mentioned by three respondents, but only as a side note. They largely linked a possible improved code quality to the revision strategy, see Section 7.5.5. However, one respondent discussed how PCR could lead to increasing the readability of the code, because the students know that their code will be read by their peers:

> It is pretty reasonable to have a good structure of the code (...) it should also be readable to others and it will be necessary to add comments from others. (...) That you program with the though of others actually will use it and make changes and modify it. In a project it will quickly be that you send it between, so then it must be a bit readable. (...) The one who creates the code must think: «here I am also communicating something to the other one. The code is not something I only have for myself», and think that here the code will actually be used by someone else. Like it often might be if you will work with programming later and then you have to structure it properly, but at the same time your coding gets better. You get better yourself because you think more systematically yourself and over time it might make you better. (R2)

### 7.3.2 Student engagement and social implications

PCR has positive implications towards student engagement, which may be amplified by the social implications of the collaborative nature of the process. One part of these implications is a 'forced' engagement in what is happening in the classroom, discussed by two of the respondents:

> Because here you more or less get forced into looking at how other have tried to solve a task. But in another setting, you might picture that students present to each other how they have done it, then they might not be as dedicated to pay attention, or get an understanding of it. But here you have to get an understanding of it. (R3)

The process also facilitates more student dedication towards reading someone's suggestion of how to solve a task carefully. At least when compared to reading e.g. a teachers solution proposal:

> Because it becomes different than if they create their own proposal and maybe see the teachers' proposal. Then they might not care that much about the teachers' proposal, if theirs work. Of course if theirs does not work, it is something else. But if they have gotten it to work they oftentimes will be happy with it I think. (...) because they have to go very detailed into the code they have in front of themselves when they are to evaluate it. If they take the task seriously at least. It is probable that they do it [take it seriously] when they are to give feedback to a peer.

There is also a social perspective to this, emphasized by four of the respondents. A part of the social implications are an increased student participation when reviewing, because they feel a commitment towards their peers. Moreover, the students might also feel like they need to properly look at the reviews they receive. Another similar perspective more focused on not losing any social status was presented by R6:

> But the fact that they are writing feedback, having the knowledge that it is to be read by a peer, I would think that student, well most students, think that I need to try to do this somewhat properly. So that I do not get a social backlash. Or that the others see that I am bad at this or something similar. (R6)

### 7.3.3 Reviewing benefits

The benefits from actually reviewing someone else's code was a very important topic in the interviews, and there were multiple perspectives to it. The most prominent one was the learning connected to the reviewing, which all eight of the participants emphasized. The most important aspect was perhaps seeing different solutions to a problem, R3 elaborated upon this:

> The student might get a broader understanding for programming in this case. Because they see problem-solving from different angels. It is not just the fact that they are trying to solve a task, but also that they are looking at how others on their level try to solve a task. It will expand the horizon of the student in a way, and in the long term if this works, next time the student might have a broader repertoire, on how to solve their task themselves. Because here you more or less get forced to look at how others have tried to solve a task. (R3)

This was a sentiment shared by all eight respondents, indicating that this aspect might be a major reason for use if teachers actually implement this into their practice. Reviewing is a skill in itself, which will be improved if the students repeatedly are exposed to a process like PCR. Moreover, students will have a better understanding of how the grading of the teachers work:

> At least when I have had students to evaluate, take for instance in maths they can get their test back and they have solution proposal, then you get better when you set a grade on it yourself. And it is with hope that they will sit down and put themselves into our [as teachers] situation. That is what I look at as the value for those who give feedback and that they put themselves into a kind of teachers' role. Because them they understand more themselves. (R2)

When reviewing, the students have to read others' code. This is also a skill in itself, which is put in use in PCR. R4 says "*And then you have to try to understand how it [the code] works, perhaps by looking at some thoughts of what strengths this code and way of solving it, versus what I did myself, or what are the weaknesses with it*" (R4).

### 7.3.4 Receiving reviews

Receiving reviews in themselves does not seem to be a major factor of what the respondents focus on in PCR, but it is nevertheless discussed by five of the respondents. Every time it was mentioned, it was in connection with something else, merely as a side note. The benefits of receiving reviews seem to be focused on learning, but this benefit is mostly mentioned briefly. R4 describes the importance of also getting students to understand that these reviews are meant to help them:

> That you get feedback on your own work, then you get tips to do it better next time or do it better when you possibly will revise the code or deliver it to the teacher. So I think that you should make them aware that it is meant as an advantage for them, not just as something the teachers make them do to do more work, more meant as help to get further. (...) That it will help them in becoming better, and do better in the subject and understand things better. (R4)

### 7.3.5 A time effective method

Implementation of PCR does not need to require a lot of time by the teacher, and it can be time effective in class. This section must be read with respect to Section 7.4.8, to get the full picture. The burden of assessing students' submitted work might be lessened by implementing this method, says R2. R6 elaborates further "*and it can be time saving for the teacher as well. Then the teacher does not have to write feedback. (...) you save time on not having to write the feedback yourself, because the peers are doing it instead*" (R6).

R3 and R4 explains how time can be saved as a teacher, when going over the submissions because the students already have reviewed the submissions. This again frees up time for the teacher, which can be utilized in some way to create better learning says R3. Several of the respondents say there is a connection

between the opportunity for PCR to be time effective and the availability of a technical solution helping the implementation, see Section 7.6.2 for more.

PCR is time effective, but it largely depends on what subject it is used in. This result is very context dependent, and this study is set in a Norwegian context. The respondents largely link its potential time efficiency, meaning its gains outweighing the costs of time, to single courses and not always programming as a whole. They describe it as a very feasible and possibly time effective method in e.g. IT1, while when teaching programming in mathematics it might not be a good option to go for. R2 says that "*I think there should be plenty of time in the pure IT subjects. I think there should be. So I do not think it should require much more time. But we do not know that before we have tried.*" (R2)

## 7.4 Challenges

The challenges as described by the respondents are reported in this section. The challenges are more comprehensively reported upon, as the interview was slightly more focused on discussing challenges with PCR.

### 7.4.1 Implementation challenges

Challenges regarding the implementation of PCR are presented here. Some implementation challenges relate to different strategies, those will be discussed in Section 7.5. The implementation of PCR differs in different CS courses, e.g. Programming in lower secondary school, IT1 and IT2 in higher secondary school. The question of the flow of the process is also one that was highlighted by R2, though not a major challenge posed by any of the other respondents:

> This can quickly become into that when they have solved the task, right, then they sit and wait for the feedback from the other. If it is to be done in the classes or not. Are we to start in the meantime here, while we wait, or should we start with something new? Right. In a project. The students might not get further before they have gotten feedback from others. (...) The flow of the process might not be very smooth compared to what would be the normal, that you make a task (R2)

One challenge that might be very interesting and that might have a very practical implication, although not a major result, is if one student has some underlying mental condition which can make that student especially difficult to engage in a comprehensive process like this. R3 says "*and especially with some students it might be hard to motivate them. (...) students a bit in the Asperger's spectrum and a bit like that, they would refuse to do this, right.*" (R3) One challenge that might become prevalent for some teachers, is the reality of differences in different programming languages. R4 elaborates upon differences when it comes to Scratch and Python:

> We use Scratch and if I was to distribute that code to someone else I would not be able to get it anonymized in any other way than by taking a screenshot of the code. Because when I take a copy, it gets remixed, and then you can always trace it back to the one who has created the code. (...) But for example if we use programming in python with repl.it, it can be anonymized. You can copy the text. And then it is excellent, or it makes it a lot easier at least. (R4)

Not a major result here, but the challenge of students plagiarising the others might become relevant, as PCR is built upon sharing your code with your peers.

### 7.4.2 Level differences between students

Level difference among students seems to be a very central challenge, as many respondents viewed this challenge as something that has to be addressed when implementing PCR. Six of the eight respondents discussed level differences among students. One of the problems of level differences among students relate to what lower achievers of do when receiving submissions they do not understand:

> Yes, when they get, I mean when they get another test, another text, or code, that you have not made yourself and are to evaluate. And it is written in a pretty high level that this student does not understand. What kind of evaluation is he to write then? (R7)

The implication of the above mentioned challenge is described is very connected to the contents of the next section, as R3 says "*But he who has a high level might not get any feedback that is appropriate.*" (R3)

In total, too large level differences might lead to neither student, either the one with high or the one with low competence, profiting from the potential of the process. The results indicate that it might not be entirely clear, or at least not intuitively clear, for all teachers what is a good way to face the difference in level problem, which in itself is a challenge.

### 7.4.3 Lack of ability to code, write reviews and low quality of reviews

Lack of ability to code, lack of ability to write reviews, and low quality of reviews are challenges that are very often related to each other. As mentioned in the last section, there is a problem when one of the respondents lack ability to code. R6 explains "*if you do not master loops yourself it might be hard to go through a code and see if it is used correctly.*" (R6)

Not being able to write reviews is itself a problem, but not as big of a problem as the lack of ability to code. The respondents seem to view the lack of ability to code as a more imminent challenge to PCR. Students might not innately have the ability to write good reviews, R3 explained "*It is a challenge. Because they might not be used to doing it. And you do not quite know what they are used to doing in primary school when they get here.*" R3

The low quality of reviews can be connected to the lacking ability to write reviews, but the respondents most significantly connected this as a challenge to lacking ability to code and level differences. R8 explained how the lack of ability to write reviews might be more significant the younger the students are:"*And the higher up in the levels you get, the more competence you have, the easier it is to give feedback if you have more experiences. And having seen more that type of tasks*" (R8).

A significant challenge is the possibility of a disconnect between the review and the actual curriculum, R8 says "*The students have not carefully read the curriculum and those things. It is a challenge that the feedback of the students do not necessarily agree with the curriculum, for example.*" (R8)

### 7.4.4 Maturity

Maturity was a challenge and is not a major takeaway from the results, as only three of the respondents talked about this. Age reported as a significant indicator on this, indicating that the process is easier to implement the older the students are:

> Let us say that the youngest in lower secondary education, 8. grade, they are just 13 years maybe. And then it somewhat concerns maturity and in what capacity they are able to give relevant feedback. (...) It would have to come at a later time. (R3)

R4 explains that the maturity might be more dependent on the individual group of students at lower secondary school than it necessarily is in higher secondary education:

> In lower secondary school I find that it is very dependent on the group of students you have, and the maturity of the group of students. With regards to challenges. If they can handle this and which students participate and how the teacher deals with classroom management etc. In upper secondary school you might at least think they are mature enough to be able to do it just fine. That they understand the value of it. (R4)

### 7.4.5 Social challenges

Social challenges seem to cement themselves in many different ways if one implements PCR in secondary education. PCR seems to be very connected to the social aspect of the classroom, as all eight respondents discussed some form of social challenge. R1 briefly says "*[it is a challenge that] the one who gives it away is not ashamed of if what they submit is too bad, because you know it is evaluated by another student.*" (R1). This challenge becomes even more significant if students choose not to partake in the process, because they are not confident in their own abilities. R2 says "*Not everyone has as great confidence on these things. In a worst*

*case a very weak student that might not deliver at all, because he is a bit afraid of getting feedback, simply.*" (R2)

Some students might find it unpleasant to know that one or more of their peers are going to see their code. R6 says "*It might happen that some students find it unpleasant that peers are going to see what they have done. If you do not know yourself, or get to participate in deciding who gets to see.*" (R6)

One respondent mentioned the challenge of students wanting to find out who had reviewed their code. This challenge is further problematized when considering the possibility of social exclusion. R5 says "*There will be challenges if the students do not take it seriously. Of course there might be challenges then. Can it be used in an unfortunate way, if the students use it to give someone a kick [negatively meant].*" (R5)

Another aspect is if the reviews are not honest, if the students are afraid to be honest:

> one challenge can be that someone might not dare to be honest, possibly. To their peers. Dare to say what should become better. That it becomes more 'kindism', that they think «that was good, that was good, that was good. Done, now I have provided feedback». That it might, it might be a challenge at least in higher levels. (R4)

R8 discusses the problem of students giving good reviews, or good scores, to their friends not because of the quality of the submission, but for the social aspect of it. Taken together, both the aspect of social exclusion and the aspect of being too kind, PCR can work as an amplifier for existing social situations. These social challenges might be more prevalent the younger the participants are:

> But I think the challenges will exist on all levels. You can perhaps think that the social aspect is more challenging when you are younger, but that is not necessarily true either. I think that in higher education the everyday is not as colored by a social game where there are challenges with those around you. (R8)

### 7.4.6 Lack of student engagement

Lack of student engagement in the method is an existing challenge, and will always have to be considered. R1 states that the challenge of low interest will almost always be a case:

> Yes, but exactly that low interest for the subject will reflect in everything you do. You just got to take it into account, that in a class with 25 students there is surely someone who will not be motivated to do it. (R1)

A challenge is if the students do not want to read their received reviews, and if the students mostly do the steps of PCR just to finish. R2 says "*That might bewhere they put in the most energy, to be done. So that it dependent on that good feedback is given, but it does not help with good feedback if they are not read.*" (R2)

Students might see it as extra work and just want to keep doing what they came to do - programming:

> It is possible that they see it as hassle because they want to sit and program themselves, and solve their own problems. It is kind of why they are there. And that you then are to force them to look at others' programming like this, randomly. Maybe they get a task from a student they don't even have any relationship towards, that might be hard to motivate for. (R3)

The challenge of students not being interested in the subject should, however, not be a major issue multiple of the respondents argue, as CS subject are taken by their own choice. R6 explains " *There [in CS courses] everyone has a certain level of motivation, as they have chosen the subject because they are interested in programming. And that is true for the majority.*" (R6)

A very real challenge might be that the teacher sees the potential of PCR, while the students do not see the benefits the process brings, R3 argues. Lack of engagement might also come to show with the quality of the review, even though it has been done. The review is not necessarily good just because a student has high programming competence, the student has to be engaged in the method. R6 says "*You quickly get surprised by, yes, how simple [in a negative manner] you can do it. Even though the competence is there.*" (R6)

As a consequence of the possible lack of student engagement, R5 argues the process should be done in the classroom and not as homework:

> I would be careful with giving this as homework. Then I fear it could be impactful. Then you in a way eat of... Everything with homework and spare-time is in a way at competition with each other. And very much here at lower secondary education and there is a very varying degree of motivation and degree of work at home. But if you have it in a class it is not quite the same. At least in the classes I have had I think that when they first are in class, then they are willing to work. (R5)

Comparing different education levels, R7 discusses the possibility of students in higher education might be more inclined to participate in the method when compared to secondary education:

> it is possible that students generally are a bit more serious. More dedicated to working when they are not in a lecture. Because I probably have some students that work very well in class, but not outside. And for some it works, it is enough. But not for everybody. (R7)

### 7.4.7 Technical challenges

Technical challenges were discussed by four of the teachers, and about half of them seem to be related to technical problems that are not special for PCR, but

are met by teachers in their everyday life. Most of the teachers were not especially worried about different technical difficulties. One of the reported challenges regards how to implement for different programming languages:

> We use Scratch and if I was to distribute that code to someone else I would not be able to get it anonymized in any other way than by taking a screenshot of the code. Because when I take a copy, it gets remixed, and then you can always trace it back to the one who has created the code. (R4)

Another one being the changing of learning management systems, possibly directly impacting the teachers' ability to implement PCR. R5 says "*But they are changing systems periodically. So the one we have in two-three years, that we do not know today.*" (R5)

### 7.4.8 A time-consuming method

PCR can be perceived as a time-consuming process, as the teacher must prepare different things and it can be time-consuming in class. This section, of course, stands in contrast to Section 7.3.5, and should be considered accordingly. However, the results are to be reported as they were presented in the interview following the protocol of the analysis. PCR might take longer than if the teacher does what is 'normally' done, both in preparation and in the classroom. R1 says "*That it might take more time, both the planning and the execution, than it might have done if we just do like we normally do.*" (R1)

Many of the respondents highlight that if there is no system to aid them in the process, PCR might be viewed as a method that requires too much administrative work from the teacher. The same goes for other material to aid them in the process, e.g. tasks, forms, see Section 7.6.2 and 7.6.3 for further elaboration on this. Many of the teachers, however, view this as a part of their job. As something very similar to other things they have to do as teachers to be able to educate students, not viewing the administrative work as a very relevant problem.

All eight respondents say it is a challenge if this method is time-consuming in the classroom, which also was the view of many respondents. This was largely connected to the individual course it is taught in, exemplifying that it would be very doable in IT1 and might be more challenging in courses with a more comprehensive curriculum to get through. This was especially true for programming in e.g. mathematics.

## 7.5 Implementation strategies

This section provides a detailed description of how the respondents talked about implementation strategies of PCR. All of the teachers responded with implementation strategies in the part of the interview focusing on challenges reported in PCR use in higher education, but also throughout the entirety of the interview,

highlighting the importance of implementing different strategies. Many of these strategies were viewed as mandatory by the participants, but the ones coded the most throughout the transcriptions were by far "Designation strategy" and "Structured review process".

### 7.5.1 Anonymity

Anonymity is brought up as a central strategy by five of the respondents, but presented mostly as a choice. A few of the respondents viewed this as a mandatory part of the process. It was somewhat unclear if anonymization is a good strategy or not among the respondents, which in itself is a challenge. R4 says "*if there is to be given a score or evaluation, grading, then you have to think a bit about if it should be anonymized or not, and why do it, why not.*" (R4)

If there is no anonymization the level difference cannot be substantial, R1 argues " *if it is not anonymous the challenge is to find possibly a partner that is on the same level. And that both have dividend from it, that the gap is not too big.*" (R1)

Anonymization is possible to avoid if informal PCR is sought after, but not really clear why not to do it otherwise:

> I do not know what is the advantage of not anonymizing it. I do not see any other obvious reasons for it. Perhaps if you are to do it a bit informal, then you can't have it go through a system and send it directly. (R5)

The social aspect of the process becomes less prevalent if the distribution is anonymized, R5 argues:

> I think it gets a social aspect even though it is anonymous. If it is not anonymous, the social aspect might come even more into play. (...) At least if you suspect that this can be misused or you do this multiple times through the year and you have seen that someone understands (...) It is enough that they understand that they have been evaluated poorly by someone, because they feel like they do not get along with them. So it would be an option to anonymize it. Then you get rid of the possible problem. (R5)

### 7.5.2 Assessment

Assessment is the addition of having the student evaluate each other in their review, and has both disadvantages and advantages. There were differences in how the respondents viewed an assessment in PCR, with seven of the eight respondents discussing this extension. R2 raised some critical questions about including assessment:

> Well there is a big difference on if this is meant for the learning or if you are to as well have a formal evaluation on it or what. There is a

> big difference. The evaluation I would probably not trust that much. And if you do a lot of it, it might become a problem if you get the impression that they are on one level but all other evaluations point towards something else, that creates a bit of a conflict. (R2)

Students do not have the competence to evaluate each other, R2 says "*They do not have that competence in evaluating. I would not think so.*" (R2) R6 says "*What is high goal attainment and what does it take for this and this. That's why I think setting a score might be challenging. I would perhaps not do it.*" (R6), also being critical of the addition of an assessment. A slightly more open approach is taken by R8 saying that it is fine that they give each other a score, but not a grade: "*don't think that the students should give a grade. They can possibly give a score where it is possible to set one, but then you need to have a guide with how many points you give at each place. But I would not say the students should give a grade.*" (R8).

Favoritism is something students might feel like they are exposed to, and an assessment done by a peer might help mitigate that problem. R5 explains this:

> I think this might not be a big problem, but I have seen it a bit in the media that students feel that there is a 'snout factor' in the evaluations. I do not think this is central to that matter, but a conceivable consideration if you get evaluations from three students in addition. Then it might either strengthen or weaken the grade that the teacher gives. It might at least give a measure of safety that it is coherent with the evaluations they get normally. (R5)

Especially if it is done by multiple peers like elaborated upon here, more on this in the next section. Assessment by the teacher of the reviews highlighted as a strategy by two teachers:

> it would be very fun [*in a practical way*] for teachers to know who has evaluated who. The teacher can also evaluate the evaluations and show a lot of competence in the evaluation done of another student. (...) I think that could be a very useful input to the teacher if that opportunity exists. (...) It might be a bit random from time to time if you manage to get the good help and get to a solution or not. But when they are to start writing about what the others have done right or wrong, then I think it will become very clear if they have a good understanding or not I think. (R5)

### 7.5.3 Designation strategy and number of reviews

The decision of how the reviewers are distributed was central among the teachers, as it was the only strategy all eight teachers mentioned. The designation strategies proposed were mostly random as exemplified by R1 discussing "*if it is completely random or if it is controlled*" (R1). Another strategy was proposed by R2:

> how you distribute task seem like it is anonymized might be challenges with that. In higher secondary school you have learning-pairs where you sit next to each other. There might be positive and negative sides with them having permanent partners. That might be. That it is a bit more controlled who gets what. That might be connected to safety as well. (R2)

The problem of how to distribute was one that all of the teachers said was something they need to think about, to take an active stance upon. The extension where one must review multiple submission, and accordingly receive multiple reviews was elaborated upon by R4:

> And the part about giving each student multiple tasks. It is really just about time, how much time you are willing to let them use on it. If a student gets three tasks they are to evaluate and not just one, then they learn even more programming and other ways of solving it. Get insight into it, but of course if you increase it to very mange it will lose its purpose again I think. But a few more, that might be smart to give. And that every student also gets multiple reviews, which makes it so that the reviews they receive might give a bit more to the student. In that way it is not only subjective, but it might approach an objective opinion. Not just one students opinion, or review. (R4)

R5 also explained benefits of having multiple reviewers, as opposed to one:

> Exactly that you get more peoples' views of your own code. That might be nice. Different people have different points of view. (...) a conceivable consideration if you get evaluations from three students in addition. Then it might either strengthen or weaken the grade that the teacher gives. It might at least give a measure of safety that it is coherent with the evaluations they get normally (...) To make the students feel a bit safer that the evaluations are correct (R5)

Further emphasizing how the problem of receiving low quality reviews can be faced by having multiple reviewers:

> If you then have received feedback from three students, and the students have an understanding for it and knoow each other often at least as well as we as teachers know them. So that they are aware of who they can expect a good review from and that there will be variation in the feedback. Because for most students I do not think it would be a problem. If you have for example a system with multiple reviews. Possibly reviews from a teacher in addition could also be an important support in this. It is supported by the scheme that you can mark evaluations. In that way we get a middle ground solution where everyone does not need to get a teachers' evaluation. You can get it if needed. (R5)

Implementing PCR multiple times, and the implications for how to designate reviewers in that case is discussed by R4:

> If you think that you do this somewhat regularly. Then you should also rotate on who gives reviews to who. Ehm.. which makes the students, over multiple turns, might have received reviews that have good quality as well. And/or make sure of the expansion that is mentioned here [*points to PCR information sheet*] with that every student gives multiple reviews. (R4)

The problem of level difference among students, and how it connects to the designation strategy, R6 talked about:

> Yes it can be challenging with concern to which, in what way to put together who corrects each others tasks. Because it is clear that should you be allowed to correct someone who is on the same level as yourself, should you correct someone who is a lot better, worse? There might be discomfort connected to this. And it might be difficult for those who do not have that much competence in the course to correct others work. Because they are not able to see, they do not have the competence themselves, that is a challenge. (R6)

While not the main strategy proposed by the respondents, R7 suggested a manual distribution process by the teacher might be beneficial: "*Perhaps it is okay that the system is the teacher? Then you think that «this I can give to that student» because it was on a low level.*" (R7)

### 7.5.4 Learning to review

A need for training on reviewing was stated by six of the respondents. All of their opinions seem to be captured by a statement by R6: "*Evaluation and competence in evaluation is something that must be trained on. It is not something you can lay forward as a task and ask them to evaluate this and then they will intuitively be able to do it. This is something that must be trained on.*" (R6).

The impressions stated were along the lines of that reviewing is a skill in itself, and it needs to be learned before the students can partake in the entirety of the PCR process. Explained by R2 like this:

> Let us say that you have a subject over one year, then you might the first time only have a few things to look for, then this evaluation will be able to evolve gradually. In that way the first time you are only getting feedback on if it was a readable code, do you have comments, those things as well. Then it can become more specific gradually. Then you can have a kind of progression at least if you see that this is a subject that goes over multiple years, that you see progression in the reviews as well. That way they get better at that too. (R2)

One respondent viewed reviewing skills as something the students might have competence in from primary education, but also stated that it is no guarantee. R5 stated that there does not need to be structured review process with pre-made comments, but that the students nevertheless need to learn how to review:

> that you give a sort of training to the students in the start if they are not used to doing this from before, then they need a bit of training about what kinds of reviews are you supposed to give. Even though it does not need to be a complete system with comments we have talked about before, but that you give examples. (R5)

### 7.5.5 Revision

A revision step was one of the examples in the information schema presenting PCR to the teachers, and was discussed by four of the eight teachers. It was mainly discussed as a strategy for making sure the participants use the reviews they receive. Revision, i.e. the process of having to resubmit one's work, was viewed by all the respondents discussing the strategy as highly beneficial. As R8 briefly puts it:

> When I have used similar ones this last step where the students are to revise the code, I think this is an important part of it. That the feedback from the students are a step of the way on-wards. Without that extra step if becomes incomplete to me. I think as a final grade on a task the teacher should do it no matter what, but this is a very nice thing to do along the way. The students might see things that the teacher has not thought of. (R8)

R4 further elaborates that the students at least have to consider the reviews they have received, even though they might choose not to do as the reviews suggest:

> But challenges could be: how to facilitate for the students to read and use the reviews at the very end, in their further work. How to do it in a good way so that it actually has a further purpose, to the student getting the evaluation. It is challenging regardless, evaluation really, to get them to use the feedback constructively. But then the last proposal is also brilliant if you facilitate for that they are to evaluate the code because then they have to use it, I think. Or at least then they have to think about the reviews, some of them might choose not to use any of them , but they at least have to take it into consideration, in some way. (R4)

Here we also see that the respondent compares this way of doing assessment to a normal practice where they might not always get to work more on their submitted work, and the respondent is very positive to the option of revision because it ensures that the students get a chance to further learn.

### 7.5.6 Structured review process

Having a structured review process was highlighted by seven of the eight respondents as an important prerequisite for implementing PCR. As explained by R1 "*the one who receives the task and is to evaluate it must have criteria for what he is to evaluate it by. Or else it just becomes thoughts and opinions.*" (R1)

A structured review process would ensure that more students felt that they could partake in the process in its entirety:

> Even though what comes might be more complex than what he himself could have made as a solution, you can have criteria that secures that everyone is able to evaluate some things. And that gives confidence to the person who evaluates even though the level of what he gets might be more complex than what he would have been able to figure out for himself. (R8)

A few of the teachers suggested checklist as a structured kind of review:

> But you can also think that if you have included, if you have six areas that the students are to look at. «Has he commented programming (...) yes or no. Yes, he has done it, check, then that is fine. Does he have a tidy code or something like that, yes, check.» Then you get a, a type of score is okay. I think. It must not become too subjective. It must as much as possible be «this is ok, this is not ok, this is ok, this is not ok». (...) But it could be nice with a checklist because then the review is not completely random. I think it is important that there must be a very rigid structure. Proper criteria. Perhaps a bit form-based. (R3)

R6 emphasized that the students should know why they are giving reviews, making sure that their reviews follow a certain style: "*be clear on that anyhow it should be so that the reviews are possible to take use of.*" (R6)

Critique of having a structured review process, in lower secondary education, by R4:

> But gradually they will be able to understand and manage to weight it a bit themselves. How many positive things you say and how much that could be better. That the teacher does not need to be so rigid. The rigid system might contribute to that the review does not become that good, in lower secondary education, right. When they have to have this and this many things that have to be positive and one thing that could become better. (R4)

To face the challenge of students not being predisposed to writing reviews, R8 suggest having a structured review process with a guide or bullet points to go through:"*to have a good instruction about how to give a review. A kind of guide or a few bullet-points one has to go through. I think that is the most important. A good run-through of how you are to do it.*" (R8)

## 7.6 How to support teachers

The last part of the interview regarded how to support teachers if deemed feasible, see Section 7.1.2. The results are divided into four main parts here, but there was no single consensus among the respondents. The results varied from not needing any more support, to welcoming everything that helps.

### 7.6.1 No further support needed

A few of the respondents stated that they were ready to use PCR after the interview and being provided the information sheet about PCR. R6 stated that with the knowledge of the method provided in the interview, it is now usable: "*Well, it is just about getting the idea. And now I have gotten it.*" (R6)

### 7.6.2 A technical system/tool is paramount

Five of the respondents highlight a technical system as the main thing to help them in implementing PCR.

R2 emphasizes that a system that does the distribution for them is the most important thing to be able to implement PCR into their practice: "*But again as said, back to the system that has to be in place. (...) If the system is in place. Easy and simple, and it is safe to start using it.*" (R2)

Also stating that the system has to be safe, connecting this result to a technical challenge.

### 7.6.3 Any and all material helps

Premade information, tasks, forms and other material are welcomed and would be helpful. Examples could advantageously be in easily printable formats, and a fully written out explanation of PCR would be helpful. A full-out example of the implementation would also be helpful: "*all out concrete ways of how to use it. That are written all the way out, in a way. A lot of teachers, or, many are not able to picture it before they have seen it all the way.*" (R3) R4 summarizes the results quite well with:

> Well, everything is good to have. It is, really. Well, this model that you have made here, is also of help. To understand a bit what the thought is. A kind of explanation as well perhaps, if this is to be sent out to multiple teachers and not just me that sit here and look at it now, a model and a short explanation of how it works. Arguments of why to to it and challenges or dilemmas you need to be wary of, that can arise, but that you as a teacher have to try to deal with. I think that could be smart. (R4)

### 7.6.4 Information and promotion of PCR

Although many of the respondents highlighted a technical system/tool to help them in implementing PCR, many of them also stated that the main thing, which comes before the need of the system, was getting information about the method. R8 explains it like this, an explanation that encompasses many of the other responses: "*I did not know of this system from before. So the part about information about it and perhaps a simple approach with tips is perhaps the most important.*" (R8)

Information is good, but there also needs to be a promotion strategy in place for teachers to actually get the knowledge of the method. This could be in a schoolbook, suggested by two respondents, through teachers conferences, suggested by one respondent, but most importantly is the general idea of actually getting the idea: "*For teachers to know about this way of doing evaluations you have to read it somewhere.*" (R6)

### 7.6.5 Preparation course

Conducting a preparation course for the usage of PCR was an option only fronted by one of the respondents, but nevertheless an interesting idea. The respondent elaborated upon different ideas of how this could be done, one of the ideas being a test-run conducted by colleagues on each school: "*It does not need to be a course you have to go to. Help for an internal workshop, possibly one that comes in and leads it. Or if a teacher can go to something and come back and share it with the oter colleagues.*" (R5)

# Chapter 8

# Discussion

In this chapter, I will discuss the results. Not all of the results will be discussed, but parts that seem relevant to discuss either for being interesting, for their relation to existing literature, for their relation to the SLR results presented in Chapter 4 or because they are contradictory. The benefits and motivations are not that thoroughly discussed, as the RQs we are trying to answer are helped answered by the knowledge of what kinds of benefits and motivations we might expect from PCR in secondary education - and to a large degree that is sufficient. The status quo, challenges, implementation strategies and how to support teachers are more pivotal to discuss towards providing a sufficient answer to the RQ. The research questions in focus is:

RQ1.1: What do secondary school programming teachers know about PCR?

RQ2: How to support secondary school programming teachers in using Peer Code Review in their practice?

## 8.1 Key takeaways

### 8.1.1 Status quo of PCR in secondary education

Because the teachers easily relate to PCR and already are, perhaps unknowingly prior to the interviews, implementing informal PCR and generally have a positive attitude towards PCR, it seems obvious that this method should seriously be considered implemented in different CS courses throughout secondary education. The challenge of implementing PCR into other courses that have small parts of the curriculum as programming is one that this research will not address, as it might involve questions not explored in this study and the group of respondents would have been entirely differently chosen. Despite the above stated, none of the teachers were using PCR, making it a challenge to implement it from scratch. This calls for a foundation to build on, i.e. support of some kind.

### 8.1.2 Why should teachers implement PCR

Learning is at the center when teachers in secondary education discuss PCR, other motivations comes in second rank. PCR has the potential to be motivating as a variation both for teacher and for students. The main motivation is different for PCR in secondary education than those normally promoted for PCR and similar methods, see the next section for more on this.

### 8.1.3 Benefits to be gained

Learning was at the center of all of the benefits and should be a focus when implementing PCR. An unexpected finding, at least from reading relevant literature, was how the social aspects could be beneficial. This is interesting, as it might indicate that teachers of secondary school might have to consider the context of the social situation to a greater degree than what has to be done during code reviews in the software industry or the PCR implemented in higher education. Learning from reviewing was an expected result, however the potential for increased engagement in the reading of others code was somewhat unexpected. To properly reap the potential of this benefit, the student engagement needs to be facilitated for and the designation strategy should be thought of, more on this in Section 8.1.5. Learning from receiving reviews gained less attention than foreseen, as the reviewing part of the process got a lot more attention. This might be due to the teachers overly focusing on the "new" aspect to them, i.e. that the students get to see each others works. The learning aspect of receiving reviews comes to fruition when there is a good designation strategy in place, possibly just with having more than one reviewer per submission, but definitely if it takes the challenge of competence level differences into account. The reviews one gets should be followed up with the extension of the revision strategy, to create a real incentive for the students to actually make use of the reviews they receive. The time potential of the strategy is very connected to the availability of a technical solution, see Section 8.1.6 for more on this.

### 8.1.4 How to face the challenges

Different implementation challenges need to be thought of if they occur, but it is hard to take every thinkable challenge into account. The implementation challenges, one might argue, would largely be the same in most non-traditional didactical arrangements. When considering the level differences among students it is important to consider having a designation strategy, and inside that strategy have multiple reviewers. This will largely face the challenge, however this is a challenge that will never go away, but can be thought of and considered. The same goes for the lack of ability to code, which might often be connected to the level differences. The challenges of writing useful reviews should be met by training on reviewing and having a structured review process, more on this in Section 8.1.5. Low quality of reviews should be considered when dealing with the chal-

lenge of writing reviews, and the strategies are mostly the same to meet the two challenges, with the addition of adding a designation strategy and making number of respondents higher, more on this in Section 8.1.5. The challenge of students not being aware if their review is conducted on the same grounds the curriculum is possibly dealt with both by having training on reviewing, and by having a structured review process to create some frame to conduct the review within.

Maturity is something that is important for the individual teacher to consider - if the teacher thinks the students might not be mature enough to be parts of the PCR process, they should probably not go through with it. Creating a separation in the PCR process where all of the students get the same task, but some of the students participate in PCR and some do not is an interesting thought, but far too complex to discuss any further here. The social challenges of the method is something the individual teacher must evaluate if they think is going to be a major problem or not - the method can both have social benefits and challenges, the teacher should probably oversee the reviews. The strategy of anonymization might be one of interest to face this challenge, at least strongly mitigate the challenge in such a way that it should always be deemed to small of a challenge to be relevant in the choice of implementing the method or not. This is however subject to discussion, and it is paramount that the teacher is confident in their own ability to oversee the process and that the teacher is comfortable in the situation. Lack of student engagement brings many challenges in itself, but can to a degree be met to maintain the benefits of the process. When thinking of how to meet this method, it is largely the same as the facing of the social and maturity challenges. The teacher needs to know if the students are mature enough, which might vary in the individual group. This is especially true in the earlier levels of lower secondary education.

Technical challenges is a challenge one might see in the perspective of if there is, or is not, an available technical solution. An available, working technical solution would hopefully eliminate some of the technical difficulties teachers might encounter, but it is hard to imagine it would eliminate all challenges. This, again, makes for the confidence of the teacher in their own ability to push through challenges that might arise to be paramount if they are to implement PCR into their practice.

The time the method takes needs to be viewed in light of the strategies chosen, the size of the programming task undertaken and the time allotted to the process by the teacher. It also need to take into account the availability of a technical solution, as well as considering the possible time benefits and not only the costs. The cost might be further added to if there are technical challenges, again highlighting the importance of the teachers confidence in their own ability to see it through. If the teacher is not confident it is worth the time, despite all the benefits PCR seemingly would bring, PCR might not be worth to implement for that teacher. This might of course also be consistent in the mind of the teacher, but the same teacher might find PCR is too time consuming depending on the group of students. This concerns their age and maturity, their ability to code and their general

willingness to participate in didactical schemes.

### 8.1.5 Which implementation strategies to adopt

Assessment is a strategy that almost works like a double-edged sword it would seem from the results. It is possible to implement the strategy in a variety of ways, with different choices bearing different fruits. No assessment might be the low-risk, medium-reward option. The teacher should probably view their own assessment of deliveries during the semester as something they have to do sometime, meaning that there should not be any risk attached to the prospect of assessing the works as a teacher, indicating the strategy of assessing the work as a teacher is a viable option to consider. The assessment among the students is what at first glance appears as a high-risk high reward strategy. This part is however very up to how the teachers themselves feel about and consider the option of students evaluating each other - some might view this as a strategy that lowers the risk the method brings, even going as far as cementing the assessment the teacher conducts, while others might view it as problematic because they don't trust the students to assess each other. Either option seems to be fine, as it is only an extension of PCR and not a critical part to make the process work. It should not be presented as a recommended strategy, but an option for teachers that find that the benefits outweigh the negatives and thus finds it enriches the process.

Further discussing the anonymity extension, there seems to be little reason to avoid the anonymization of the process. Anonymization helps with mitigating social challenges, such as students feeling discomfort that their peers will read their work, pre-existing social conflicts and the students discomfort with providing honest reviews. The anonymization greatly depends on a technical solution to be easily implementable. That said, if it is available, the process seems to run itself. The only reason for not using it would be if the teacher wanting to implement PCR seeks to do an informal version, which they all already are doing. This indicates anonymizing is to be considered a smart implementation strategy.

Designation strategy and number of reviewers, the choice of how the reviews are distributed and how many reviews one has to do, are identified as important elements of PCR. The results from the interviews indicate that it is not clear what strategy to choose, but that there are multiple strategies that can be viable. The strategies mentioned in the results range from random designation, which was the most frequent one, to learning-pairs and rotating between reviewers. Number of reviewers is a strategy that might require time, but brings more benefits than it costs it would seem from the results. The strategy helps with challenges the level difference and the low quality of reviews, the results indicate. Based on this, it would perhaps be best to lean on the connected strategy of increasing the number of reviewers first and perhaps go with randomization as a designation strategy at first. This is in itself a strong argument for having a higher number of reviewers than one. Going too far and setting a too high number is arguably counterproductive the results indicate, and might even amplify some of the prior

challenges, e.g. low quality of reviews and low student engagement.

Learning to review seems like it should be implemented, but this strategy is important to look at while considering having a structured review process. Learning to review will mitigate some other challenges, but it might be time consuming. However, I would argue it is worth the cost because once some of the challenges are less prevalent the benefits will come to their own and the process might run smoother creating a better experience for all involved. Structuring the review process might be a way to skip the learning to review all by itself, but I would argue that one would quality assure the reviews to a greater extent by implementing both strategies. The strategy of learning to review might be time-consuming, but I would argue it is possible to create good reviews and to save some time as well by having a structured review process.

Revision should be implemented into the process, with the only downside as the time it takes. The indication is that the strategy brings more than it costs, making it viable for recommendation to teachers considering implementing PCR.

### 8.1.6 Supporting teachers

The result of already being capable to conduct PCR is based on the fact that the teachers got information from the interview, rendering this result only useful as an indication that some teachers might only need information and then be ready to implement PCR. It is important to realize that the result does not indicate that all teachers are already ready to implement it, as the indication from the study is that this is not a known method. One cannot implement PCR if one does not know about it. This brings us to the information and promotion of PCR being crucial if the method is to be used. A preparation course could be held, but that would bring much more logistical challenges than information and promotion. For the problem of what the information should contain, the result of any and all material being helpful is a good place to start. The teachers seemed to want a full explanation of the method, indicating it would be helpful to create this, plus recommendations of what strategies are recommended to extend PCR with and which ones can be considered if the teacher wants to use them. However the support of the teachers is done, the discussion above would indicate that it comes down to the individual teacher if they want to implement PCR.

## 8.2 Comparing the results

The results are in this section compared to existing material on PCR in secondary education, and PCR in higher education.

### 8.2.1 In relation to the existing material on PCR in secondary education

In this section the results are compared to the findings of the SLR presented in Chapter 4. This is done as an attempt to compare the results to existing literature on PCR in secondary education, yet, it is important to realize the limitations of this literature as the results in Chapter 4 show.

The first interesting find is the result of "learning to review", which is viewed as an extension. This extension was central in the findings from the SLR, and was also highlighted in the results from the interviews. This together indicates that it is a promising extension and should be further researched and used in preliminary practice implementing PCR. The teacher assessing the work was seen as an option in the results from the interviews, yet it was a part of the process from the SLR with the teacher also commenting on each of the phases.

The motivations behind conducting PCR seem to be at odds. The main motivation, seen again and again in the results here, is the learning aspect of PCR. In the SLR results the motivation for conducting PCR are largely linked to known advantages of the software industry. This is a very interesting result, although it needs further exploration to be of real significance. A minor result from this study was the motivation of students learning to read code because this was to be done later if the student continued with programming education, which is also seen as a significant factor in the results from the SLR.

Many of the benefits are the same, although with some differences. The benefits will not be compared one by one, but a general takeaway is that the results from the interview have a focus on social implications not found in the SLR. The results from the SLR also focus more on typical standards from the software industry as the ones from modern code review (Fatima et al. 2019, p. 2). The challenges are hard to compare, as this study has focused more on challenges and how to meet them, and the results from the SLR were more focused on the positive sides of PCR to effectively promote the method. One interesting comparison is that of the level difficulty of the task reported from the SLR results. This was not found in the results in this study, as this study merely reported the students level difference and lack of ability to code as challenges, and not the difficulty of the tasks themselves. This might be because the respondents did not see this as a valid problem. The challenges of low quality of reviews are however reported upon, but having already done training on reviewing their solution seems to be lowering the difficulties of the programming tasks.

The SLR did not provide any comparison of secondary and higher education regarding PCR, but the results from this study mentioned it to a limited degree. It is, however, important too recognize the limitation of this, as many of these views presented in the results might be strictly hypothetical as all of the teachers are in secondary education.

The feasibility is agreed upon. This study found that PCR is feasible and should be further implemented into secondary schools, which the results from the SLR

also did.

### 8.2.2 Results compared to higher education

There are both differences and similarities when comparing the results to PCR in higher education. One very interesting difference is the increased focus on learning in secondary education when compared to higher education. In higher education many of the benefits are closely linked to that of the software industry (Indriasari et al. 2020, p. 13), while the results presented here are very much so taking the setting of a classroom and all that brings into account. When comparing the challenges of PCR in higher education to the results of the interviews here, it shows that the teachers see many problems that are practical and very directly will affect them. This includes e.g. implementation challenges and technical challenges. The size of the programming task, and the lines of code it creates has not been mentioned in the results here, but seems to be important to consider in higher education (dos Santos and Nunes 2017, p. 28). The focus of the results here that are similar to this is to not have too many reviews to be done per student. Another interesting aspect is the mentioning of reading code, and learning to do it better, to prepare for an eventual later education within CS, which is also a factor in higher education (Li 2006, p. 3;Anewalt 2005, p. 150).

## 8.3 Implications of research

The findings of this study are important because they explore a relatively unexplored field. This research gives motivation for more research on the field and gives an indication of where to start - by teachers trying the method out. The findings somewhat differ from the literature referred to, mainly because the literature the SLR found does not have a comprehensive view of the challenges and literature on PCR in higher education is from a different educational standing point.

The research has the practical implications of quite comfortably being able to assume that the method is implementable in CS courses in secondary school. Limitations to the research is context of which the research was carried out, i.e. Norwegian secondary schools. It is plausible that the results are transferable to other countries, but this cannot be known by me as a Norwegian researcher conducting research in the Norwegian school system. For further on the reliability and validity of the research, see Chapter 9.

## 8.4 Further research

This project has brought interesting results with indications of what could be further researched. The research indicates a promising future for PCR in secondary education, the next step would be to implement the strategy in secondary education with data collection from both the teachers' perspective, as well as the

students. There are many focuses that are relevant. A study on PCR in secondary education in another context, e.g. another country, would also be very interesting. The action to be taken as the first step is, however, for a few teachers to try it out in their own way and for those teachers to recommend the method to other teachers. If that process were to continue, PCR would quickly gain a lot more traction than it currently has.

## 8.5 Towards practical guidelines

With the contents of this thesis, it could be possible to create some temporary guidelines for practical use. These guidelines would need to be evaluated and empirically tested. Creating a set of guidelines should include a simple description of how it could be done, with elaboration further down in the document for those who would want to read it. The readability of the guidelines would be paramount, as they should be easy to read and use. The guidelines could be built upon the conceptual framework in Chapter 2, the findings from the SLR in Chapter 4, the results from the interviews in Chapter 7 and the discussion in Chapter 8. The definition of PCR could come from the conceptual framework in Chapter 2, with the simple three-step method there. A visual representation could be built from the PCR information sheet the teachers were handed, as it was easily understood by the teachers during the interviews. Benefits could be based on the results from the interviews in Chapter 7, because the benefits need to reflect what the teachers see as valuable when considering to implement the method. It should which challenges to be prepared for should be based on what the reported ones by the teachers in Chapter 7. There could be a list of suggested expansions as a means to face those challenges. They could be based on the discussion in Chapter 8, which other possible expansions also could be. Technical systems would also be smart to have a recommendation of, as well as examples for each one. In the Norwegian context of schools the systems Canvas, ItsLearning and Google Classroom would be a good place to start with finding systems that integrate with these, as these are utilized in many schools. This was an intended contribution of this research, but due to time constraints it was not done. It seems from the results that the most important thing might be to find a way to promote the method.

# Chapter 9

# Ethics, validity and reliability

In this chapter the ethics, validity and reliability of the study is elaborated upon. When conducting research, in this case research where teachers are involved, many ethical considerations need to be taken and there are formal requirements to consider. The research also needs to have considered its own validity and reliability, through assessing if the research has been done rigorously and thoroughly.

## 9.1 Ethics

Many ethical aspects when conducting research, especially with people involved in the process. For this reason, it is helpful to have guidelines to abide by. This research has followed the guidelines of NESH, *Guidelines for Research Ethics in the Social Sciences, Humanities, Law and Theology* (NESH 2019). NESH has its own committee that lies under *The Norwegian National Research Ethics Committees*, who "shall ensure that all research is conducted in accordance with recognized research ethical norms" (Committees 2019).

In this research project, the ethical aspect brought into the light is that of the teachers, who were the respondents of the interviews. This includes having respect for the teachers' human dignity, their privacy, the right to be informed, gathering their informed consent, confidentiality, the limitations of the data, the storage of personal data (NESH 2019). To interview the teachers and gather their personal information I applied to NSD (NSD n.d.[a]). I gave the teachers required information and gathered their written consent prior to the recordings. The recordings of the interviews, and the personal information pointing to which was which, was encrypted and kept in different physical locations.

The teachers as individual humans were considered, as well as teachers as a group. Because all of the respondents are of the same profession, and all have programming of some sort, they are also to be looked at as a group when making ethical considerations. All teachers work at schools, making the interests of the school also prevalent in the research, and in need of consideration. Teachers were the main entity who had to be considered as a vulnerable group, and it is important that teachers do not feel degraded, disrespected, marginalized or feel like they

have gotten bad representation in the report. This is attempted to assure through following the five next steps.

### 9.1.1 NSD

All research conducted at NTNU has the mandatory obligation to apply to NSD, Norwegian Centre for Research Data, if deemed necessary (NSD n.d.[a]). NSD "ensures that data about people and society can be collected, stored and shared, both safely and legally, today and in the future (NSD n.d.[b]). For the application I made an interview-guide and an NSD information form, see Appendix A and C. The data I needed to collect for my project was the contact information of the respondents, i.e. their names and e-mail, the recordings of the interviews and their written consent. This was put into the application. The application was thoroughly filled out with the objective of the research, all the types of data to be collected and other relevant information such as length of the project. I got approval for the research project 25.09.20.

### 9.1.2 Information to respondents

In advance of the interviews an information form was sent out to the respondents. Information was sent in plain text over email, and the information was also sent as a more formal way of informing the respondents to what they were participating in. However, one of the respondents did not get this in advance of the interview due to planning difficulties. This is critiqueable, but the respondent gave verbal consent that the information was sufficient to receive in paper format the same day. In the email sending out that information, however, it was clearly stated twice that the respondents did not have to read any of the attached information before the actual interview. This was an effort to not to stress the teachers about the interview, clearly stating that except from being teachers in programming, they did not need any prior knowledge to partake in the interview However, they were given the option to ask questions and they did receive the information form stating what the interview would be about if they wished to prepare themselves. In the information sent out there was also an attached model of the PCR process, created as a means of scaffolding the teachers through the process to get their informed thoughts on PCR. This was not a required read before the interview. The information form was gone through at the start of the interview, in advance of the start of the recording. This took the time it needed, and was done to make sure we had a mutual understanding of what was going to happen. All respondents were clearly given the option of withdrawing their consent and withdrawing from the research project.

### 9.1.3 Consent

The research requires consent from the teachers. In the last page of the information for there was a line where the respondents had to sign, where they gave

their consent to the contents of the information form. There is a risk when giving consent to a semi-structured interview, because you don't have a fixed rubric of how the interview unfolds. Another risk is the validity of the informed consent, as the researcher might not want to give the respondents all information as it might affect their answers (Kvale and Brinkmann 2015, p. 105). These are however legitimate risks to take, as both my own assessment and NSD's assessment is that I have not collected sensitive data or information, when all of the data gets anonymized.

### 9.1.4 Data processing and storage

The data was processed and stored in a meticulous manner. There are several guidelines and laws to follow. In this research project I have trusted that NSD and the guidelines from NTNU have properly assured the data has been processed and stored correctly. The research has followed the NTNU's guidelines of GDPR and privacy (NTNU n.d.[e]). This regards many aspects, one of them is the requirement of not storing personal data longer than it is strictly needed for the project. The collection of data has followed NTNU's procedures for this, having requirements of which units to collect the data with (NTNU n.d.[b]). The subsequent data processing and storage also have guidelines to follow, having strict regulations for how to store the data (NTNU n.d.[a]). I have, in line with the guidelines, chosen to store it on separate physical locations with NTNU's own method of encrypting the files (NTNU n.d.[d]). The recordings were transcribed and anonymized. There was created a mapping of aliases to interviews in an encrypted file, separated from the other data. All information that could in some way lead to the respondent being identified has been changed, removed or in some way altered in ways that do not affect the rest of the dataset, whether it was name, name of school, years of work at a certain place or some other information. All of this was done to protect the integrity of the respondents.

### 9.1.5 Reporting of data

The reporting of the data took the respondents, teachers as a group and schools as institutions into consideration. When reporting the data the respondents need to feel like their given information is being reported in a representative way (Robson and McCartan 2016, p. 489). The data was only used in this project, and the respondents of the interview know this. This is hopefully assuring in itself for the respondents. To have met in person, talked to and feel like they have a personal connection to me as a researcher should also help. The anonymization of the data makes the data confidential (Kvale and Brinkmann 2015, p. 106). Although the interviews were done in Norwegian, for the sake of the flow and readability of the report, when used in the report information and quotes have been translated to English. This is itself a challenge, as part of the intention of what was said can be lost in translation, and the respondents might feel like their data is being manipulated. This, however, is something I have had in mind when analysing the

data and reporting it, somewhat minimizing the risk of losing the intent of the spoken word in translation. The data needs to be reported in a way that ensures the integrity of the respondents, and because all respondents are teachers, in a way that does not paint a negative picture of teachers as a group. This also applies to the individual, as one respondent should not read a report where they might feel like they are being made to look bad. Teachers are human being and can make mistakes, meaning that although they might have bad practice, lack knowledge of a certain subject or have made mistakes in their career or life, they should not feel that the report makes them look bad. However, the report should not appear biased, making for the reporting having tried to be precise and explanatory. If one respondent is reported on I have tried to not generalize the results, and to keep the respondents integrity.

## 9.2 Validity and reliability

In a research project such as this it is important to assess its validity and reliability. Validity and reliability relates to the quality of the research. This includes the validity and reliability of the results produced. Validity regards the truth, accuracy and strength of the results produced (Kvale and Brinkmann 2015, p. 276). These terms encompass that the produced material is only as good as the premises it was derived from, if the correct approach was taken and if the research regards what it intended researched (Kvale and Brinkmann 2015, p. 276). Reliability regards the produced results' consistency and credibility (Kvale and Brinkmann 2015, p. 276). This concerns if the results are possible to reproduce, and in this project how the interviews were carried out (Kvale and Brinkmann 2015, p. 276). All of these aspects were taken into consideration, and will be further discussed in the coming paragraphs.

The method for collecting data was interviews, but others were considered. Some of these were having workshops and focus-groups with teachers or other types of experts on the field. Focus-groups are essentially multi-person interviews where the situation changes from a normal interview and the interviewer works more like a moderator, but they were not chosen as they give the researcher less control over the flow of the conversation than traditional interviews (Kvale and Brinkmann 2015, pp. 179–180). Implementation of the method into the classroom was also considered, with methods such as interviews of teachers, students and possibly questionnaires being thought of as viable approaches to gain valuable insights. These options were however quickly discarded as the initial phases of the project showed little empirical research to build an implementation on. One might argue that more hands-on results would have been produced with an implementation, but interviews were deemed more feasible as the focus of the study is a method that is in its early stages when it comes to secondary education.

The interviews were set up quite openly, not trying to hide anything about the method. This openness of the interviews allowed for the respondents to largely make an informed consent. It does however pose some quite serious questions

about the results. Would the results be the same if the questions had been structured in a way that did not reveal how the method was set up, and the challenges reported in higher education? The results might have been slightly different, because it might have altered the flow of the conversation. It is however up to the interviewer how this is done, with both an "open" approach and a more "closed" approach being acceptable approaches (Kvale and Brinkmann 2015, p. 162). However, the respondents were not presented with all of the known literature that I as a researcher know of. This can be seen as an argument that they somewhat were kept in the dark, and that one can never really keep them in on the entire loop as they then would have to be experts on the field. The results were not gone through with the respondents ahead of the writing of the report. This was done because the projects' size was deemed small enough that it is not a major issue.

The group of respondents chosen is worthy of critique. All of the respondents are teachers who got the option of participating. This means that they all volunteered, perhaps indicating that this exact group of respondents are more eager to learn new methods and are more open to new ways of seeing things. This is important to keep in mind when assessing the quality of the results. Aside from this, it is important to keep in mind that the respondents all had different backgrounds, subjects taught, years of teaching, years of teaching programming and so forth. They all came into the interviews with different perspectives. The group of teachers are all teachers in the Norwegian school system, suggesting the representativeness of the respondents to be questioned. When assessing the quality of the qualitative data analysis it is important to assess the data quality, test the patterns found and test the explanations made (Kvale and Brinkmann 2015, pp. 479–480).

I myself conducted all of the interviews, but did not transcribe all of the interviews myself. This might have been a mistake, as the transcription process is an important step of the initial phase of the analysis conducted. This was a concern that was addressed by me going through all of the transcriptions, allowing me to make the argument that I did get a good understanding of the contents of the transcriptions from early on as well as quality-checking the transcriptions. More on this in Section 5.3.6.

The analysis has some challenges. One question is the flexibility that thematic coding allows for (Kvale and Brinkmann 2015, p. 470). The flexibility can be hard to deal with as a researcher when choosing which aspects to focus on. The analyst is important in thematic coding, and human deficiencies are therefore also important to consider and be aware of as a researcher (Kvale and Brinkmann 2015, p. 462).

# Chapter 10

# Conclusion

This research project has focused on PCR in secondary education. This was based on the following problem statement: What is known about PCR in secondary education, and how can we support teachers in implementing it? To address this problem statement, two main research questions were created. These research questions are answered here.

> **RQ1:** What is the current knowledge of PCR in secondary education?

This research question was split into two: the literature and the real-world context. Each of these got its own sub-question.

> **RQ1.1:** What is the state of the art in the research about PCR in secondary education?

To answer this question a SLR regarding PCR in secondary education was conducted to find out what is known about PCR in secondary education and to which degree it is used. From this SLR the main result was that the literature on PCR in secondary education is very scarce as only four papers were selected, confirming that there exists a gap in the literature on PCR in secondary education. Other results included that there is no clear consensus about a definition of PCR and only two studies report empirically using PCR in secondary education. Further, some different implementation strategies were used in the literature, but there was no clear consensus of what was the best practice. The motivations for implementing PCR was closely linked to that of peer reviews and the software industry's code reviews. The papers reported many benefits such as promoting student learning, improved code quality, and different benefits from reviewing. Challenges reported in the papers were not as much elaborated upon, yet they reported challenges such as how difficult the tasks should be, negative student attitude towards the process, and time demands both in class and for the teacher. The studies also indicate that PCR is applicable in secondary education.

To answer the next two questions, RQ1.2 and RQ2, qualitative semi-structured interviews with eight respondents were carried out. The interviews were transcribed and analyzed, to find answers to the two questions.

> **RQ1.2:** What do secondary school programming teachers know about PCR?

Programming teachers are not familiar with PCR, but easily relate to the method. They are not using PCR, but they are implementing an informal version of PCR where they encourage their students to help each other when they encounter problems. They also generally show a positive attitude towards the method, indicating that PCR should be considered implemented in different courses throughout secondary education. Because PCR is not a well-known method generally being used by teachers, implementing PCR can be challenging. This calls for support in using the method.

> **RQ2:** How to support secondary school programming teachers in using PCR in their practice?

To figure out how to support secondary school programming teachers in using PCR, finding out which motivations, benefits, challenges, and implementation strategies related to the method was pivotal. The results indicate that the main motivation for implementing it was learning. The method can be motivating for students and the teacher, but learning is prioritized. Benefits teachers connect to the method include beneficial social implications, learning from reviewing, and increased student engagement. More benefits are to be gained with the right implementation strategies. PCR has a set of challenges, which include level differences among students, students not having the ability to write reviews, low quality of reviews, the maturity of the students, social challenges, lack of student engagement, technical challenges, and being time-consuming both for the teacher and in class.

When adopting PCR in their practice, secondary school teachers need to choose good implementation strategies. The results from this study indicate that many of the challenges can be met in a good way, at least mitigated, when implementing PCR. Strategies the results indicate would be beneficial to adopt include having more than one reviewer, thinking about who reviews who, having a session before the actual PCR process where the students learn to review, structuring the review process, and implementing a revision step at the end of the process. Another thing to consider is adding an assessment by either the reviewer or the teacher, if the teacher wants to. Any way these motivations, benefits, challenges, and implementation strategies are viewed by the teachers, one of the most important takeaways is that the strategy is dependent on the teachers' confidence in their own ability to conduct the method.

Teachers lack the information about the method. As teachers easily relate to PCR and show positive attitude towards it, what they lack is knowledge of the method. To support them it would be beneficial with promotion of PCR, examples and other material connected to PCR such as forms to use in the review process and a set of best practices to adopt.

A very important limitation of this study is that it was done in a all-around Norwegian context, both the researcher, the respondents and the school system

were Norwegian. This might indicate that it is not directly transferable to other countries.

# References


Aalberg, T. & Lorås, M. (2018). Active learning and student peer assessment in a web development course [Publisher: Bibsys Open Journal Systems]. *Active learning and student peer assessment in a web development course*.

Alfasoft. (n.d.). NVivo for tekstanalyse og kvalitativ analyse av video og lydopptak. Retrieved November 26, 2020, from https://www.alfasoft.com/no/produkter/statistikk-og-analyse/nvivo.html

Allen, H., Cury, A., Gaston, T., Graf, C., Wakley, H. & Willis, M. (2019). What does better peer review look like? Underlying principles and recommendations for better practice [Place: Chichester, UK Publisher: Wiley]. *Learned publishing*, *32*(2), 163–175.

Almeida, J. R. d., Camargo, J. B., Basseto, B. A. & Paz, S. M. (2003). Best practices in code inspection for safety-critical software [Conference Name: IEEE Software]. *IEEE Software*, *20*(3), 56–63. https://doi.org/10.1109/MS.2003.1196322

Anewalt, K. (2005). Using peer review as a vehicle for communication skill development and active learning. *Journal of Computing Sciences in Colleges*, *21*(2), 148–155.

Badampudi, D., Britto, R. & Unterkalmsteiner, M. (2019). Modern code reviews - Preliminary results of a systematic mapping study, In *Proceedings of the Evaluation and Assessment on Software Engineering*, Copenhagen Denmark, ACM. https://doi.org/10.1145/3319008.3319354

Blum, K., Jacobs, W., Modestino, E. J., DiNubile, N., Baron, D., McLaughlin, T., Siwicki, D., Elman, I., Moran, M., Braverman, E. R., Thanos, P. K. & Badgaiyan, R. D. (2018). Insurance Companies Fighting the Peer Review Empire without any Validity: The Case for Addiction and Pain Modalities in the face of an American Drug Epidemic. *SEJ surgery and pain*, *1*(1), 1–11.

CASP. (n.d.). *CASP CHECKLISTS* [CASP - critical appraisal skills programme]. Retrieved December 12, 2020, from https://casp-uk.net/casp-tools-checklists/

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit [Publisher: American Psychological Association]. *Psychological bulletin*, *70*(4), 213–220. https://doi.org/10.1037/h0026256

Committees, T. N. N. R. E. (2019). Who are we and what do we do? Retrieved December 10, 2020, from https://www.forskningsetikk.no/en/about-us/who-are-we-and-what-do-we-do/

dos Santos, E. W. & Nunes, I. (2017). Investigating the effectiveness of peer code review in distributed software development [event-place: Fortaleza, CE, Brazil], In *Proceedings of the 31st brazilian symposium on software engineering*, New York, NY, USA, Association for Computing Machinery. event-place: Fortaleza, CE, Brazil. https://doi.org/10.1145/3131151.3131161

Fagan, M. E. (1976). Design and code inspections to reduce errors in program development [Conference Name: IBM Systems Journal]. *IBM Systems Journal*, *15*(3), 182–211. https://doi.org/10.1147/sj.153.0182

Fatima, N., Nazir, S. & Chuprat, S. (2019). Understanding the Impact of Feedback on Knowledge Sharing in Modern Code Review, In *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*. https://doi.org/10.1109/ICETAS48360.2019.9117268

Fu, Q., Grady, F., Broberg, B. F., Roberts, A., Martens, G. G., Johansen, K. V. & Loher, P. L. (2017). Code review and cooperative pair programming best practice.

Hamer, J., Cutts, Q., Jackova, J., Luxton-Reilly, A., McCartney, R., Purchase, H., Riedesel, C., Saeli, M., Sanders, K. & Sheard, J. (2008). Contributing student pedagogy. *ACM SIGCSE Bulletin*, *40*(4), 194–212. https://doi.org/10.1145/1473195.1473242

Indriasari, T. D., Luxton-Reilly, A. & Denny, P. (2020). A Review of Peer Code Review in Higher Education [Place: New York, NY, USA Publisher: Association for Computing Machinery]. *ACM Trans. Comput. Educ.*, *20*(3). https://doi.org/10.1145/3403935

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele Univ.*, *33*.

Kunnskapsdepartementet. (2020). Rekordmange søkere til høyere utdanning [Publisher: regjeringen.no]. Retrieved December 9, 2020, from https://www.regjeringen.no/no/aktuelt/rekordmange-sokere-til-hoyere-utdanning/id2699329/

Kvale, S. & Brinkmann, S. (2015). *Det kvalitative forskningsintervju* (3. utg.). Oslo, Gyldendal akademisk.

Li, X. (2006). Using peer review to assess coding standards - a case study, In *Proceedings - frontiers in education conference, FIE*, Institute of Electrical; Electronics Engineers Inc. https://doi.org/10.1109/FIE.2006.322572

MacLeod, L., Greiler, M., Storey, M.-A., Bird, C. & Czerwonka, J. (2018). Code Reviewing in the Trenches: Challenges and Best Practices [Publisher: IEEE Computer Society]. *IEEE Software*, *35*(4), 34–42. https://doi.org/10.1109/MS.2017.265100500
cited By 25

NESH. (2019). Guidelines for Research Ethics in the Social Sciences, Humanities, Law and Theology. Retrieved December 10, 2020, from https://

www.forskningsetikk.no/en/guidelines/social-sciences-humanities-law-and-theology/guidelines-for-research-ethics-in-the-social-sciences-humanities-law-and-theology/

NSD. (n.d.[a]). Norsk senter for forskningsdata. Retrieved November 19, 2020, from https://nsd.no/

NSD. (n.d.[b]). Norwegian Centre for Research Data. Retrieved December 10, 2020, from https://nsd.no/en/

NTNU. (n.d.[a]). Behandle personopplysninger i student- og forskningsprosjekt - Wiki - innsida.ntnu.no. Retrieved November 19, 2020, from https://innsida.ntnu.no/wiki/-/wiki/Norsk/Behandle+personopplysninger+i+forskningsprosjekt

NTNU. (n.d.[b]). Datainnsamling. Retrieved November 19, 2020, from https://innsida.ntnu.no/wiki/-/wiki/Norsk/datainnsamling

NTNU. (n.d.[c]). *Department of computer science (IDI) - NTNU*. Retrieved December 7, 2020, from https://www.ntnu.edu/idi

NTNU. (n.d.[d]). Kryptering av filer. Retrieved November 19, 2020, from https://innsida.ntnu.no/wiki/-/wiki/Norsk/Kryptering+av+filer

NTNU. (n.d.[e]). Personvern og GDPR. Retrieved November 19, 2020, from https://innsida.ntnu.no/wiki/-/wiki/Norsk/Personvern+og+GDPR

Prince, M. (2004). Does Active Learning Work? A Review of the Research. *Journal of Engineering Education*, *93*(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Rigby, P. C. & Bird, C. (2013). Convergent contemporary software peer review practices, In *Proceedings of the 2013 9th Joint Meeting on Foundations of Software Engineering*, New York, NY, USA, Association for Computing Machinery. https://doi.org/10.1145/2491411.2491444

Robson, C. & McCartan, K. (2016). *Real world research : A resource for users of social research methods in applied settings* (4th ed.). Chichester, Wiley.

Sabarinath, R. & Quek, C. L. G. (2020). A case study investigating programming students' peer review of codes and their perceptions of the online learning environment. *Education and Information Technologies*, *25*(5), 3553–3575. https://doi.org/10.1007/s10639-020-10111-9

Sentance, S. & Csizmadia, A. (2017). Computing in the curriculum: Challenges and strategies from a teacher's perspective [ISBN: 1360-2357 Publisher: Springer]. *Education and Information Technologies*, *22*(2), 469–495.

Søndergaard, H. & Mulder, R. A. (2012). Collaborative learning through formative peer review: Pedagogy, programs and potential [Publisher: Routledge _eprint: https://doi.org/10.1080/08993408.2012.728041]. *Computer Science Education*, *22*(4), 343–367. https://doi.org/10.1080/08993408.2012.728041

Topping, K. (2016). Peer Assessment Between Students in Colleges and Universities [Place: Thousand Oaks, CA Publisher: Thousand Oaks, CA: American Educational Research Association (AERA)]. *Review of educational research*,

*68*(3), 249–276. https://doi.org/10.3102/00346543068003249 1935-1046

Utdanningsdirektoratet. (2006a). Læreplan i informasjonsteknologi - programfag i utdanningsprogram for studiespesialisering (INF1-01). Retrieved December 6, 2020, from https://www.udir.no/laring-og-trivsel/lareplanverket/finn-lareplan/lareplan-hele/

Utdanningsdirektoratet. (2006b). Læreplan i informasjonsteknologi - programfag i utdanningsprogram for studiespesialisering (INF1-01). Retrieved December 6, 2020, from https://www.udir.no/laring-og-trivsel/lareplanverket/finn-lareplan/lareplan-hele/

Utdanningsdirektoratet. (2016). Forsøkslæreplan i valgfag programmering (PRG1-01). Retrieved December 9, 2020, from https://www.udir.no/laring-og-trivsel/lareplanverket/finn-lareplan/lareplan/

Utdanningsdirektoratet. (2020). Læreplan i valgfaget programmering (PRG01-02). Retrieved December 6, 2020, from https://www.udir.no/lk20/prg01-02

Utdanningsdirektoratet. (n.d.). Fagfornyelsen – prosess for å utvikle nye læreplaner. Retrieved December 9, 2020, from https://www.udir.no/laring-og-trivsel/lareplanverket/fagfornyelsen/

Yanqing, W., Hang, L., Yanan, S., Yu, J. & Jie, Y. (2011). Learning outcomes of programming language courses based on peer code review model, In *2011 6th International Conference on Computer Science Education (ICCSE)*. https://doi.org/10.1109/ICCSE.2011.6028746

Zotero. (n.d.). *Zotero | your personal research assistant*. Retrieved December 11, 2020, from https://www.zotero.org/

# Appendix A

# Interview guide

## Før opptaksstart

### Formalia

- Formålet med intervjuet
- Hvordan intervjuet dokumenteres og hva det brukes til
- Mulighet for å trekke seg underveis
- Mulighet for å unngå å svare på spørsmål
- Mulighet for å trekke seg i etterkant, ta kontakt
- Informasjonen behandles konfidensielt - taushetsplikt
- Data anonymiseres

### Forklaring av intervjuets gang

- Ute etter erfaringer, refleksjoner, tanker - tenk høyt
- Peer code review (medstudentvurdering i programmering)
- Noe kan være vanskelig å svare på, men det er et veldig viktig resultat for meg at det er ting lærer ikke syns er lett å svare på - forklar om du syns noe er fjernt å svare på, eksempelvis
- Spørsmål og oppklaring før start?

## Opptak startes

### Åpningsspørsmål

- Hvilke fag underviser du i?
- Hvor lenge har du jobbet som lærer?
- Har du noen formell kompetanse i informatikk/programmering?
- Hvor lenge har du undervist programmering?

### Egne erfaringer med peer code review

- Har du undervist og brukt metoden "peer code review" (norsk: medstudentvurdering i programmering)?
  - Hvis ja:
    - Hvilke klassetrinn? (på hvilke utdanningsnivå)
    - Hvorfor?
  - Hvis nei:
    - Hvorfor ikke?

### Hva er peer code review (Scaffolding) (forklares mens de får utdelt ark som gir eksempel på prosessen)

1. Elev skriver kode
2. Elev leverer kode i et system (eller til lærer)
3. Elev får utdelt en annen elev sin kode og "reviewer" denne (skriver en tilbakemelding)
4. Tilbakemeldingen blir levert tilbake og elevene leser over tilbakemeldningen de har fått på koden sin

#### Eksempel på muligheter for å utvide prosessen

- Alle elever vurderer 3 innleveringer, slik at alle får 3 tilbakemeldinger
- Man legger inn en faktisk vurdering, slik at elevene får muligheten til å vurdere hverandres kode med en score, samt at de får muligheten til å si ifra til lærer om de ikke er fornøyd med en medelevs vurdering av deres kode
- Det blir lagt inn et iterativt steg der elevene skal forbedre koden sin etter de har fått tilbakemelding

### Generelt om peer code review

- Hvilke fordeler ser du at peer code review har?
  - (læringsutbytte, tidsbruk, aktiv læring, tilbakemelding, vurdering)
  - (tidseffektivt eller tidssluk?)
- Hvilke utfordringer ser du at peer code review har?
  - (tidsbruk, vurdering, "peer designation" - random, 3-1, one level higher)
- Ser du forskjellige fordeler og utfordringer på ulike utdanningsnivå?
  - (ungdomsskole, VGS, høyere utdanning)

### Infobit

PCR har lite litteratur om seg på ungdomsskolenivå og VGS, viser et systematisk litteratursøk jeg har gjennomført. Det mangler derfor også et grunnlag å bygge på, derfor forsøker vi å undersøke med bakgrunn i de utfordringer som er blitt rapportert om i høyere utdanning. Spørsmålene nå vil altså ha fokus på å se på disse utfordringene på ungdomsskole og VGS-nivå.

### Hvordan eventuelt tilrettelegge for peer code review

- Vil du si at peer code review er praktisk gjennomførbart på de forskjellige utdanningsnivåene?
  - (ungdomsskole, VGS)
  - (Tidsbruk i faget, elevforskjeller, tidsbruk å planlegge)
- Reflekter rundt følgende problemer, og hvordan det gitte problemet kunne sett ut og blitt håndtert på ungdomsskolen/VGS:
  - Elevers mangel på kunnskap og evne til å skrive en skikkelig tilbakemelding
  - Lav interesse for å skrive en tilbakemelding (ekstra jobb, lav interesse for faget, etc)
  - Lav kvalitet på tilbakemeldinger (for lite arbeid lagt ned, etc)
  - Ueffektiv administrativ prosess for læreren (tidskrevende, etc)
  - Andre faktorer som tidsmangel for gjennomføring av opplegg, arbeidsmengde for elever, for lite tid gitt til elever for å gjøre arbeidet skikkelig
- Hvordan kunne det eventuelt blitt tilrettelagt for at du skulle brukt peer code review i ditt virke som lærer?
  - (Rammeverk? Retningslinjer (guidelines)? Ferdige opplegg?)

### Avslutningsvis

- Er det noe du sitter inne med nå, enten angående peer code review eller noe relatert?

# Appendix B

# PCR information sheet

The following spreadsheet was handed to the teachers during the Scaffolding phase of the interview. It provides a visual representation of how the PCR process is.

## Peer code review

| Steg | Lærer | | Elev n | | System | | *Eksempel for utvidelse av PCR* | |
|---|---|---|---|---|---|---|---|---|
| | Lage oppgave | -> | Løs programmingsoppgave | | | | | |
| | | | ↓ | | | | | |
| 1 | | | Lever oppgave i system | -> | Motta oppgaver | | | |
| | | | | | ↓ | | | |
| | | | Skriv tilbakemelding på utdelt oppgave | <- | Distribuer alle oppgaver til tilgjengelige elever | | Inkluder en vurdering som skal settes (score) | Gi hver elev flere oppgave å gi tilbakemelding på |
| | | | ↓ | | | | ↓ | ↓ |
| 2 | | | Lever tilbakemelding i system | -> | Motta tilbakemeldinger | | Mulighet for å "flagge" en vurdering, si ifra til lærer om at de er misfornøyde | Hver elev får flere tilbakemeldinger |
| | | | | | ↓ | | | |
| | | | Motta tilbakemelding | <- | Send tilbakemeding tilbake til deltakere | | | |
| | | | ↓ | | | | | |
| 3 | | | Les tilbakemelding | | | | Legg til et ekstra steg der elevene skal revidere koden | |
| | | | | | | | ↓ | |
| | | | | | | | Lever kode på nytt til lærer | |

# Appendix C

# Interview information form

The following information document was sent to the teachers in advance of the interview, and handed out before the interview began.

# Vil du delta i forskningsprosjektet “How to support teachers in using Peer Code Review”?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å kartlegge bruk av medstudentvurdering i programmering, eller “Peer code review” (PCR), i ungdomsskolen og videregående skole, for så å bygge videre på denne kunnskapen for å støtte lærere i deres bruk av metoden. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

## Formål

Programmering er et fag som byr på både muligheter og utfordring når det gjelder å skape et engasjerende undervisningsopplegg. PCR er en metode som er foreslått å kunne hjelpe å talke noen av disse utfordringene på en hensiktsmessig måte. Formålet med dette intervjuet er å kartlegge bruken av Peer code review på ungdomsskolen og i videregående skole, og bakgrunn for eventuell bruk.

To av hovedspørsmålene vi prøver å belyse vil være forskningsspørsmålene:

> *Har du brukt metoden medstudentvurdering i programmering?*
>
> *Hvilke utfordringer og muligheter ser du med medstudentvurdering i programmering?*

Prosjektet er tilknyttet Magnus Tvildes masteroppgave høsten 2020.

**Hvem er ansvarlig for forskningsprosjektet?**

Institutt for datateknologi og informatikk ved NTNU Gløshaugen er ansvarlig for prosjektet.

**Hvorfor får du spørsmål om å delta?**

Du får spørsmål om å delta fordi du har jobber som lærer i informatikkfag på ungdomsskolen eller videregående skole. I første omgang får 10 lærere denne henvendelsen.

## Hva innebærer det for deg å delta?

Metoden for datainnsamling som brukes i dette prosjektet er et en-til-en intervju. Opplysningene som innhentes vil være et samtykkeskjema med ditt navn og signatur på, din profesjon og hvor lenge du har vært lærer.

- Hvis du velger å delta i prosjektet, innebærer det at vi avtaler et tidspunkt å gjennomføre intervjuet. Intervjuet vil vare i ca. 30 minutter, avhengig av hvordan samtalen utarter seg. Intervjuet vil bli tatt lydopptak av som lagres elektronisk, for å kunne transkribere dette i ettertid. Intervjuet inneholder spørsmål direkte knyttet opp mot det som står under "Formål". Ved behov kommer jeg også til å ta notater under selve intervjuet.

**Det er frivillig å delta**

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykke tilbake uten å oppgi noen grunn. Alle opplysninger om deg vil da bli anonymisert. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

**Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger**

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- Veileder for prosjektet, Monica Divitini ved Institutt for datateknologi og informatikk, vil ha tilgang til all data tilknyttet prosjektet
- For å sikre at ingen uvedkommende får tilgang til personopplysningene i prosjektet vil navnet og kontaktopplysningene dine bli erstattet med koder som lagres på egen navneliste adskilt fra øvrige data. Alle slike data vil også til enhver tid være lagret på krypterte enheter.

Ved en eventuell, men høyst usannsynlig, publikasjon vil alle data som kan indikere personer være utenfor publikasjonen.

**Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?**

Prosjektet skal etter planen avsluttes 18.12.2020. Etter avsluttet prosjekt vil personopplysninger anonymiseres og lydopptak av intervju vil slettes.

**Dine rettigheter**

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg,
- å få rettet personopplysninger om deg,
- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

**Hva gir oss rett til å behandle personopplysninger om deg?**

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Institutt for matematiske fag ved NTNU har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

**Hvor kan jeg finne ut mer?**

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

· Institutt for datateknologi og informatikk, NTNU ved professor Monica Divitini. Epost: divitini@ntnu.no, tlf: 735 94 462.

· Vårt personvernombud: Thomas Helgesen. Epost: thomas.helgesen@ntnu.no, tlf: 930 79 038

· NSD – Norsk senter for forskningsdata AS, på epost (personverntjenester@nsd.no) eller telefon: 55 58 21 17.

Med vennlig hilsen

| | |
|---|---|
| Prosjektansvarlig | Student |
| Monica Divitini | Magnus Tvilde |
| (Veileder) | |

---

# Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet “How to support teachers in using Peer Code Review”, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i intervju

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, ca. 18.12.2020.

---

(Signert av prosjektdeltaker, dato)

# Appendix D

# Codebook

This is the final version of the codebook used in the analysis.

# Interviews

## Nodes

| Name | Description |
| --- | --- |
| Benefits with PCR | |
| Code readability | |
| Improved code quality | |
| Improved student engagement | |
| Learning from reviewing | |
| Learning programming | |
| Level closeness among students | |
| Other benefits | |
| Reading code skills | |
| Receiving reviews | |
| Review skills | |

| Name | Description |
| --- | --- |
| Social benefits | |
| Soft skills | |
| Time effectivity for teacher | |
| Time effectivity in course | |
| Challenges with PCR | |
| Administrative burden on teacher | |
| Implementation challenges | |
| Lack of ability to code | |
| Lack of ability to write reviews | |
| Lack of student engagement in method | |
| Level difference among students | |
| Low quality of reviews | |
| Maturity | |
| Other challenges | |

| Name | Description |
| --- | --- |
| Social challenges | |
| Technical challenges | |
| Time-consuming in course | |
| Education level | |
| Comparing different levels | |
| Higher education | |
| Secondary education | |
| How to support teachers | |
| Feasibility of PCR | |
| Preparation course | |
| Prepared material | |
| Prepared tasks | |
| Promotion of method | |
| Tool or technical solution | |

| Name | Description |
| --- | --- |
| Implementation strategies of PCR | |
| Anonymity | |
| Assessment | |
| Creating benefits | |
| Designation strategy | |
| Facing challenges | |
| Revision | |
| Structured review process | |
| Training on reviewing | |
| Motivation behind PCR | |
| Developing soft skills | |
| Learning | |
| Learning industry standards | |
| Motivating for students | |

| Name | Description |
| --- | --- |
| Motivating for teacher | |
| Other motivations | |
| Status quo | |
| Attitude towards PCR | |
| Informal PCR | |
| Knowledge of PCR | |
| Reported use of PCR | |
| Similar methods to PCR | |

# Appendix E

# NVIVO coding example

This is an example of how the transcriptions were coded.

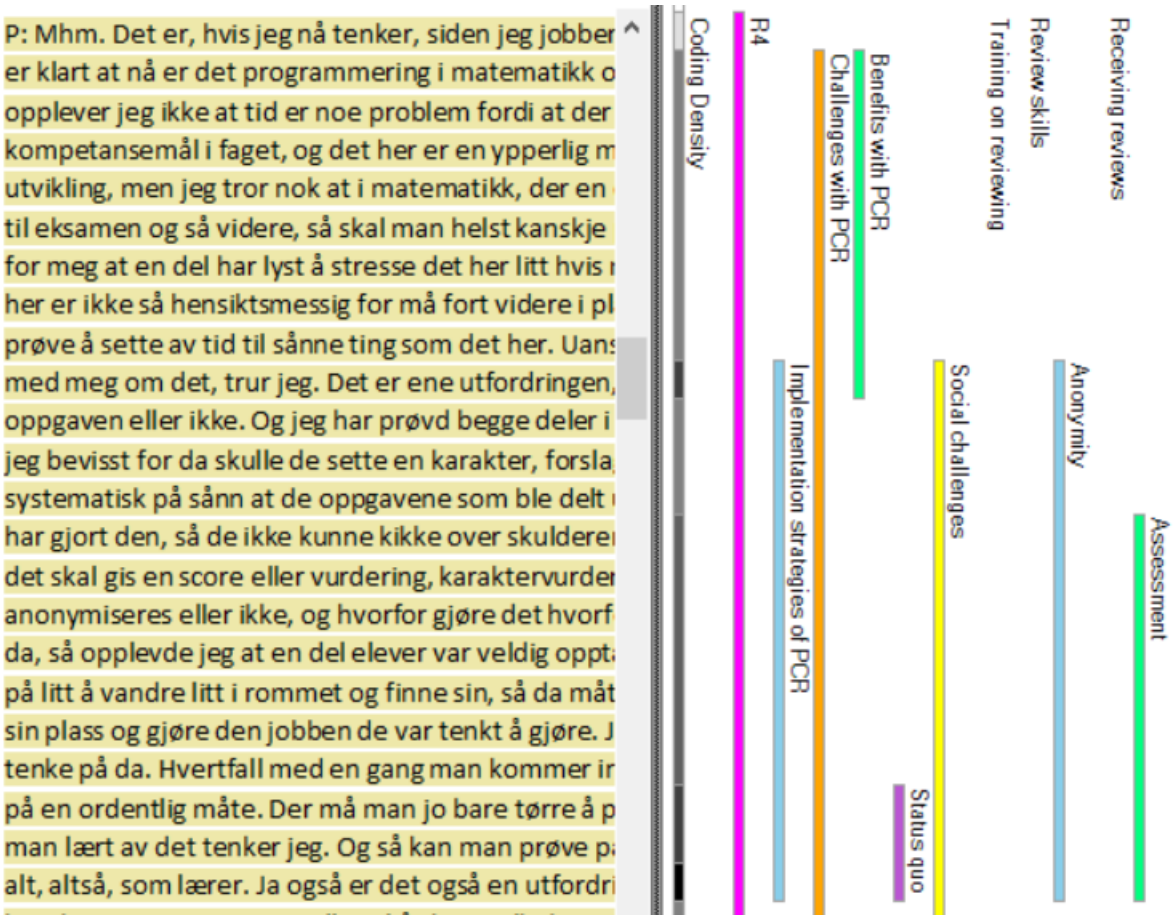


**Figure E.1:** NVIVO coding example

Magnus Tvilde

# Peer Code Review in Secondary Education: A Systematic Literature Review and an Interview Study

NTNU
Norwegian University of
Science and Technology