

# A Model for Short Peer-Assessment

Kjellove Bjørge

05.08.2020



## 1. Abstract

Formative assessment has been proven to have positive impacts on students' learning gains. There has been an increased focus on formative assessment, both nationally and internationally. All of the potential effects of formative assessment are not fully discovered, and there are untested areas that can potentially have a great effect on student learning gains. Self-assessment is an integral part of formative assessment and is needed in a future where learning strategy is vital for dealing with a rapidly changing society. Peer-assessment, the act where students provide feedback on other students' work, has shown to have significant benefits for students in several areas. One of these benefits is related to self-assessment. By providing feedback to others, students are raising their skills in self-assessment. This study explores a possible avenue of utilizing a digital scaffold to conduct short peer-assessments by proposing a model for short peer-assessment. The model is expanded into a method to explore some of the potential effects the model can have on self-assessment. It is suggested that a digital scaffold with support for both teachers and students can be potentially vital for the proposed model and method. The model and the method reveals several areas that possibly can provide positive benefits for students' self-assessment skills. The potentially most important benefit revealed in this thesis is the increased number of feedback opportunities that the model can provide. Shorter sessions of peer-assessment take less time and can allow for a higher number of peer-assessment sessions. It also allows for a higher number of feedback opportunities within each peer-assessment session. Further study is required in order to determine how much impact, if any, shorter peer-assessment sessions can have on student gains in self-assessment skills. However, arguments based on the proposed model and existing literature presented in the thesis suggest that a higher number of feedback opportunities can have a significant effect.

## 2. Table of Contents

1.	Abstract.....	2
2.	Table of Contents.....	3
3.	Introduction .....	5
3.1.	Background.....	5
3.2.	Research Question and Goal .....	8
3.3.	What We Know.....	9
3.4.	Short Peer-Assessment in Norway at College Level.....	11
3.4.1.	Technology – a Scaffolded Approach .....	13
3.4.2.	Integration, Gamification and High Ability Versus Low Ability .....	16
4.	Method.....	18
4.1.	Design Method .....	19
4.2.	Literature .....	20
5.	Building a Model .....	22
5.1.	Digital Scaffold .....	24
5.2.	Setting Out.....	24
5.2.1.	Introduction .....	24
5.2.2.	Co-created Goals .....	25
5.3.	Task.....	26
5.3.1.	Open and Uncomplicated.....	27
5.4.	Feedback of Task .....	28
5.4.1.	Feedback.....	28
5.4.2.	Anonymous.....	28
5.5.	Feedback on Feedback .....	29
5.5.1.	Reflection.....	29
5.6.	Dialogue.....	30
5.7.	Culmination .....	31
6.	Expanding the Model into a Method.....	32
6.1.	Limitations .....	32
6.1.1.	Technological Limitations .....	33
6.2.	The Original Assignment Process .....	33
6.3.	Description of the Method .....	35
6.3.1.	Setting Out.....	36

6.3.2.	Task .....	38
6.3.3.	Feedback of Task .....	40
6.3.4.	Feedback on Feedback .....	40
6.3.5.	Dialogue .....	41
6.3.6.	Culmination .....	42
7.	Discussion.....	43
7.1.	Discussing and Revising the Model .....	43
7.1.1.	Setting Out.....	43
7.1.2.	Task .....	44
7.1.3.	Feedback of Task .....	45
7.1.4.	Feedback on Feedback .....	46
7.1.5.	Dialogue .....	47
7.1.6.	Culmination .....	48
7.1.7.	Not in the Model.....	49
8.	Conclusions .....	50
9.	Literature.....	53
10.	Figures and Tables.....	56

## 3. Introduction

### 3.1. Background

There is a global discussion, and in many cases, a shift, in terms of *what* and *how* students should learn (Schleicher, 2018; St.meld. nr. 28 (2015-2016)). We do not know what challenges today's first-grade students will face when they exit the school system, and what challenges students will encounter in a future that is rapidly changing. One of the skills required that seems to be important in the future is learning how to learn (Black & Wiliam, 2009; Schleicher, 2018; St.meld. nr. 28 (2015-2016)). Students will, on a larger scale than in decades passed, have to be self-dependent for learning necessary skills (Black & Wiliam, 1998, 2009; Lise Vikan & Trond, 2014; Schleicher, 2018). The Norwegian Ministry of Education and Research is rolling out the *Knowledge Promotion: Reform - Subjects – In-depth learning – Understanding*, which is a complete rework of the core curriculum and an updated view of how and what students should learn, in all Norwegian schools starting the fall of 2020 (St.meld. nr. 28 (2015-2016)). A substantial part of the Knowledge Promotion Reform has a focus on making students able to learn on their own (St.meld. nr. 28 (2015-2016)) so that when a student faces a challenge, she can assess the situation and know what steps she can take to overcome the challenge. If a student is to be responsible for her learning, she needs to be able to assess herself so that she knows if she has done an adequate job or if she needs to adjust some of the steps (Hattie & Timperley, 2007). Siemens states that “Our ability to learn what we need for tomorrow is more important than what we know today” (2005). This points to the fact that we do not currently know what problems might arise in the future and that skill of learning and being able to learn can help us overcome problems when they occur. The Covid-19 pandemic that swept the world during the final months of writing this thesis can be seen as an example of a situation that required more from us, the teacher's, ability to learn than a normal school year might have required.

Over 20 years ago, Black and Wiliam (1998) showed that formative assessment is a powerful tool for enhancing learning outcomes. Formative assessment takes many forms and is present in most classrooms every day (Black & Wiliam, 1998). A teacher walking around the classroom and giving individual feedback to students as they are completing a task, a teacher addressing the whole class after seeing a pattern of misunderstanding, a student using a calculator to check his calculations, or a student asking another student what they answered on a task are all examples of formative assessment in the classroom. The conventional sources of formative assessment are self-, peer-, expert(teacher)-, task-, and automated-assessment. The most important of these from a learning perspective is self-assessment (Black & Wiliam, 1998; Yurdabakan, Gelbal, Olkun, & Aydoğan, 2012). The other ones have benefits on their own, but are, however, necessary for developing self-assessment skills (Black & Wiliam, 1998), and to give the students confidence in the validity and reliability of their self-assessment (Nicol, 2014). Black and Wiliam (1998) argued for a theoretical connection between self-assessment and students' ability to self regulate learning. In a 2017 meta-analytic review, Panadero, Jonsson, and Botella (2017) found evidence that supported Black and Wiliam's argument that self-assessment interventions have a positive effect on

students' ability in self-regulated learning. *Peer-assessment* has been observed to have powerful effects on the quality and depth of students' self-assessments (Black & Wiliam, 1998; Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006; Reinholz, 2016). If using peer-assessment interventions is beneficial for increasing skills in self-assessment, it should, therefore, have a positive impact on students' skillsets that are needed in a changing future. A form of peer-assessment intervention is the main talking point of this thesis, but with a specialized focus, which will be elaborated on later. This thesis focuses on peer-assessment and some of the impacts it can have on self-assessment and does not contain an in-depth look at the ramifications for self-regulated learning by using peer-assessment interventions as this would likely require a different approach and methodology.

The Norwegian Directorate of Education and Research had a national project called Assessment for Learning (Vurdering for læring) from 2013 to 2018 with a focus on educating teachers in methods and theory of formative assessment (Norwegian Directorate for Education and Training, 2019). Formative assessment is a core element of both self and peer-assessment. This is elaborated on later in this thesis. The Norwegian Directorate of Education and Research presented a rapport (Norwegian Directorate for Education and Training, 2019) on the results of the project early in 2019. The rapport contained, among other reflections on the project, the results from a student survey. Only 20,4 % of students in 8<sup>th</sup> to 10<sup>th</sup> grade responded yes to the question, "Do you get to participate in assessing your own school work?" (my translation). This is progress from 2013, where only 17% percent of the students answered yes to the same question. The authors of the rapport argue that self-assessment is an area that has had improvements in Norwegian schools from 2013 to 2018, but is still in need of further improvement (Norwegian Directorate for Education and Training, 2019). The question from the survey is a subjective question that reflects the student's views and does not necessarily reflect how much students evaluate themselves. However, at a minimum, we can extrapolate that most students are likely not aware of situations at school were they evaluate themselves. I suggest that the low results stem from either that students are not aware of the situations were they assess themselves, or that students actually have few situations were they are asked to assess themselves. Both situations points towards a need for self-assessment to have a higher priority in Norwegian schools if they are to follow the aspirations of the afformentiond *Knowledge Promotion*.

A problem might be that a shift in pedagogical assessment strategies might lead to an up-front increase in work-load for the teachers, as was the experience of Haugan and Lysebo (2015), and without clear visible results, teachers might get discouraged from pursuing further activities that enhance student abilities for self-assessment (Black, Harrison, Lee, Marshall, & Wiliam, 2004). In this thesis, I propose a scaffolded approach to peer-assessment that tries to remove some of the extra work-load associated with peer-assessment. The scaffold also serves a second purpose as support for both teachers and students in the peer-assessment process. A web-based interactive system that serves as a scaffold is explained and proposed later in this thesis.

Nicol and Macfarlane-Dick (2006) argue in their article that an increased amount of peer-assessment activities is beneficial for gains in peer and self-assessment skills as students get

a broader view of other students' work and have a larger frame of reference for self-assessment. This collides somewhat with the more traditional use of peer-assessment, where students are primarily assessing written assignments of some length. I have not found a source that compares the length of peer-assessment assignments. However, examples of peer-assessment assignments that take at least several hours, and in some cases up to several weeks, to complete can be seen in the works of Burner (2016); Cho and Schunn (2007); Liang (2010); Lin, Liu, and Yuan (2001); Nicol (2014); Walker (2006). The traditional forms of peer-assessments might make it difficult for students to get a broad perspective of assignment quality as they are only reviewing a few assignments. In a Norwegian primary school setting, for example, students in eighth grade might have 3 \* 45 minutes of English classes each week. Writing an assignment might take a full week of class hours, and they might have to spend an entire week conducting peer-assessments of other students' assignments. In these two weeks, the students might only get one additional view other than their own on how to solve the problem. This method of conducting peer-assessments does not provide students with a wide variety of student work. This study will examine the possibilities of conducting shorter peer-assessments in an attempt to address the problem of few perspectives and low variety in peer-assessment assignments.

My personal motivations for writing an assignment about formative assessment and peer-assessment stems from my experiences as a teacher, and from learning more about formative assessments in my studies. Even though the teachers and administration at the school where I work has been a part of the Norwegian national project on formative assessment, Assessment for Learning (Norwegian Directorate for Education and Training, 2019), there seems to be a limited focus on formative assessment at the school. In my attempts at utilizing formative assessment strategies in the classroom, I have not found many digital tools that are of great support for formative assessment. By creatively using tools such as Microsoft Forms in combination with an elaborate Microsoft Excel-sheet with a bunch of automated processes, I have found some ways to support formative assessment with digital solutions. This is time-consuming and not beneficial for teachers or students on a larger scale. A notable exception, and one of the biggest inspirations for this thesis, is the Discussion tool that is a part of Campus Inkrement (Campus-Inkrement, 2020). The Discussion tool is an interactive classroom activity based upon students discussing in pairs and providing answers. The teacher can then display an overview of submitted answers, and conduct a classroom discussion revolving around the submitted answers. From my experience, this tool creates opportunities for learning that earlier would have been hard to imagine in a traditional classroom. I hope that this thesis can, in some way, provide insight into some potential avenues for formative assessment in classrooms.

### 3.2. Research Question and Goal

In the introduction, I argued that peer-assessment could be part of a process that leads to greater abilities in self-assessment and, as an effect of that, enhances vital skillsets that are relevant and important for students. However, students do not get a wide variety of material for reflection if they only view one other student's work. So, in an attempt at finding some possible solutions, here is my research question:

**In what way does short formative peer-assessment interventions, using a scaffolded approach, contribute to self-assessment for students?**

To answer this question, I propose a model for shorter peer-assessment later in this thesis. The model is heavily influenced by the research of Nicol and Macfarlane-Dick (2006), but as their research is on higher education and longer forms of peer-assessment, other data on shorter peer-assessment is required. I will further expand on this in this chapter, but from my experience of searching for literature for this thesis, there does not seem to be many studies on shorter peer-assessment in academic literature. I have detailed my search parameters in chapter four. As a substitute for the lack of literature, I am using a design-based method that can provide insight into the details of how shorter peer-assessments might look. I argue that this should provide enough saturation for a novel first perspective of short peer-assessment interventions. The model is inspected and theoretically tested by proposing a method for short peer-assessment in chapter six. In building the method, I found that using a technological component for conducting the peer-assessment interventions would suit the model. The technological part is used as a support or scaffold for both teachers and students. I see it as probable that the technological component will aid greatly in keeping the peer-assessment interventions short. This is discussed in more detail later in this thesis. Chapter four is dedicated to proposing the model for shorter peer-assessment based on scientific literature. Chapter five is dedicated to proposing a method based on the model for conducting shorter peer-assessment interventions.

There are different approaches to developing self-assessment-skills, and self-assessment is a complicated matter and can involve peers, technology, society, and many other factors (Black & Wiliam, 1998; Pachler, Mellar, Daly, Mor, & Wiliam, 2009). The approach in this study focuses on peer-assessment. Models, frameworks, and structures for feedback or peer-assessment have been made before (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006; Reinholz, 2016). These models do, however, not discuss or revolve around shorter versions of peer feedback. Likewise, they do not include in-depth discussions of technology as a resource for facilitating the peer-assessment process. A 2016 article (Tenório, Bittencourt, Isotani, Pedro, & Ospina) does contain a model for peer-assessment with the support of technology. As they state in the article, they try to solve the problem of having a large number of students enrolled in MOOC classes, and the article is focused on the reliability of grading and the motivation of students by using gamification. It is not a study with a focus on the potential benefits of using peer-assessment as a way of enhancing self-assessment or in using peer-assessment as a formative tool. Utilizing technology as a tool for increasing the effectiveness of peer-assessment is not discussed in any detail in the aforementioned articles.



This study focuses on formative assessment. Using peer-assessment for summative assessments in, for example, MOOCs that show that students are as reliable as teachers in grading assignments (Tenório et al., 2016) is interesting in many aspects, but students grading other students have shown in some cases to have a negative effect on the peer-assessment process (Nicol & Macfarlane-Dick, 2006). This study is aimed towards formative assessment and will not be including considerations to peer-assessment as a tool for grading or marking.

Two notes regarding terminology: The term scaffold that I have used in the research question and in the introduction of this thesis, is in this context a term for a support structure that holds or guides the process of peer-assessment. The term method is used in a traditional academic way of describing the methodology of this study in chapter three. It is also used as a term for describing my proposal for a peer-assessment tool in chapter five.

### **3.3. What We Know**

Formative assessment has been shown to have some of the most potent positive impacts on learning achievement (Black & Wiliam, 1998; Hattie & Timperley, 2007). Formative assessment is a focus on using assessment to achieve better learning outcomes. It stands in contrast to summative assessment that focuses on measuring learning outcomes. Formative assessment takes place during the process of learning, and summative assessment usually takes place after the learning process is finished (Black & Wiliam, 2009).

This study uses the definition of formative assessment set forth by Black and Wiliam (2009):

*Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited.*

There are several crucial elements in this definition worth mentioning. Formative assessment needs an initial set of instructions or a task set out and is not the first step in any process. This does not, however, mean that the initial instructions or task is not an important factor in the peer-assessment process. Nicol (2014) argues that students should work on tasks that are within the same topic domain as that of their fellow students and argues that it will help the students develop their own concept of quality if they are exposed to them.

Another element from the definition is that assessment or feedback needs to be used for it to be formative. Unused feedback is not forming the answer. It does not mean that the answer to a task has to change after feedback has been given, but that the answer might be better founded. This is one of the key elements for the proposed model in chapter three.

Assessing one-self is not an easy task. It takes practice and reflection to be able to do it properly (Black & Wiliam, 1998; Nicol & Macfarlane-Dick, 2006). There is not always a strong correlation between self-assessment and the teacher's assessment (Chang, Tseng, & Lou,

2012). If a student's self-assessment does not correspond with others' perceptions of their efforts, there will be implications for motivation and the student's view of their abilities (Nicol, 2014). One example of such missing correlation is when students evaluate their own work almost solely based on the time they spent on the task and not on the quality. If the student gets a low grade on her work, and she fails to understand why, there will not be growth in her self-assessment skills, and the student will probably make the same mistake again. It is not difficult to assume that repetitive failures of accurate self-assessment can lead to a drop in motivation for the student. Giving students a perspective of other students' work, and as Nicol (2014) argues, a perspective of high quality work will contribute to students' ability in self-assessment. This is one of the strengths of peer-assessment.

As mentioned earlier, peer-assessment can be a strong tool for enhancing student self-assessment skills. By reviewing other students' work, students get insight into several steps of the process (Nicol & Macfarlane-Dick, 2006). When a student reads another student's assignment, she gets a different perspective on how the task was solved. She already has written her own answer, and can now compare them both. She also has to formulate an assessment and give feedback about the task. This is a completely different approach to the task than before. The student providing feedback has to describe and formulate her feedback so that it makes sense to the first student. And then, the first student gets back the original assignment; the student gets feedback about her task from the perspective of a peer. All of these steps have shown to be beneficial for students' learning achievements and have a positive effect on self-assessment (Black & William, 1998; Hattie & Timperley, 2007; Heidi, 2019). Nicol (2014) argues that producing feedback for peers might be more beneficial than receiving feedback. This highlights the value of students being part of the assessment process.

Peer-assessment is often used on completed or nearly completed written assignments (Cho & Schunn, 2007; Lin et al., 2001; Nicol, 2014; Walker, 2006). If an assessment is too complex, it might not yield any beneficial learning outcomes for students (Hattie & Timperley, 2007). Several studies suggest that extensive training is needed for students in how to assess others (Nicol, 2014; Walker, 2006) and that there are many possible pitfalls during the process of peer-assessments. The students need to have clear and understandable criteria to be able to reliably and confidently assess each other. If there is a threat to a student's self-esteem or if they feel unsafe and do not trust their peers, the learning outcome might decrease significantly (Hattie & Timperley, 2007).

Peer-assessment has in multiple studies shown to be beneficial for learning outcomes. The studies suggest that learners get to benefit from seeing other learners' mistakes and useful solutions (Nicol & Macfarlane-Dick, 2006). They also get a deeper understanding of what is considered good or bad solutions to the problem at hand. The studies also suggest that students get a deeper understanding of the assessment criteria (Hovardas, Tsivitanidou, & Xacharia, 2014).

The results from peer- and self-assessment have shown varying degrees of reliability and validity in comparison to expert assessment (Hovardas et al., 2014). Even when students had previous experience with and training in using the assessment criteria, they failed to

guarantee reliable or valid quantitative results. An argument can be made that differing feedback might confuse the student that receives the feedback. However, students can use this as a basis for a learning strategy similar to the triangulation method researchers use for checking validity (Hovardas et al., 2014). It has been found that students have used this method to evaluate the responses and disregarded expert feedback from the teacher in favor of peer feedback (Hovardas et al., 2014). This can encourage a reflective process where the student has to assess the feedback and make choices on how to proceed.

Different variations of formative assessment or feedback have varying degrees of effectiveness. Some even have a negative effect on student performance (Hattie & Timperley, 2007). As an example pointed out by Hattie and Timperley, praising students usually does not have a significant positive impact on a student's performance. It is therefore important for teachers to have competency on the most effective ways of giving feedback, and what strategy to apply in different situations. A tool that is inscribed for peer-assessment, and that aims at increasing students' skills in peer-assessment as they work on different school subjects can potentially support teachers' executions of peer-assessment in the classroom. An inscribed tool not only has *an option* for peer-assessment but is built with a goal of supporting both the teacher and the students with the peer-assessment process. As with many other things, regular training, and high exposure to many instances might increase the learning outcome for students (Nicol, 2014). So a tool that focuses on shorter bursts of peer-assessment might be preferred. An IT-tool can alleviate and assist peer-assessment (Nicol, 2014), and has also shown to have a positive effect on peer-assessment (Hattie & Timperley, 2007). The method proposed in chapter six will, therefore, be based on a technology-based solution. An important argument that is worth noting from a report by Pachler et al. (2009) is that no technological or digital solution is inherently formative or that it will automatically create formative situations. It has to be connected with a social process involving students and teachers. It is the actions of the participants that, in the end, determine if the teaching sessions are successful formative assessment situations.

Black and Wiliam (1998) point to several studies highlighting the need for consideration of group compositions in peer-assessment assignments. A mixed body of students seems to solve open tasks better, and low achievers benefit from a mixed group of students as long as the low achievers' contributions are not drowned out. There is not a complete absence of learning gains if the composition is sub-optimal. But for peer-assessment purposes where the discussion should be open (Nicol, 2014), a mixed group of students might be a good starting point. An IT-based solution could be beneficial in composing groups based on earlier results.

### **3.4. Short Peer-Assessment in Norway at College Level**

This is a look into a real-world scenario where short peer-assessments were the primary formative assessment method in a handful of courses at Høgskolen i Oslo og Akershus, now Oslomet. It provides valuable insight into the process and challenges that occurred. They removed all of the obligatory class assignments and replaced them with voluntary peer-assessments. Haugan and Lysebo wrote two articles (Haugan & Lysebo, 2015, 2018)

describing the background for this decision, a description of the execution, and provide reflections of the completed process. These two articles discuss peer-assessment from a Norwegian perspective, and they revolve around fast and short peer-assessments situations.

One goal of the shift was to see if the teachers' work-load could be lowered to free up the teachers' time to give the students better and more thought out assignments. Another goal was to stimulate students to work consistently throughout the year. They also argue the shift with arguments of extensive literature pointing to the benefits of timely formative assessments. They also point to two other Norwegian examples of using peer-assessment, and in both of those, the percentage of failing students dropped dramatically. They conclude that it is not necessary with traditional obligatory assessment and that peer-assessment worked (in this one case) for getting the students to work at a consistent pace throughout the year.

They do, however, find that the students are not one-sidedly pleased with how the peer-assessment process functioned, and Haugan and Lysebo suggest that peer-assessment has not found its form yet. They also point out that peer-assessment is not a quick fix for saving resources. Even though they did not assess any student assignments, they found it time-consuming to manage the process and make assignments suitable for peer-assessment (Haugan & Lysebo, 2015).

In the latter of their two articles, Haugan and Lysebo (2018) present grade results from before and after the shift from obligatory assessments to peer-assessments. The number of failed students dropped by five percentage points. Based on a pre-test done at the start of the semester, it seemed like the student with some of the worst results benefited the most from the peer-assessment structure. Haugan and Lysebo did find that students with the highest number of study hours out-performed those with fewer study hours, and also suggest that peer-assessment might have motivated them to work more.

The articles did not discuss if the higher number of weekly work hours for the students were the main reason for low achieving students getting the most benefit from the peer-assessment structure. But in follow-up e-mail correspondence, both authors highlight other factors than increased work hours as the most plausible reason. Haugan argues that a more flexible course program with more flexibility is beneficial to a diverse group of students and that incremental victories are important for continued motivation. Haugan points out that peer-assessment makes the students articulate and debate the subject at hand, and that this leads to a better understanding and might organize each student's thoughts in a manner that is less present with traditional assignments. Haugan also argues that the removal of obligatory student hand-ins that is required before a student is allowed to participate in the final exam had a significant impact on the result. This, argues Haugan, allows the student to come forward with something they do not understand and get help to solve it instead of sitting alone and not getting the feedback that he needs. These are perspectives that the two authors elaborate about via e-mail, and are not necessarily scientifically tested. Haugan and Lysebo do, however, have over five years of experience with shorter peer-assessment.

The students in Haugan and Lysebos' articles complete the peer-assessment in a classroom situation. The students that want peer-assessments sit together when they assess each other's work. When the authors are asked via e-mail if they have used any digital solutions for the peer-assessment sessions, both Haugan and Lysebo say that they have not. Haugan points to the difficulty of typing mathematical calculations quickly and effectively on digital platforms, but also that they have focused on quick feedback and that a non-digital classroom situation is preferable to facilitate quick feedback. As the authors point out in their articles, timely feedback was one of their goals, citing the work of Chickering and Gamson from 1987 and Gibbs and Simpson from 2002 as arguments for this goal. Both are indicating that timely feedback is more effective than delayed feedback.

### 3.4.1. Technology – a Scaffolded Approach

A project report for JISC (Pachler et al., 2009), highlighted a set of areas that were uncovered after a literature review, where digital solutions could aid formative assessment. The areas are speed, storage capacity, processing, communication, and construction and representation. Following is a brief explanation of some of the mentioned areas that are most relevant for this thesis.

Rapid iteration and the speed of response enables students to get feedback quicker, but also gives the opportunity to respond and have more frequent iterations of their work (Pachler et al., 2009).

Technology-supported communication allows ideas to flow between a range of audiences, and technology can aid in limiting the flow or the audience so that participants are not overwhelmed with information (Pachler et al., 2009). It can also aid in retaining communication and provide students the ability to revisit earlier communication to focus on areas of improvement (Pachler et al., 2009).

Technology can enable a representation of ideas in differing ways so that students can compare and translate between these representations. Shared objects can then be changed and reviewed easily (Pachler et al., 2009).

In the same report, the authors present what they call patterns that are key in the context of analyzing and planning for formative assessment. The authors do state that they “offer them here in no particular order as possible focal points for discussion in the design of instructional episodes and/or the development of future e-assessment tools but wish to stress that they are neither systematic nor have they necessarily been empirically tested and verified” (Pachler et al., 2009). So the patterns that are listed below (Pachler et al., 2009) are only areas that might be worthy of exploration when developing e-assessment tools.

<i>Classroom Display</i>	<i>Share learners' work with a trusted audience. Create a space within the learning environment where learners' works can be displayed side by side.</i>
<i>Feedback on Feedback</i>	<i>Feedback given to learners should provide opportunities to improve the learning experience. It should comprise</i>

	<i>constructive feedback to improve learning as well as socio-emotive feedback. Tutors in large courses often resort to grading devoid of effective feedback. To support them in improving their feedback, they need effective feedback on the feedback they give.</i>
<i>Narrative Spaces</i>	<i>Constructing narrative is a fundamental mechanism for making sense of events and observations. To leverage it, we must give learners opportunities to express themselves in narrative form.</i>
<i>Objects to Talk With</i>	<i>When we talk we point at objects. When we talk on-line we should be able to do so too. When providing tools for learners to discuss their experience, either as part of the activity or at a reflective meta-level, allow them to easily include these artefacts in the scope of their discussion.</i>
<i>Round and Deep</i>	<i>Use the students' experiences to complement your own and provide the alternative perspectives required.</i>
<i>Showcase learning</i>	<i>Publicly celebrate student work.</i>
<i>Soft Scaffolding</i>	<i>Scaffolding is a term commonly used in educational design to describe structure that directs the learner's experience along an effective path of learning. Technology should be designed to scaffold learners' progress, but an interface that is too rigid impedes individual expression, exploration and innovation.</i>
<i>Try Once, Refine Once</i>	<i>A two-step question-answering system which encourages students to consider their initial answers to skills-based questions very carefully, and, on receiving feedback on their errors, to give as much thought to the refinement process.</i>
<i>Wear Your Skills on Your Shirt</i>	<i>Use virtual appearance to reflectabilities. The visual representation of your avatar shows the extent of your skills. Skills can be gained or given, and be personal skills or avatar skill.</i>
<i>Use My Stuff</i>	<i>Use learner supplied artefacts as raw materials for new learning activities.</i>

*Table 1 - Patterns for e-assessment (Pachler et al., 2009)*

Some of these patterns are revisited and discussed later in this thesis.

Wasson and Hansen (2014) interviewed six teachers from three schools to determine what types of ICT were used by Norwegian teachers to facilitate for 21<sup>st</sup>-century teaching. Even though this study is from six years ago, they selected schools that had prioritized and focused on ICT. Therefore, their findings should still be relevant as other schools would have had to catch up with the schools in the study. From the interviews, they found few tools for peer-assessment in use at the schools, and only wikis were specifically mentioned as a tool for peer-assessment. The teachers argued that "ICT, as it is now, is not the best choice for enabling dialog," and that "they did not have the skills required to use the tools." Tools like ItsLearning and Canvas do provide tools for peer-assessment in 2020 based on the

information on their respective websites, but the extent of use is uncertain. And if such tools are widely used, it is likely to be a new development in Norwegian schools.

Wasson and Hansen found in their interview of Norwegian teachers that even though Norwegian schools seemed to take more advantage of formative assessment than other European countries' school systems, they were still too closely connected with summative assessment.

*... formative assessment was still closely tied to summative models of grading and assessment criteria and that more needs to be done to develop formative assessment practices that are more qualitative, student-centred, and which enable the rich detail of student learning to be visualized and used effectively (Wasson & Hansen, 2014).*

Wasson and Hansen also found that especially foreign language teachers found it challenging to incorporate formative assessment. The teachers claimed that the number of aspects involved with learning language made it difficult (Wasson & Hansen, 2014).

Wasson and Hansen described Norwegian teachers' use of technology as "stretching the pedagogical advantages of technology and integrate it into a pedagogical setting" (Wasson & Hansen, 2014). Many of the existing ICT tools are not inscribed for peer-assessment but are translated and used as such. Wasson and Hansen (2014) highlight wikis as an example of this. The teachers use the wikis activity log for seeing student activity, and the editing capabilities also make it possible for peers to give feedback. The teachers did, however, talk about the difficulties of sorting through all the information and found it time-consuming. Wasson and Hansen proposes that cloud-based and Web 2.0 applications could be the key for some possible solutions to lower the teacher burden and innovating the classrooms.

Seufert, Guggemos, and Sonderegger (2019) used ItsLearning for peer-assessment intervention. According to them, ItsLearning has functions for peer-assessment assignments. They did, however, find that the students might have a lack of acceptance of the process. They also argue that the competency development of the students and teachers is a key requirement for conducting successful peer-assessment interventions. I, the author of this thesis, have looked at the peer-assessment functionality of ItsLearning, and it is predominantly a tool for solving the administrative issues of assigning a submitted task to another student. It takes student assignments and delivers them anonymously to other students for them to assess. There is not any further scaffold or support for how to best execute a peer-assessment intervention other than an option for adding goals and criteria as a part of the initial instructions for the task. There are, of course, a lot of other LMSs, and this is not an in-depth look at the peer-assessment functionality of all of these. As highlighted earlier, Wasson and Hansen (2014) mapped ICT use in Norwegian schools through interviews with teachers. They found that the technology that teachers used was not always made for pedagogical use, but that teachers formed the technology in some ways to fit a pedagogical purpose. This fits with my own experience with two popular LMSs; Google Classroom and Microsoft Teams. Evidence of this can be seen with a Google-search of *Peer-assessment in Microsoft Teams* or *Peer-assessment in Google Classroom*, where the search results show various ways of almost forcing the LMSs to allow peer-assessment. As a

result, the teachers who choose to use peer-assessment interventions do not always get the necessary support from the technological systems that they are using. This might prohibit more teachers from fully utilizing peer-assessment in classrooms as they are not confident enough with the method or do not see enough of a potentially positive effect against the amount of work they are required to do.

In presenting the learning theory of connectivism, Siemens (2005) argues that the result of learning can not be understood as an ability to recite information from memory and can not be restricted to a single person's sole abilities. He proposed that obtained learning can be viewed as nodes and the connections between them. As we learn, we strengthen the nodes and the connections between them. The nodes can be a persons' own memorized information, but it can also be peers, experts, the internet, an app, a process, and much more. The connections between these nodes can be weak, or they can be healthy. If you need to learn how to use a power drill, it does not help you that your neighbor is an expert if you do not know that she can help you. Siemens (2005) claims that "Nurturing and maintaining connections is needed to facilitate continual learning." Evaluating and assessing the information from the nodes is also very important. Therefore, the nodes are not only linked with a connection to ourselves, but they are connected across each other. Plentiful and robust connections are required to have confidence in the result (Siemens, 2005). As noted by Siemens (2005), learning (or knowledge) may reside in non-human appliances. A teacher with a basic understanding of formative assessment with a digital tool that is inscribed with solutions to support formative assessment might be as competent in formative assessment as an expert. The expert has the nodes regarding formative assessment in his head. The teacher has a strong connection to a useful tool, and can, therefore, benefit from the same knowledge as the expert. To fully see the potential for peer-assessment in classrooms, I argue that the technological solution used for carrying out peer-assessment interventions needs to be a major part of the equation. The technological tools teachers use are vital nodes that are part of a teachers combined competency. This is the main reason for using a substantial part of this thesis on proposing a digital method for conducting peer-assessment interventions.

### **3.4.2. Integration, Gamification and High Ability Versus Low Ability**

A Taiwanese study called "From integrative to game-based integrative peer response: high ability versus low ability" (Wang, Chen, Chang, & Chan, 2015) proposes two peer response approaches to address some of the problems associated with traditional peer response.

This study matches the search criteria for both short and digital peer-assessment and gives valuable insight into a digital solution for shorter peer-assessment interventions.

They call the approaches *integrative peer response* and *game-based integrative peer response*. They attempt to combine the benefits from face-to-face mode and computer-mediated communication mode so that these modes complement each other. They present a summary of the advantages and disadvantages.



The article does not talk in detail about the implementation of the *integrative peer response*, but it is a forum-based system that tries to alleviate some of the disadvantages of both face-to-face mode and computer-mediated communication mode.

The *game-integrative peer response* is described in a bit more detail. It functions as a role-playing game where the student plays as the head of a publishing company. As part of this game, there was a review department where the tasks focused on peer-assessment. They used a structured checklist to facilitate readers to deliver comments in a useful manner. The comments were again assessed by the writers. The process is described as «learning by doing» by the authors. The *game-integrative peer response* also provided game elements to enhance the students' engagement. Examples are virtual currency, leader boards, and trophies.

The study had 75 student participants, all in third grade. They were divided into three groups. One did not use a peer response system, one used *integrative peer response*, and the last used *game-integrative peer response*. The participants all took a pre and post-test to give insight into each student's ability.

The study revealed some differences between the three student groups. The *integrative peer response*, a non-game context, was useful to improve written expression, while the *game-integrative peer response*, the game context, was helpful to enhance writing quality. As the *game-integrative peer response* used a structured list of guidelines, the assessors had better knowledge of issues to address than the *integrative peer response*, and this might have led to a higher focus on these elements. In the forum-based *integrative peer response*, the students had a more open structure, and this resulted in free discussions and generated interesting ideas to improve the content of their peers' work. The positive impact of the written expression in the *integrative peer response* only appeared for the high achieving students.

Both high achieving students and low achieving students performed significantly better in the *game-integrative peer response* than the students using *integrative peer response*. The low achieving students with the *game-integrative peer response* had significantly better results than the low achieving with *integrative peer response*. On the other hand, the high achieving students had better results with the *integrative peer response* than the high achieving students with the *game-integrative peer response* in the aspects of directive feedback on content features and the length of composition. The authors conclude that the *game-integrative peer response* was more helpful for the low achieving students, and the *integrative peer response* was more helpful for the high achieving students. The authors argue that game elements may have various impacts on motivation and that the motivation changes based on the achievement level of each student, and they conclude that student's abilities should be considered when game-elements are incorporated into peer response.

The authors argue that their results can be generalized for most game-based learning system involving peer-assessment, as the game-based elements are the most commonly used, and that other studies have shown similar results. The model proposed later in this thesis does

not include much of traditional game elements, but the set-up for the study of Wang et al. (2015) has enough similarities for it to be worthy of consideration.

## 4. Method

Peer-assessment has, as discussed in the previous chapter, had some traction in different parts of education and has been shown to provide beneficial learning gains for students. There might, however, be other versions of peer-assessment that might provide more or different learning gains for students. The main goal of this thesis is to find out in what ways shorter peer-assessment interventions, using a scaffolded approach, contribute to self-assessment for students. By proposing a new model and an accompanying method that is built on a framework of existing literature, a new perspective and possibly new avenues for peer-assessment might emerge and provide valuable insight into further development of peer-assessment interventions. It is important to state that the purpose of this study is not to create the ultimate peer-assessment model or method. It aims at exploring an avenue of peer-assessment and looks at what benefits this particular avenue might have in regards to self-assessment.

Discussing and trying to answer in what ways it contributes will be in focus, and not how much it works. The answer to what ways short peer-assessments contribute is not fully explained in the literature that I have found on the subject. There are models and guidelines for general peer-assessment (Nicol, 2014), but they are not specific enough to adequately provide enough data to answer the research question. Building and proposing a model for a specialized version of peer-assessment is something that Nicol and Macfarlane-Dick (2006) suggest as possible paths for further research.

In a critical review of research on student self-assessment, Heidi (2019) proposes that future studies should try to answer how to best scaffold effective formative self-assessment. Proposing a model that uses peer-assessment to enhance self-assessment is an avenue that might yield a better understanding of how to scaffold effective formative self-assessment.

One of the weaknesses of proposing a theoretical model is that there is no proof of it being viable in a real-world context (March & Smith, 1995). To alleviate this, I propose a method that is built for conducting a specific assignment in a specific school subject at a specific school level. In creating a method with a narrow area of use, it should provide some perspective into the viability of the model. The method is then used as a base for discussing the model. This provides a perspective on the feasibility and probability of real-world use to the model (March & Smith, 1995). The technological limitations of the method is also discussed in some detail with the same purpose; giving a perspective into the viability of the model. As I have argued earlier, a technological scaffold can add value to a peer-assessment process by expanding a teacher's competency and expertise. Proposing such a digital scaffold, and discussing the different elements of it, yields a new perspective to the peer-assessment process. The scaffold is also vital for the method as it provides opportunities for conducting peer-assessment interventions with a better control of the communication involved and provides more opportunities for automation and effectiveness, as argued by Pachler et al. (2009).

## 4.1. Design Method

March and Smith (1995) propose a framework for doing design-based IT research. The framework consists of four research activities and four research outputs. Combining each research activity with each research output, they create 16 individual focus areas that, on its own or in combination with the others, can be the basis for scientific research.

March and Smith (1995) describe four types of design science. They define constructs as concepts that form the vocabulary of a domain. Model is defined as a set of propositions or statements expressing relationships among constructs. They (March & Smith, 1995) describe method as a set of steps used to perform a task. Methods are more detailed and specialized than the model and are used in translating from a model to a more specific situation. Instantiations, they define as the realization of an artifact in its environment. I use these terms throughout the thesis. The exception is the title of this chapter, which refers to the research method of this thesis.

According to March and Smith (1995), an IT system may be instantiated based on intuition, necessity, or experience, and does not need to have clearly defined concepts of constructs, models, or methods for it to have an instantiation. Even when such an IT-system works, it does not provide much insight into how or why it works. An IT-system can be measured on *how much* it works without considering the constructs, model, or method it might use, but it will be hard to determine why or how without these research outputs. This study will not propose or conduct an instantiation, and will therefore not be able to answer how much effect one can expect from a short peer-assessment intervention, but by examining the constructs, the model and the method, it might provide insight into why it can work, and in what way it can work.

March and Smith (1995) list the four different Research activities: build, evaluate, theorize, and justify. They define *build* as the construction of an artifact. March and Smith (1995) argue that «building the *first* of virtually any set of constructs, model, method, or instantiation is deemed to be research, provided the artifact has utility for an important task.» They also argue that performance evaluation is not required at the building stage, and would best be suited for a later revision of the model.

*Evaluate* refers to the development and assessment of criteria and how the artifact performs in relation to those criteria.

Novel instantiations have, according to March and Smith (1995), had less enthusiasm in IT-literature compared to novel constructs, models or methods. A novel instantiation that builds on unchanged constructs, models, and methods might not provide much value for research. March and Smith (1995) do, however, point out that constructs, models, and methods might work «on paper» but fail in real-world contexts. Since the process in this thesis does not include an attempt at instantiation, concerns regarding a real-world scenario should be discussed, and that is what is done on chapter 5 where I propose a method based on a real-world scenario. This will alleviate some, but not all, of the weaknesses of not having a real-world instantiation with concrete proof of functionality. Building a method might also provide insight into potential benefits that the model might yield.

An IT-tool usually never ends development (March & Smith, 1995). It is a constant process to keep evaluating and developing new solutions for an everchanging world. This is where instantiation and evaluation can give the best answers as to how the artifact should evolve and change. So, it is important to note that the model and method that is proposed in this thesis should not be considered final versions. The model and method might function differently in different circumstances and yield differing results based on the parameters of its context. Another consideration is that our society is in constant change, and as our world changes, so must our tools. It is also highly likely that I fail to take into consideration all aspects that would make the optimal peer-assessment model.

Theorizing in the framework of March and Smith (1995) refers to analyzing and developing the characteristics of the artifact that is being studied and the environment it is in. As mentioned earlier, this study aims to propose a new model for peer-assessment. Some of the constructs for the model will be theorized from existing constructs and models in literature regarding formative assessment and peer-assessment. There is also a need for new constructs to address shorter forms of peer-assessment, and these will come from a look at relevant literature formative assessment, feedback, and short peer-assessment. The results from building the method are the basis for further discussion of the model and justifying it.

## 4.2. Literature

Peer-assessment, in general, is thoroughly researched (Li, Xiong, Hunter, Guo, & Tywoniw, 2020; Nicol & Macfarlane-Dick, 2006; Reinholz, 2016). However, the “short” part of the main question in this thesis is essential. The volume of research papers regarding short peer-assessment is less, but not nonexistent. I have searched for relevant literature using Idun, Ebsco Host, and Google Scholar. I did not find a common set of terminology discussing short peer-assessment, and in most cases, I was unable to determine the relevance of the article based on the title or abstract. I included literature that, in some form, had shorter, less than thirty minutes, versions of peer-assessment. I excluded literature that focused on peer-assessment as a system for grading.

In searching the Norwegian database Idun, I used the terms *medstudentvurdering*, *hverandrevurdering*, *kvarandrevurdering*, and *peer-assessment* with the latter yielding the most results with 16 results. I specifically searched through Idun to find articles that had a Norwegian perspective as the method is aimed at Norwegian students that has English as a second language. I am proposing a method aimed at the Norwegian school system, and a Norwegian perspective might illuminate specific cultural or geographical considerations to take. After reading the articles, four of them were selected for further consideration.

I used two search-terms on Ebsco Host that gave some results, only including peer-reviewed literature.

- 1- Short AND peer-assessment
- 2- peer-assessment AND educational games

Number one gave fifteen hits with one relevant article, and number two gave twenty-one hits with three relevant articles that are used to a greater or lesser extent in this thesis. I included the term *educational games* as a parameter because educational games, from my experience, often have a quicker pace to them, and game-based peer-assessment solutions have a technological aspect to them, which might make them relevant for my study.

I further used the references-lists of these articles to find other relevant literature. Four of these are used in this article, including the article written by Nicol and Macfarlane-Dick (2006), which is the most influential source for the model that is proposed in chapter four.

## 5. Building a Model

Research done by Hattie and Timperley (2007) and Black and Wiliam (1998) is widely read and used in assessment and feedback research (Lise Vikan & Trond, 2014; Wisniewski, Zierer, Hattie, & Wisniewski, 2020). It would probably be a mistake to not use their research papers as a basis for further research in the field. David Nicol's work (2014) on peer review has 4915 citations on Google Scholar, and it seems to be one of the more comprehensive works on peer review. It is, however, based on college-level students and long-form peer reviews, so it is not certain that all of the arguments have the same impact on younger students or shorter forms of peer-assessment. The following is a set of constructs that have been built with a heavy influence from Nicol's article (2014) but also an influence from Hattie and Timperley (2007), Black and Wiliam (1998), and the literature that I found while searching using Idun, Google Scholar and Ebsco Host.

Nicol (2014) uses the term "review" instead of "assessment." In his article, he refers to peer-assessment as the work of grading other students' work. In this thesis, both terms are used to describe the same thing: formative peer-assessments. Another note regarding Nicols article that I refer to in this chapter is that it is a book chapter mostly based on an earlier scientific article that Nicol wrote together with MacFarlane-Dick (2006). I refer to the book chapter as it is of a newer date, and Nicol has made some updates to the text.

Table 2, below, is the model I propose for shorter peer-assessment interventions. The following subchapters are elaborations, clarifications, and literature-based discussions of the constructs that make up the model.

The model is not connected to a specific school subject, theme, or topic. It should be relevant for a wide variety of different school curriculum topics. In chapter five, the model is used to propose a method for a peer-assessment intervention for a specific school subject at a secondary school level.

The model is built chronologically, which means that it follows the process from start to finish. *Setting Out* contains the elements that should be considered before students start working on the initial task, and *Culmination* is the end or completion of the task. There is an explanation of the process of the model directly below Table 2. The individual elements are explored further in this chapter, and a more direct and practical interpretation of the model can be found in chapter five.

	Setting Out	Task	Feedback of Task	Feedback on Feedback	Dialogue	Culmination
Activity	Introduction Co-created-goals	Open and uncomplicated	Feedback Anonymous	Perspective	Dialogue	Self-reflection
<b>Digital Scaffold</b>						

Table 2 - Proposed model

In the *Setting Out* stage, the students are introduced to the study material, the learning goals, and the process of the upcoming peer-assessment process. Together with the teacher, they formulate learning goals. The *Digital Scaffold* reminds the teacher of these steps and can display the learning material. The teacher can put in the co-created learning goals as well as an introduction to the peer-assessment process. The *Digital Scaffold* is a technological tool that both students and teacher are logged in to. In chapter five, the proposed scaffold is a web-based system, but it could also be an app or take another form altogether.

In the *Task* stage, the students complete a task set out by the teacher related to the study material and hands it in. The role of the scaffold here is to provide appropriate solutions for handing in the task. A text-box for a written answer, or options for multimedia if the task requires it.

In the *Feedback of Task* stage, each student gets an answer that another student has written. They do not get any information about who might have written the answer. They are prompted with a task to write feedback on the answer. The scaffold handles the administration of assigning feedback tasks.

In the *Feedback on Feedback* stage, each student gets an answer to the initial task accompanied by the feedback that was written in the previous task. They are now prompted to give feedback on the feedback and answer what about the feedback they believe to be correct or false and give an explanation of their reasoning.

When all students have completed the two feedback stages, the class moves on to the *Dialogue* stage. In front of the whole class, the teacher goes through the submitted answers and the corresponding feedback. The teacher has the option of discussing each element and can involve the whole class in the process. The scaffold provides an overview of each set of tasks and feedback.

In the final stage, *Culmination*, each student gets the opportunity to modify their original answer. They are also asked questions about their reasoning where they need to formulate a response on why they did the changes they did.

The teacher has a live overview of student progress and can monitor each student or the class as a whole throughout the process.

## **5.1. Digital Scaffold**

The digital scaffold is something that is present throughout the whole model. It provides support to both the teachers and students. A digital addition to formative assessment can aid the process by speeding up the process (Pachler et al., 2009). Speeding up the process is important with regards to this model, as it aims to provide a way for students to conduct peer-assessment interventions within a short timeframe. Another benefit that can be present with a digital addition to formative assessment is better control over the flow of communication (Pachler et al., 2009). A better controlled flow of communication is key for the proposed model to function, and so is the benefits of automating the flow of communication. A digital addition to formative assessment can also be beneficial by being able to mutate the artefacts that the students are working with and keeping track of the mutations throughout the process (Pachler et al., 2009).

## **5.2. Setting Out**

### **5.2.1. Introduction**

As seen in the results in the report from Norwegian Directorate for Education and Training (2019) regarding formative assessment, not all students are familiar with different types of formative assessment. If students are not familiar with giving and receiving feedback from fellow students, it can be an unsure and potentially upsetting setting to be asked to participate in for the students (Nicol, 2014). Nicol (2014) suggests six approaches in how to ensure an atmosphere of trust and respect in peer-assessment situations, as can be seen in Excerpt 1 below.

- 1. explaining why peer review is being implemented and what students will get out of it;*
- 2. illustrating how peer review operates in professional contexts and in life beyond university;*
- 3. clarifying that reviewing is not about finding fault with, and undermining the work of, others, also showing students examples of both constructive and less constructive feedback critiques;*
- 4. emphasising that students will still learn even if they receive poor reviews, as it is the reviewing experience itself that matters;*
- 5. dealing with concerns about copying by emphasizing that learning, at its best, is a collaborative endeavor or by explaining how you have designed the activities so that plagiarism is not an issue; and*
- 6. making it clear that you are not asking students to mark others' work.*

*Excerpt 1 - Six points of encouraging an atmosphere of trust and respect (Nicol, 2014)*



Instruction is, in this model, the learning material that students go through before they start working on the task. The introduction can take any number of forms. It can be a teacher talking about the subject, watching a video about the subject, reading about it, or several other methods of learning about a new subject.

Hattie and Timperley (2007) state that if a student is not able to reflect on the study material, there will not be any benefit from a feedback process (Hattie & Timperley, 2007). If the reason for lack of reflection is unfamiliarity with the subject, there will be a greater effect on giving instructions than in providing feedback, according to Hattie and Timperley (2007) and Nicol (2014). The model can potentially be used for incorporating short peer-assessment interventions in almost any school subject, and at almost any school level. The way the teacher introduces the students to the learning material would, therefore, depend on the specific school subject, the student group, school level, and other factors involved with classroom learning. What shape or form the instructions takes is not of great importance for this model. The important part is that *there is* an introduction and that it gives the students enough background so that they are able to envision some form of coherent response to the following task.

### **5.2.2. Co-created Goals**

Hattie and Timperley (2007) introduce three questions in their model of feedback. The first one is “Where am I going?” and refers to the goals for the task. Feedback, according to Hattie and Timperley (2007), works by “reducing the discrepancy between current and desired understanding.” If the desired understanding (the goals) is clear, students are more likely to increase effort in the task. The teacher and the students should strive to have a common understanding of the overarching learning goals for the subject as a part of the introduction.

The creation process of goals is one of two core activities highlighted in formative assessment by Black and Wiliam (Black & Wiliam, 1998). Students' involvement in the creation of the goals and assessment criteria have shown great effect as it gives students a higher understanding of what is to be expected to be accomplished and how (Nicol, 2014). Teachers might think that students share their enthusiasm for academic achievement, but they often do not, according to Nicol (2014). By involving the students in the process, they are more likely to get motivated to achieve the goals (Black & Wiliam, 1998). Focusing on process goals instead of task goals has shown to give better results (Black & Wiliam, 1998). The goals for the peer-assessment intervention should, therefore, be made clear, or preferably be created by the students themselves. The goals should include both goals for the specific class topic, but it should also contain goals regarding peer and self-assessment as those are vital goals for these types of exercises (Nicol, 2014).

Having clear and precise goals has been shown to be effective for formative assessment (Black & Wiliam, 1998). Nicol (2014), however, argues that even with clear and precise criteria set out by the teacher, students still reflect on their work along with their own criteria alongside with the teacher set criteria. The students' work becomes a reference point and serves as a form of criteria in assessing other students' work. Further, Nicol (2014)

argues that even experts use their internal tacit criteria in combination with the pre-set criteria when they are assessing someone's work. Throughout the process, features of the work might become more prevalent and give more weight than the initial criteria set out that it would have. By providing the students with criteria, the students might grow and develop their own criteria, and the students' criteria might be sounder than if the students were not provided with criteria (Nicol, 2014). If you eliminate or reduce the criteria set out by the teacher, there is evidence of students developing a stronger ability for holistic judgments (Nicol, 2014). Nicol (2014) stresses that for this to be the case; students need to review a larger quantity of peer texts, and that some of those need to be of high quality. By conducting shorter peer-assessment interventions, it allows for a greater quantity of interventions. If a teacher has regular peer-assessment interventions, it makes it possible first to have clear criteria so that students learn how to use them, and then, in time, remove the criteria so that the students can better develop their own holistic judgments. To achieve some of the same effect, Nicol (2014) argues that students could assess a high-quality answer and discuss the criteria in class. I will return to this argument in the *Dialog* section of the model.

The study of Wang et al. (2015) suggested something similar where students that were given criteria showed student improvements in writing quality. And assignments without criteria showed student improvements in written expression. This can suggest that deeper, more reflective outcomes can occur when criteria are less expressed. However, the study also showed that weaker students had better outcomes from more direct criteria. In my model, I, therefore, suggest that criteria always should be present in the first peer-assessment interventions that students partake in, but that they should be removed after the students are comfortable with different kinds of criteria and have a better vantage point for creating internal criteria when they are presented with a peer-assessment assignment.

### **5.3. Task**

The type and form of the task depend on the subject the class is studying. As an example, a language-based task would often be a written task, and a mathematical task could be a handwritten calculation. The specter of available task options is too big to be discussed here, but, as a part of the model, I suggest a set of guidelines that can be beneficial for tasks in shorter peer-assessment interventions.

Each student gets their individual task. The task can be, but does not have to be, precisely the same as the other students. Nicol (2014) argues that it should be similar enough so that the student has a basis for comparison. Having slightly different tasks provides the students a basis for discussion later as they are familiar with the topic and have worked on solving it themselves. If the tasks are identical, the student would, in this model, work on solving the same task five times. Black and Wiliam (1998) argue that formative tasks should be novel for the student. Repeating the same task five times would work against that.

Nicol (2014) presents findings that “suggest that, up to a certain point, the more assignments that students are asked to review, the richer the evaluative process they engage in and the more likely they are to be exposed to works of different levels of quality

and to engage in productive learning transfer.” If the students get slightly different tasks, it gives them a wider range of exposure to different tasks within the same topic area.

### 5.3.1. Open and Uncomplicated

*Task* is the actual assignment that the students are to complete and later get and give peer-assessments on. As with the introduction, the form of the task can vary greatly depending on the subject at hand. In chapter six, I use an example task where the students translate a sentence from Norwegian to English. It can, however, be a wide variety of tasks, like drawing a sketch, taking a photograph, writing a short story, or almost anything else.

Black and Wiliam (1998) argue that if a task is too challenging, it might stop a student from getting involved. This corresponds with the findings of Hattie and Timperley (2007), who state that if the initial task is too complicated, the benefit from feedback seems to diminish. For a shorter peer-assessment process, this should perhaps carry extra weight as there is less time for understanding or working with a complicated task.

The experiences of Haugan and Lysebo (2018) suggest that the initial tasks should be open for discussion. A task that has a straight answer can make the feedback process stumped. A feedback statement that says if something is correct or wrong does not provide any information that the student can use. The very definition of formative assessment requires information to be used (Black & Wiliam, 1998). Hattie and Timperley (2007) highlight the question of *How am I going?* as essential for the feedback process. The initial task should, therefore, be open. It can be a task with several correct answers, a task that where the answer might seem obvious, but where there are other more optimal solutions, or other kinds of open tasks.

The suggestions here are to make the tasks open, but not so complicated that the weaker students fail to even begin the task. There can not be any peer-assessment intervention if the student is unable to start the task.

Black and Wiliam (1998) state that stronger students can better make sense of incomplete, missing, or poor instructions, and that weak performing students seem to be more reliant on instructions to be able to perform well on a task. The findings of Wang et al. (2015) support this. They found that stronger students would still have good learning outcomes even with less structure and instructions. This ties into the *uncomplicated* argument in this chapter. There should be as little room for misunderstanding of the initial task as possible. If a student has misunderstood the instructions for the initial task, it might make the following peer-assessment process almost meaningless as the students are, essentially, not working on the same task.

## 5.4. Feedback of Task

### 5.4.1. Feedback

Hattie and Timperley (2007) present four levels of feedback. Feedback about the task, feedback about the process of the task, feedback about self-regulation, and feedback about the self as a person. They argue that all of them except feedback of the self are beneficial for students, but that feedback about self-regulation and feedback about the process seems to give the best results for learning outcomes.

In this stage of the model, the students are prompted to provide feedback to another student's completed task. Feedback about the process of the task comes in the *Feedback on Feedback*-stage of the model, and feedback about self-regulation comes in the *Dialogue*- and *Culmination*-stage of the model.

Feedback about the task is most useful for clearing up a misunderstanding, and it is most powerful when there are only cues about the correct answer (Hattie & Timperley, 2007). So feedback given at this stage should aim to give feedback about elements of the task that the student has misunderstood or overlooked. Black and Wiliam (1998) point to a study showing that indirect feedback made students out-perform their peers that were given a complete solution, and could better apply their knowledge to similar tasks. The students should, therefore, not be prompted to correct all mistakes of the original task. They should instead be prompted to write suggestions or give cues to improvements that the first student can do. Nicol (2014) argues that written feedback creates new knowledge for the student as it engages students in revisiting and rehearsing their current knowledge. To promote reflection, students should write full sentences when they are reviewing other students.

Wang et al. (2015) argue that game-based elements like points or awards can be used in a peer-assessment setting. They show that it can give weaker students a boost in motivation. This, however, might not give the students as many opportunities for reflection as written sentences, but could be considered as a tool for increasing motivation.

The timing of feedback seems to affect learning achievements (Hattie & Timperley, 2007). More complex tasks are more likely to have a higher effect on learning outcomes if the feedback is somewhat delayed. At the process level and for less complex tasks, a more immediate response time proved to be beneficial (Hattie & Timperley, 2007). As mentioned earlier, this requires the task to be somewhat uncomplicated.

### 5.4.2. Anonymous

I suggest as part of the model that all students are anonymous in all parts of the peer-assessment process. The only exception is the teacher, who should have a complete overview of all activity. Below are several issues that all have the possibility of being mitigated by having the students be anonymous.

Chang et al. (2012) argue that high school students might not possess enough skills to perform peer-assessment on a social level. They argue that the teacher needs to oversee

and control the assessment process. Reviewing students' texts might not be enough, and perhaps regular interviews or questionnaires should be implemented to oversee the inner work-relationships, or, as Chang et al. (2012) state, the peer-assessment process could be anonymous.

Introducing peer-assessment introduces a shift in authority (Nicol, 2014). This might not be welcomed by everyone. Students' attitudes towards having other students review their work have been shown to be negative in some cases (Nicol, 2014). Nicol (2014) argues that the value of peer-assessment and a mutual agreement surrounding the process should be discussed with the students before starting peer-assessment assignments. When the student does not know who gave her feedback, she is forced to evaluate the feedback in a vacuum. The student does not know if the feedback came from someone they perceive as a strong or weak student. The student needs to critically evaluate the feedback and use and develop her own critical thinking skills.

Studies have shown that threats to self-esteem in students can nullify the benefits of formative assessment (Hattie & Timperley, 2007). They state that «The climate of the classroom is critical, particularly if disconfirmation and corrective feedback at any level is to be welcomed and used by the students (and teachers).» Lack of trust leads to a perceived lack of validity in feedback received from other students (Gamlem & Smith, 2013). In regards to peer-feedback, trust is an issue that should be considered. Students appreciate safe work environments where students support and help each other (Gamlem & Smith, 2013). If the students are confident that their work can not be recognized by their peers, it might alleviate the concerns the students have about threats to their self-esteem, and provide the safety needed for peer-assessment interventions.

## **5.5. Feedback on Feedback**

In this stage of the model, the students are reviewing and giving feedback on the previous feedback that another student provided. The focus shifts from discussing the task and over to evaluating the feedback that was given in the previous stage.

### **5.5.1. Reflection**

Giving feedback to others should provide students with a basis for reflection or mindfulness on their work (Black & Wiliam, 1998; Nicol, 2014). Students should, therefore, according to Nicol (2014), have written their assignments in the same topic area before reviewing another student's work. In this step, the students do not really address the initial task. They are addressing the input from other students. The students get prompted to evaluate the feedback that other students have provided. This line of prompting can potentially influence the students to have more elaborate written responses than if they are just correcting simple errors. This turns the feedback process into something more of an asynchronous discussion format. According to Nicol (2014), a well-reasoned written explanation helps students develop their tacit criteria, and it helps the students revisit and rehears their current knowledge in new ways. Perhaps just as important, discussion and dialogue involves the students in constructing, reconstructing, and co-constructing meanings (Nicol, 2014).

Nicol (2014) argues that students should look at peer texts from different vantage points to ensure a varied source for reflection. He identifies five possible variations:

- Holistic – Students are asked to view the entire peer text as a whole.
- Stakeholder – Students are asked to take a particular role as they review the peer text.
- Reader response – Students are asked to give their reactions to and feelings about the peer text.
- Graduate attributes – Students are asked to view the peer text from an attribute that the teacher wants the students to focus on. An example can be ethical or environmental.
- Contrastive – Students are asked to comment on a peer text from an opposite vantage point than what guided the initial assignment.

Nicol (2014) presents findings that show that giving explanations to others promotes deeper understanding and knowledge production. In giving explanations, students find holes in their own understanding and fill it by creating new knowledge.

Critically examining and evaluating the feedback is an essential part of peer-assessment (Nicol, 2014). Feedback given by experts (teachers) has another weight and meaning than feedback given by peers (Hattie & Timperley, 2007). Feedback from both experts and peers might be completely wrong and should be reviewed critically (Nicol, 2014). As a result of this, Nicol (2014) argues that critical evaluations of the peer feedback should be encouraged in the peer-review process. In support of this, this stage encourages an evaluation of the feedback from the previous stage.

## 5.6. Dialogue

In this step, the teacher leads a conversation about what the students have completed. The teacher displays an initial task with the answers given by the students. One single task contains input from three different students, as shown in Figure 1, below.

Initial task	
Answer to task	Student 1
Feedback to task	Student 2
Feedback to feedback	Student 3

Figure 1- A single task with input

The teacher can now lead the class in a dialogue about each of the individual steps. All of the students have completed similar assignments and should be able to contribute to the process.

This step serves several purposes. It gives the students a view of several other students work, and the teacher helps the students determine if what they are seeing is work of high quality. It facilitates synchronous dialogue, something that Nicol (2014) argues enhances all aspects of the peer-assessment process. According to Nicol (2014), dialogue in peer contexts “involves students in constructing, reconstructing, and co-constructing meaning together.”

By allowing students to discuss, they can test their assumptions by discussing them with their peers. However, for a student to participate freely in discussions, she needs to have an environment of trust and respect, just as in the rest of the peer-assessment process. This part of the model is different from the others because here, the students are no longer anonymous and hidden from scrutiny. This might require extra considerations in the method and in potential instantiations of the method. Having students discuss in pairs or small groups could, according to Nicol (2014), help advance the students enhance their collaborative skills, and also enabling the students to test their assumptions on a small audience before revealing them in front of a full class.

## **5.7. Culmination**

This is the final step in the model. In this step, the students get an overlook of all the different feedback that has been provided by other students. They get the opportunity to modify the initial answer to the task. A key aspect of formative assessment is that the assessment has to be used for something (Black & Wiliam, 2009). It has to be a part of forming the final result. By requiring students to modify their initial answer to the task, the final result should be more correct, but perhaps more important, the student can be more confident in their answers as they have gotten input from several different sources.

In addition, the student is prompted to answer questions focusing on self-assessment. According to Nicol (2014), self-assessment can ensure that the students learning gains are consolidated and strengthened. Increasing self-assessment abilities in students is one of the main benefits of conducting peer-assessments (Black & Wiliam, 2009; Hattie & Timperley, 2007; Nicol, 2014), and incorporating self-assessment into the peer-assessment process can aid the process further. Students can have trouble assessing themselves if they have no frame of reference (Nicol, 2014). A student, or anyone really, can not be expected to assess themselves if they do not know what to compare themselves with. The questions in this brings back the initial goals set out in the Setting Out step of the model. This could allow the students an opportunity to reflect on how well they have achieved these goals. However, the students are not only reflecting on the goals. Potentially, they have acquired differing perspectives from other students, and these will, according to (Nicol, 2014), greatly aid the students in reflecting on their own work.

## 6. Expanding the Model into a Method

### 6.1. Limitations

In this thesis, I propose a peer-assessment model based on existing theory concerning peer- and formative assessment in an attempt at answering what ways short formative peer-assessment interventions, using a scaffolded approach, contribute to self-assessment for students. In this chapter, the model is expanded into a method to provide a perspective into how the model might function. As argued earlier, the method also serves as data for discussion and also for providing some evidence of plausibility of the model. The model gets expanded by using it to propose a classroom exercise in a specific subject at a specific level.

Technology is a vital part for the method to function. Without the technological aspect, the method would be prohibitively difficult to implement in a classroom scenario. The technology part is not the main part of the research question, however, and some areas are not discussed that would need to be addressed if the method is to be implemented in a classroom. As an example, there is no discussion regarding security or ensuring safe storage of user data. There is also no discussion of Universal Design, basic usability, or general design theory needed to make a user-friendly website. These issues are essential for a final instantiation but do not provide much, if any, data relevant for discussion in this thesis.

The method proposed in this chapter is a general view of a peer-assessment. It does not consider local conditions of a school or a class, nor does it consider the individual needs of students or teachers. This would be considerations for an instantiation with a test group. The target for the method proposed is students in 8<sup>th</sup>-grade Norwegian school students and is based around an English assignment. The grade and subject are chosen somewhat arbitrarily by the author as the method could have potential in almost any subject. It is not suggested that this particular grade or subject is better suited for this method. The choice of subject and assignment is, however, chosen as the author has some years of experience with this particular age group. More importantly, the method is used to develop an existing assignment type and process that has been in use at the authors' local school for over two decades —more about that in the next subchapter.

The purpose of this tool is to enhance learning and training by using formative peer- and self-assessment. And, as mentioned earlier, it is not supposed to be limited to a specific school subject like English. Subject-specific learning, like English didactics, will not be in focus as it is more of a secondary benefit in regards to the research question in this thesis. This does not mean that the method is not beneficial for the students' subject-specific learning or goes beyond or outside of the upcoming curriculum in the The Norwegian Ministry of Education and Researchs Knowledge Promotion that is getting implemented in Norwegian Schools this fall. The first point in the list of competency goals for Norwegian 10<sup>th</sup>-graders in the upcoming Knowledge Promotion says that students shall use varied strategies in language learning, text-creation, and communication (Norwegian Directorate for Education and Training, 2020 - translation to English by the author). As the method in this chapter can be viewed as a tool for expanding and evolving learning strategies through self-assessment



while simultaneously working on a language-specific assignment and communicating with other students, I would argue that the method is well suited for this competency goal. Another competency goal in English for Norwegian 10<sup>th</sup> grade says that students should process their texts based on feedback and knowledge about language (Norwegian Directorate for Education and Training, 2020 - translation to English by the author). This is a goal directly aimed at feedback. It does not specify the source of the feedback and is probably intended to be open for differing sources of feedback. It does, however, state that the goal is for students to process *their own* texts based on feedback. The method proposed in this chapter does not use original student texts, and it can be argued that it does not directly meet the competency goal. But, it could be a suitable tool for teaching students how to process and use feedback. So, even though subject-specific didactics are not discussed in greater lengths, that does not mean that this method does not contain valuable properties for reaching learning gains in this particular school subject.

There are justifications, explanations, and thoughts described in the following description of the method. If they are not followed by a citation or directly linked to the model in chapter 4, they are the authors' own justifications, explanations, and thoughts, and not tested in any scientific process. The relevant parts of the method will be more thoroughly linked to the model in chapter six and discussed in regards to what we know from the literature presented earlier in this thesis.

### **6.1.1. Technological Limitations**

This method is proposed to function in a web-based environment. There are technological obstacles associated with creating a teaching tool. The proposed method is, however, not radically groundbreaking from a technological standpoint. The entire process could be compared to a messaging system, only that the flow of messages are diverted in other directions. From the authors' experience, technology widely used in existing web-applications like forums, messaging, or LMS is more than adequate for the method proposed in this chapter. An example of one of the most technology-complex parts of the method is providing the teacher with a live view of the students' progress. This can be handled by several types of existing server technologies without any special modifications. One of these is Node.js, which is an asynchronous event-driven JavaScript runtime (Node.js, 2020 2020). Node.js is designed to build scalable network applications and is in use by eBay, General Electric, GoDaddy, Microsoft, PayPal, Uber, Yahoo! and Netflix (Node.js, 2020 2020).

Limitations based on technology might occur in a later development stage. However, from the current perspective of the method, there does not seem to be any technological limitations that would impede the functionality of the proposed method.

## **6.2. The Original Assignment Process**

The basis for the assignment in the proposed method has its base idea from an assignment process at the authors' local school. The original assignment process has been in use for over two decades and is used by all foreign language teachers at the authors' local school. The process was created by a life long English teacher named Bjørn Krogsæter. He was frustrated

with assignments in English textbooks for Norwegian students. The weekly assignments in these books did not retain grammatical lessons in subsequent assignments. According to Bjørn, a chapter (A) revolving around some kind of grammatical rule would have assignments suited for dealing with this rule. However, in the following chapters (B, C, D, etc.), there would not be any assignments regarding the rule in chapter A. There would be no progress or cohesion from chapter to chapter. The students would, as a result, have few repetitions or reminders of skills they had acquired earlier in the school year. Bjørn's solution to this was to make weekly translation assignments that progressed along with the students' expected knowledgebase. Grammar taught in chapter A would still be present as a repetition in all subsequent chapters from eighth grade through tenth grade.

The process Bjørn developed had weekly assignments with about ten sentences in Norwegian, and the task for the students was to translate them into English. The students were set to do this task on their own, sometimes at school and sometimes at home. The teacher would then go through each sentence on the blackboard and get input from the class about how to translate the sentences. All pupils should have their solutions and suggestions for how to translate the sentence in front of them. Since all students should have worked on the assignment, they should be prepared for a classroom discussion and get opportunities to view examples of high-quality work.

From the start, this system already incorporates elements of formative assessment. The discussion in the classroom at the end is a vital part of the exercise. The students get to provide their solutions to the problem, and all of the students adjust their own sentences after the discussion is completed. As mentioned several times earlier in this thesis, using the feedback to help form the answer is essential for formative assessment.

All sentences should contain some form of trap, according to Bjørn. The traps can be grammatical, based on commonly misused words or other language-based pitfalls. This enables the teacher to control the desired learning outcomes by tailoring the sentences after what skills the students seemed to be lacking.

To highlight a couple of other reasons why this assignment process should fit well within the proposed method, it might be beneficial to look at some examples of sentences. The following are some sentences from the worksheets that 8<sup>th</sup>-grade students are asked to translate from Norwegian to English.

1. Jeg børstet håret når den sjalu ektemannen kom inn.
2. Neste år skal Japan arrangere OL.
3. Jeg fant skatten da vi leitet etter skiene mine på loftet.

According to Bjørn, all students with a rudimentary understanding of English can attempt to translate these types of sentences. It can generally be assumed that almost all students in Norwegian schools can read a sentence in Norwegian. There are, of course, students in Norway that can not read Norwegian, but these are at the far ends of the spectrum of what can be expected of students. Students in Norway that follow standard curriculum and progression have also had lessons in English since early primary school and should have a basic understanding of English. In the example-sentences above, there are basic English

words in each sentence. Words like *jeg* and *year* are not difficult to translate for most students in 8<sup>th</sup> grade. And having the student be able to attempt the task is important for peer-assessment interventions. As discussed earlier and argued by Hattie and Timperley (2007); Nicol (2014), there can not be peer-assessment if the students do not have adequate knowledge beforehand to attempt the assignment.

The first example-sentence is focused on the correct use of ing-form. *I was brushing my hair when the jealous husband came in.* Only the first verb gets -ing ending.

The second sentence focuses on a commonly mistaken translation of the Norwegian word *arrangere*. The similar-sounding English word *arrange* has a completely different meaning from its Norwegian counterpart and should rather be translated with the word *host*. It also deals with sentence structure as it can be more common to say *Japan is hosting the Olympics next year.*

The third sentence uses the Norwegian word *skatt*. This word in a vacuum can be translated into several different English words – treasure, tax, or honey. According to Bjørn, students often write *on the attic* instead of the correct *in the attic*.

### **6.3. Description of the Method**

As described in the previous segment, the proposed method has its origin from a paper-based English assignment. The proposed method is web-based and entirely digital. The digital interface for both students and teachers is from here called *the website* for ease of reading.

As was discussed in chapter three, the construction of the website and the method are not the focus of the research questions in this thesis. The method's purpose is to provide perspective to the model for how the model might function and what ramifications this might have to the model and, ultimately, in what way the model of a peer-assessment instantiation can contribute to self-assessment. The method and the website are built upon and expanded from the model in chapter four. Nevertheless, the author has taken creative liberties in creating the illustrations of the website. It is but one version of how the model might manifest in a specific school level and in a specific subject. Different subjects, assignments or age group would probably provide different perspectives to the model. The author has five years of work experience in building websites and has a bachelor's degree in media, ICT, and design.

The illustrations in the following chapter are illustrations for the reader of this thesis. Their purpose is to provide a clearer perspective of how the method functions. The illustrations are not indicative of how a final instantiation might look like from a design perspective.

A timeframe for a class to complete the entire peer-assessment instantiation might be difficult to predict accurately, but hopefully, the complete method described below could be completed in 10 to 20 minutes if the teacher and pupils are experienced with it. This timeframe does not include a thorough introduction to the learning material but assumes

that the class has had some introductory instructions to the relevant subject-specific curriculum.

### 6.3.1. Setting Out

The teacher has a class of students gathered in the same physical classroom, and all participants are using their own devices and are logged in to the website with unique user accounts.

As illustrated in Figure 2, the teacher can select a class, a topic, and a level for how much assistance the website should provide in setting goals for the assignment. As mentioned in chapter 4.1.2, students might benefit more in forms of gains in self-assessment if they form their own tacit criteria.

**Start peer assessment session**

Choose a class

Class 8A

Choose a topic

Standard: Mix all

Goals \*

Provide detailed goal suggestions. Suitable for beginners in peer assessment.

Provide room for class-creation of common work goals. Suitable for intermediates in peer assessment.

Do not provide room for goals. Suitable for seasoned veterans in peer assessment.

Show tutorial before session starts

**Start session**

Figure 2 - Start screen - Teachers perspective

The teacher can also choose to show a tutorial at the beginning if the students are beginners. The tutorial is a video that shows the peer-assessment process. It highlights that no student will know anything about who has written what during the entire process. It also shows that the teacher can, at any point, see the progress made by each student. The tutorial video also has a brief explanation peer and self-assessment and that the feedback the students provide is just as valuable as answering the initial task.

During the whole process, the teacher has a view of all the students in the class, and the students get to see examples of this in the video tutorial. The teacher can monitor the progress of each student and assist students that are stuck in a particular step of the

process. See Figure 3 for a visualization of how this might look. Additionally, the teacher can click on each student to get a full view of their contributions so far.

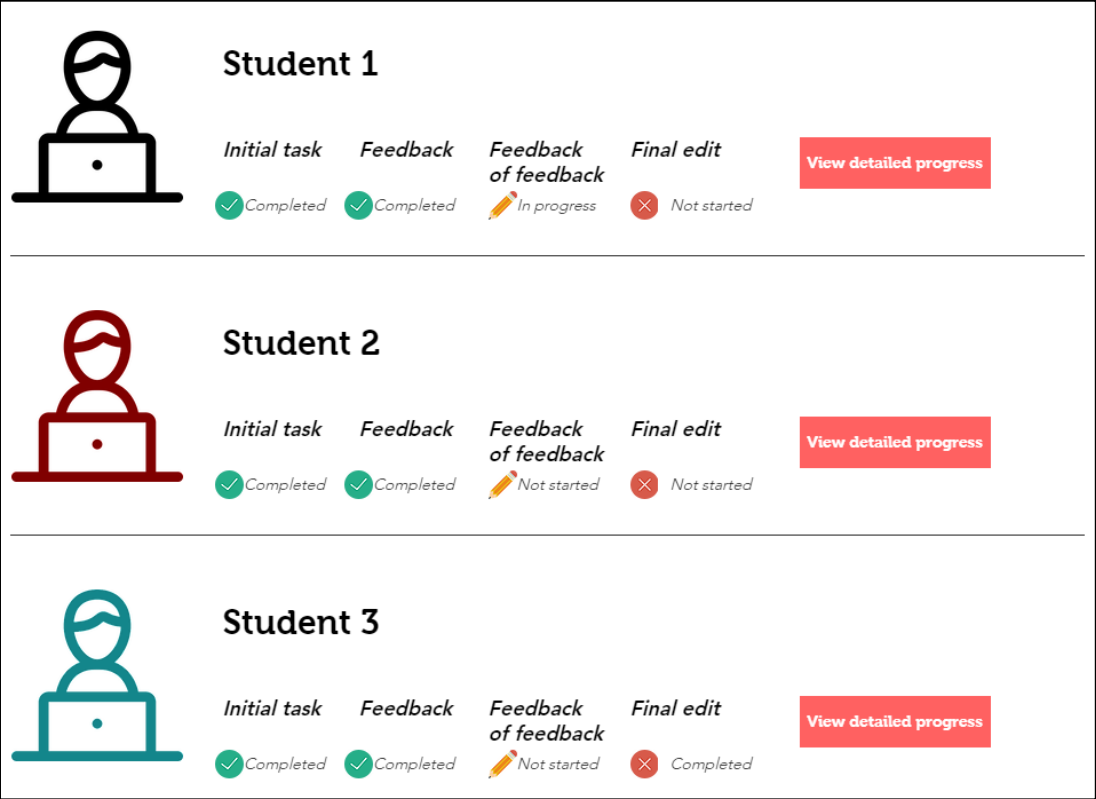


Figure 3 - Student overview - Teachers perspective

Before the students get assigned their sentence, the teacher can display a screen for the whole class where they are prompted to enter goals for the peer-assessment instantiation collaboratively. Depending on the teachers' choices in the setup, the goal fields are displayed with varying degrees of support. The class can then suggest goals for the session based on the support. Figure 4 illustrates an example of a goal screen with some support for the questions. As students get more experienced with these types of peer-assessment sessions, the teacher can choose to have less support for creating the goals, have the students enter goals individually, or not enter goals at all.

This assignment contains Norwegian words that has multiple meanings. Create a goal that suits this challenge

*Understand the meaning of different translation optio...*

This is a collobartive assignment. Make a goal that suits this challenge

*Help improve someone else's translation*

Submit

Figure 4 - Setting goals - Class perspective

### 6.3.2. Task

After the teacher has initiated the session, each student gets a single sentence to translate, as shown in Figure 5. As mentioned earlier, the sentence is suitable for all Norwegian students that have a rudimentary understanding of English. It is, however, not overly simplistic. There is a vast difference between this sentence and, for example, *I have a cat named Scruffy*. The sentence is difficult enough to provide grounds for discussion. The sentence used in the illustrations is just used for illustration purposes.

#### Translate this sentence from Norwegian to English

*Jeg fant skatten da vi lette etter skiene mine på loftet.*

Type here...

Submit

Figure 5 - First task - Student perspective

All students could be assigned the same sentence, or they could each have a unique initial sentence, or there could be a mix where different groups of students get the same sentence.

Figure 6 illustrates Student 1s perspective if all students were assigned the same starting sentence. Throughout the process, Student 1 would only see variations of the same sentence. If all students were assigned the same sentence they would become more familiar with that particular sentence, and be could potentially become more confident in discussing it in front of the class. It could, however, limit their exposure to differing approaches to solving the task. It would also be more repetitive than the other alternatives.

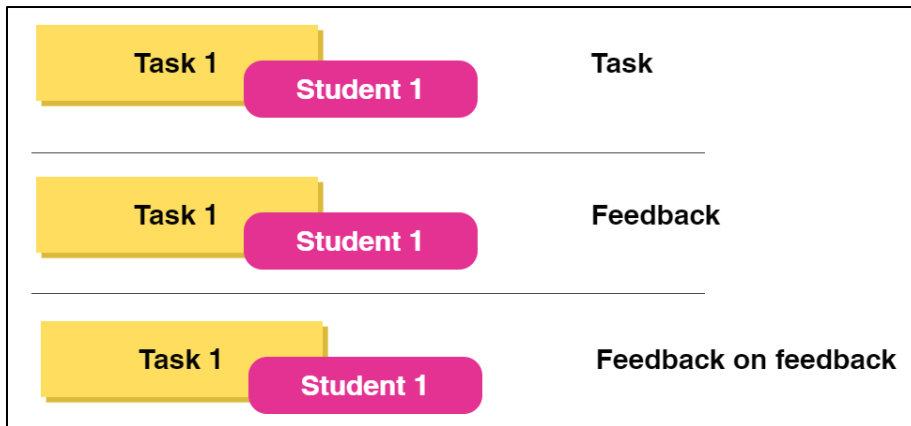


Figure 6 - Exposure to different tasks if all tasks are similar

Figure 7 illustrates Student 1's perspective if all students were assigned a unique initial sentence. Student 1 would be exposed to three different sentences, but all three with the same type of obstacle. This provides more variety for the students even though each sentence is just a variation of the same type of obstacle. But in the dialogue-part of the method, when the teacher displays the results in front of the class, the teacher would have to display each individual task if the goal is to provide each student feedback about their task. In a class of 20 students, this would take up a lot of time and effort. And discussing 20 sentences could feel unnecessary for each student as they are really only invested in the sentences that they were involved in.

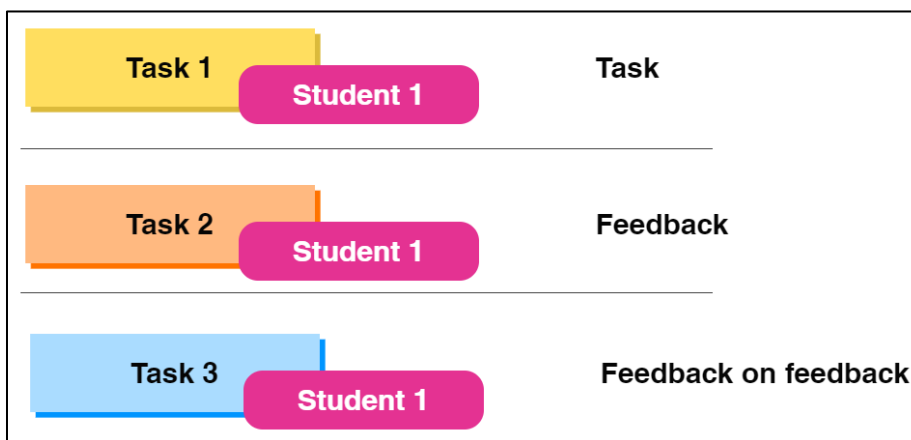


Figure 7- Exposure to different tasks if all tasks are unique

An alternative to all sentences being unique and all sentences being identical is a mix of unique and identical sentences. As each student is exposed to two other students' work throughout the individual steps in this method, there would have to be a minimum of three different sentences total in the class. If there are three total unique sentences, an illustration similar to Figure 6 and Figure 7 would be identical to Figure 7. Figure 8 is an illustration of the workflow of three students. Each student would all start with individual sentences, and only see new sentences in each step of the process. In the dialogue-part of the method, the teacher would only have to go through three different sentences, but still be able to discuss each sentence. And in that scenario, each student would have something invested in all of the three sentences that are discussed in the class as they have worked on all three. In a

whole school class, it would not be necessary to group three and three students. The sentences could be mixed all through the class as long as the students get a sentence they have not encountered before in each step.

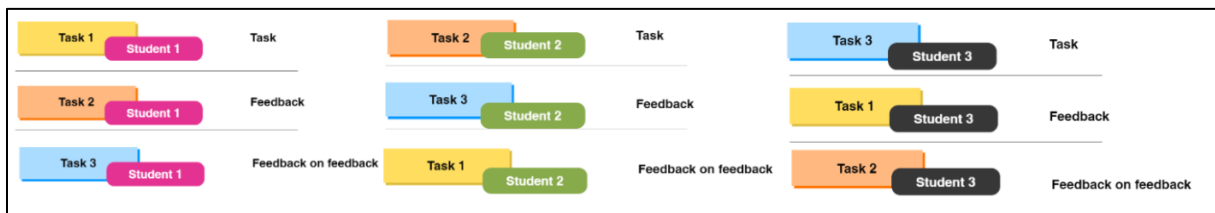


Figure 8 - Workflow of three students

### 6.3.3. Feedback of Task

When the first student has completed the initial assignment, she gets a waiting screen until the next student has completed their translation. When another student have completed the initial assignment, Student 2 will be prompted to improve to Student 1s translation, as seen in Figure 9. This sentence is a different one from the one Student 2 translated initially.

**Can you improve the translation below?**

*Jeg fant skatten da vi lette etter skiene mine på loftet.*

*I found the award when we looked for my skis on the attic.*

*I found the treasure, when we were searching for my skis in the attic.*

Submit

Figure 9 - Giving Feedback - Student perspective

### 6.3.4. Feedback on Feedback

When Student 2 has completed the task, a third student will be presented with both Student 1s translation and Student 2s improvement, as can be seen in Figure 10. Student 3 is asked to assess the two translations and evaluate the differences. She is also asked to provide a reason for the choices she makes.



**Look at the differences between the two translations. What elements are correct? Provide a reason for your selection**

*Jeg fant skatten da vi lette etter skiene mine på loftet.*

*I found the award when we looked for my skis on the attic.*

*I found the treasure, when we were searching for my skis in the attic.*

*Treasure is the correct use of the word because award means belønning*

Submit

Figure 10 - Feedback on feedback - Student perspective

### 6.3.5. Dialogue

The teacher now has the opportunity to display the complete work done by the students on a screen in front of all the students. The teacher displays one task at the time, but with the complete revision history visible. All parts of the process can be assessed and discussed, involving the whole class in the complete process.

The teacher has the opportunity to have a full class discussion or to divide the class into pairs or small groups to discuss and assess the feedback that has been presented to them.

**Jeg fant skatten da vi lette etter skiene mine på loftet.**

1 - I found the **award** when we **looked** for my skis **on** the attic.

2 - I found the **treasure**, when we **were searching** for my skis **in** the attic.

3 - *Treasure is the correct use of the word because award means belønning.*

Class and teacher suggestions \*

*I found the treasure, when we were searching for my ...*

Submit

Figure 11 - All contributions - Class perspective

### 6.3.6. Culmination

In the final step of the method, the students return back to their original task, but with feedback from two other students, and feedback from the teacher/class. This is illustrated in Figure 12. The student is then prompted to revise and change the initial translation based on the feedback she has received from other students. The student has also seen two other sentences with a similar initial issue during the process, and can compare her own sentence with the other two sentences.

The students are also asked to evaluate themselves based on the goals set at the start of the peer-assessment instantiation. The goals are displayed for the student to remind them of what they as can be seen in Figure 12. If the students are very experienced with this type of peer-assessment, and the teacher has elected to not start the session with setting goals, the student will be prompted to evaluate their performance without any other support.

1 - I found the *award* when we *looked* for my skis *on* the attic.

2 - I found the *treasure*, when we *were* searching for my skis *in* the attic.

3 - *Treasure* is the correct use of the word because *award* means *belønning*.

Teacher and class suggestion - I found the *treasure* when we were searching for my skis in the attic.

**Goals**

1. Understand the meaning of different translation options of Norwegian word.
2. Help improve someone else's translation.

Look at your initial translation and see if you can improve on it. \*

*I found the treasure while we were searching for my s...*

How well did you meet your goals? Describe \*

*I learned the different meanings of skatt and I found a...*

Submit

Figure 12 - Final review of the translation - Student perspective

## 7. Discussion

### 7.1. Discussing and Revising the Model

In developing the model and the method, it is not easy to say what insight or idea occurred at which stage of the process. But most of the insight into how the formation of the scaffolding came during the development of the method. The model presented in chapter four, Table 2, is based on scientific literature. The model below, Table 3, is a revision of the previous model. The revisions are based on the method proposed in chapter five. These adjustments are not empirically tested, and for discussion purposes and can be seen as suggestions for areas for further study. The revisions are marked with red. Below the revised model is a discussion into each of the revised elements of the model.

	Setting Out	Task	Feedback of Task	Feedback on Feedback	Dialogue	Culmination
Activity	Introduction Co-created-goals	Open and uncomplicated	Feedback Anonymous	Triple perspective	Dialogue	Self-reflection
Scaffold	Ques Process-Info	Bank	Anonymity Automatic assignment	Automatic assignment	Visualization of solutions	Self- assessment cues

Table 3 - Revised model

#### 7.1.1. Setting Out

The scaffold at the start of the method provides an insight into the process (Process-Info) by offering a video showing the process. Giving students insurance that their contributions will be handled carefully can provide some of the necessary safety for effective peer-assessment interventions. Ensuring that the students feel safe is important for ensuring meaningful peer-assessment interventions (Hattie & Timperley, 2007; Nicol, 2014; Pachler et al., 2009).

The scaffold also provides aid for the teacher with pre-set lessons with completed tasks and goal-suggestions, potentially alleviating work from the teacher. Alleviating workload from the teacher can be a goal for ensuring that teachers see that the value of effort they put into the process corresponds with achieved learning gains from the students. By utilizing the scaffold to support the teacher can potentially be beneficial for expanding a teachers

competency in peer-assessment situations by providing nodes of expertise that the teacher can utilize.

The proposed method does not directly say what the goals for the process should be. It provides hints and cues for the teacher and the students. Pachler et al. (2009) describe this as soft scaffolding. Providing support, but trying not to be too rigid and impede on individual expression, exploration, and innovation. The scaffold also provides an option for opting out of the pre-set goals and create goals without the support or leave it up to each individual student to set their own (written or unwritten) goals. In the study of Wang et al. (2015), only the high achieving students benefited from a less structured and less goal-oriented peer-assessment process. Study on the impact of learning gains by removing the goals after the students had significant experience in using an instantiation of the method would be necessary to determine any potential effect on different groups of students.

The digital scaffold proposed in the method does not address all of the approaches for encouraging an atmosphere of trust and respect suggested by Nicol (2014) and highlighted in chapter five of this thesis. The second approach presented by Nicol (2014) suggests an illustration of how peer review can operate in professional contexts and life beyond school. This can be illustrated or discussed in the tutorial video, but it might be more appropriate for the teacher to utilize the students' world-view and use examples that are closer to the experiences of the students. Nicol (2014) also argues that simply telling the students the benefits of peer-assessment could be insufficient in conveying the information to the students. He suggests workshops with discussions as a possible solution. I would argue that trying to solve all problems with a digital scaffold could potentially have less effect than the solution Nicol suggested. Since the method allows for several iterations of the process with the same student, the process itself could potentially help encourage an atmosphere of trust and respect. A student's first apprehensions about the process could be alleviated by carefully trying it out. If the student feels secure in that no one will recognize them based on the work they submitted, they could potentially be more trusting in the second iteration of the process. And since the process is not a massive investment of time for the students, it would not be a huge loss of potential learning gains if they are apprehensive and careful the first time and the more confident in the next. If the students can see the benefits of the process, they could also

### **7.1.2.Task**

The tasks presented in the method are examples of tasks that can be viewed as open and uncomplicated. Translating a sentence does not require much in the ways of introduction to the material other than relevant introductions to grammar and language, yet it can provide ample opportunity simply based on the complexity of language-learning. Teachers answered that language learning was too complex for peer-assessment in the article of Wasson and Hansen (2014). This might be the case for a longer piece of student-submitted text. But a pre-selected sentence with pre-set pitfalls is arguably less complex. Other types of tasks might require more introduction depending on the students' previous experiences with the subject. Keeping the tasks somewhat open or that the task provides some are for discussion

can be necessary for the method to yield good learning outcomes, as was the experiences of Haugan and Lysebo (2015). Black and Wiliam (2009) and Hattie and Timperley (2007) do stress the need for tasks to be simple enough for the students to understand what is asked of them. If the students are unable to answer the initial task, the feedback will be of little value. Further research would be necessary to reveal how and why the task can affect shorter peer-assessment interventions.

The scaffold in this part provides the teacher with a bank of premade tasks to administer to the student. The teacher is free to create her own tasks, but providing pre-made tasks that are suitable for peer-assessment feedback saves time for the teacher. It can also potentially influence the teacher into creating similar, equally suiting, tasks later.

Giving a student three different tasks to consider during the process can be beneficial according to the argumentation of Nicol (2014) that working of different tasks in the same topic domain allows the students to develop their own concept of quality.

### **7.1.3. Feedback of Task**

It is this step and the next, Feedback on Feedback, that the scaffold potentially has the most significant benefit to the method. Bjørn Krogsæter and other teachers have used the tasks and sentences as school assignments for decades. That process already utilizes the open, yet uncomplicated tasks and the classroom discussion. It is the anonymity and automated assignment process that allows for potentially easy peer-assessment situations.

In chapter five of this thesis, I presented arguments, based on scientific literature, saying that only hints or suggestions should be provided in feedback, and not complete corrections to the initial task. The proposed method in chapter six does not prompt for hints or suggestions but asks for complete corrections. The scientific evidence used in presenting the arguments for only providing hints in feedback is not specific for peer-assessment. It is based on a general concept where the person providing feedback is often considered an expert. In this method, the person providing the feedback is a random anonymous student. The initial student does not know if the provided corrections are correct or not. The student would, therefore, have to consider the correction as a sort of hint or clue that something is possibly wrong with the initial translation. A random anonymous student's feedback could possibly have some of the same effects as feedback cues and hints provided by a perceived expert. The differences between complete corrections from an anonymous random student and hints and cues from a perceived expert would have to be studied further to determine the actual differences it could have.

The issue of group compositions based on student levels in peer-assessment assignments was raised earlier in this thesis. The proposed method does not contain any considerations for group composition. However, there are some arguments for why group composition is of lesser importance in this particular method. The peer-assessment instantiations are short and provide the opportunity for conducting a higher volume of peer-assessment instantiations than would else be possible. If one group composition is less than ideal, it would potentially even out by conducting more peer-assessment instantiations. By having

random group compositions, you will eventually have all different kinds of group compositions. Another point of interest is the class discussion part of the method. It provides input from high achieving students and the teacher. So even if the initial group consists of low achieving students, it would still provide the necessary input to avoid mislearning. A third argument is that by having the group compositions be anonymous and randomized, it could possibly enable a stronger need for critical skills from the students. They can not safely assume that it was a high achieving student that provided the feedback.

The feedback the students receive, should, if the method functions as suggested, be delivered to the students only minutes after they submitted their initial answer. In the time between the student submitting the initial answer and receiving

#### **7.1.4. Feedback on Feedback**

As mentioned in the previous subchapter, Feedback of Task and Feedback on Feedback might be the two steps of the model that benefits most from the digital scaffold because of the possibilities it has for anonymity and automation.

In the proposed method, it is not Student 1 that reviews the feedback. It is a third student. This gives the Student 1 a second perspective on her original task. It still is not an expert providing the feedback, so the student still needs to evaluate the feedback given by both peers critically. If the feedbacks differ or deliver opposing views, the original student has arguments for both sides and gives her a better base for assessing the answers.

Involving a third student has another potential advantage. As discussed earlier, Nicol (2014) argues that exposing students to a wider range of perspectives is beneficial for developing self-assessment strategies. In this model, each student gets at least six different perspectives on solving tasks within the same topic.

Perspective 1 – The student solves a task.

Perspective 2 – The student gives feedback to another student's task.

Perspective 3 and 4 – The student sees and gives feedback to a third student's original task with accompanying feedback from a fourth student.

Perspective 5 and 6 – The student gets back her original task with the perspective of two other students.

If we compare this to a version where the original student gives feedback on the feedback about her own task, she would only get three different perspectives.

Perspective 1 – The student solves a task.

Perspective 2 – The student gives feedback to another student's task.

Perspective 3 – The student gets back her original assignment with accompanying feedback.

In both versions, all the students have given almost the same amount of feedback and completed the same amount of work. But the first one provides double the number of perspectives the student is exposed to.

Adding *Third* to the model is the only change in the model besides the additions to the scaffolding sections. If a broader range of feedback and perspective is beneficial to the students' development of self-assessment skills as Nicol (2014) argues, then this step in the

model might be the most important contribution in regards to student benefits from shorter peer-assessment interventions. This would have to be studied further in order to determine how and why it might affect the process.

As argued by Nicol (2014), providing feedback to another student is a different mental process than simply providing the answer. Providing feedback on feedback can possibly enable yet another process that enforces the learning gains by the students. Providing Feedback on Feedback is also highlighted by Pachler et al. (2009) as a vital element for improving student's feedback abilities.

As was highlighted in the construction of the model in chapter five, Nicol (2014) argues that the students providing feedback can look at the text from different vantage points when they are providing feedback. He suggests several vantage points for the student can assume. These vantage points are meant for students at higher education, providing feedback on longer texts. In constructing the method, these vantage points did not seem to fit very well with the task, and would probably not provide many benefits to the students in the forms of learning gains. The type of task suggested in the method has a somewhat narrow set of avenues for learning not suitable for the type of general vantage points suggested by Nicol (2014). As an example, asking a student to provide feedback to a translated sentence from the bases of an environmental standpoint would probably take attention away from the task at hand. If the students were prompted to provide feedback from a specific vantage point, it would have to be a vantage point that is chosen based on the current task at hand. The types of vantage points suggested by Nicol (2014) could possibly be implemented into another version of the proposed method, but they did not seem to be advantageous in this iteration of the method.

### **7.1.5. Dialogue**

The digital scaffold provides some potential benefits to the peer-assessment process in this step by utilizing the patterns described as Classroom Display and Showcase Learning by Pachler et al. (2009). The students get to view examples of solutions to assignments they have worked on with a trusted audience. As Pachler et al. (2009) puts it: the teacher has the opportunity to align instructions with students' perspectives and their current state of knowledge. In the model, I have added this part as *Visualisation of solutions*, as it provides an opportunity for all students to see the solutions to answers, feedback and feedback and feedback. This visualization allows the class to carry out a discussion in comparing the different solutions. Visualization of answers was possible with the old process as well, where the teacher would write the words on the white-board. The benefit here is that the students can see multiple all the steps from three different students in addition to having worked on it themselves. This can aid the students' confidence in suggesting reasoning or differing solutions. The teacher also has the opportunity to celebrate student work publicly. By doing this, the teacher provides the students with examples of high-quality work. As highlighted earlier in this thesis, seeing high-quality work is beneficial for developing self-assessment skills, according to Nicol (2014). Moreover, in this case, it is not only the result that can be

highlighted, but each step of the process. The teacher can point out high-quality examples of feedback from the *Feedback on Feedback* step.

The discussion step in the proposed method provides several potential beneficial aspects to the peer-assessment process. The first one is based on the arguments from Nicol (2014), stating that discussion can strengthen and consolidate all benefits of the peer-assessment process. A second potential benefit is that some students might generally be apprehensive about discussing their ideas publicly. However, in this situation, they have seen several examples of the tasks conducted by other students. Even if the student and all other students have answered incorrectly in the previous steps of the method, they at least know that they are not alone about their assumptions. They can even take cues from the line of answers visible on the screen at the front. A third potential benefit is that the students might have a broader perspective of the issue at hand after working on and getting feedback on the issue and knows how to formulate questions that can clarify or confirm the students' assumptions. A fourth potential benefit is that if the students are discussing in pairs or small groups, they can potentially develop their collaborative skills, according to Nicol (2014), and boost their confidence in similar tasks. A fifth potential benefit of dialogue based upon peer-assessment, according to Nicol (2014), are gains in the students' abilities to self-assess. By discussing their own or other students' submissions, they will likely evaluate the discussion up against their own submission.

All of these potential benefits are assumptions based on the model and the method and would have to be studied further to determine how beneficial these effects might be.

### **7.1.6. Culmination**

In the final step of the method proposed in this study, the students get prompted with two questions. The first question asks the student to improve her initial translation based on the feedback that has been provided. This is the first time the student sees all of the feedback given to her translation. Only adjusting the initial answer once is a strategy suggested by Pachler et al. (2009) as a way to discourage students from performing their minimum and only using the feedback without careful consideration, or groping along as the authors describe it. However, giving the teacher a full view of the entire process can potentially also alleviate the problem of students groping along as the feedback the student provides during the process can stand on equal ground with the final answer as an insight into student effort.

The second question reminds the student of the goals set at the start of the peer-assessment instantiation, and prompts the student to self-assess. As discussed earlier, peer-assessment can help students build skills in self-assessment, and Nicol (2014) argues that including actual self-assessment in peer-assessment interventions can help consolidate the self-assessment skills the student has developed throughout the process. Having the students return to the initial question can help the student in the self-assessment process. The student can look at the goal for a perspective of how effective the process has been and what they have learned. In regards to potential developments of self-assessment skills, the student has gotten a number of perspectives and insights from others during the process. As highlighted earlier, the student has seen and worked on six other students' perspectives by



the end of the Feedback on Feedback step. And after that step, the student was a participant in the class discussion. The number of perspectives would depend on the form of the discussion and the number of inputs from students. But if the teacher gets input from just a single student on each of the three sentences, and the teacher provides her own perspective, then there are six new inputs that can help the student form their basis for self-assessment.

As a hypothetical, let us assume that each step in the proposed method takes five minutes to conduct. Five minutes for introduction, five minutes for the first translation, five minutes for feedback etc. The total of all steps is 30 minutes. Then, let us assume that a more traditional peer-assessment process with a longer answer takes about 30 minutes to complete. If the longer peer-assessment went through the same steps, that would come to a total of three hours. Moreover, the teacher would not be able to adequately discuss openly in the class each student's assignment in that timeframe. Simply having step 1 and 2 would require a full hour of work. Utilizing the proposed method for longer peer-assessment interventions would require a substantial amount of time.

An important note is that the student has worked on all of the sentences and is, at least to some degree, invested in the process. And since all the sentences are within the same topic domain, the results from each can help the student reflect on her original task.

### **7.1.7. Not in the Model**

There are some elements that are discussed earlier in this thesis but have not been added to the model or the method. Pachler et al. (2009) and Wang et al. (2015) argue that some form of reward-system can be beneficial for the motivation of, at least, some groups of students. There is no reason that this could not be implemented, but the effects of these types of elements are, in some cases, varied and often more beneficial for increasing motivation among students rather than increasing gains in learning (Wang et al., 2015). It could also fit better in a larger eco-system of multiple school activities like a Learning Management System.

Pachler et al. (2009) argues that a digital solution for formative assessment can benefit from giving the students the ability to highlight elements that they are talking about. They argue that we point at things when we talk and discuss things in real life and that this ability can be beneficial in digital tools as well. The feedback elements of the method were developed with the benefits of having the students write complete sentences (Nicol, 2014) in mind. An earlier iteration of the method consisted of students giving visual thumbs up or down to elements of the student work they were assigned to assess. A way for students to highlight elements of what they are providing feedback is something that should be considered in potential studies into the model.

## 8. Conclusions

The research goal in this thesis includes the term *short* in addressing peer-assessments. The term has appeared as a complex construct in conducting this study. It has had several ramifications and effects on both the model and the method in regards to both students and teachers.

In the most discussed elements of literature in this thesis (Black & Wiliam, 1998; Hattie & Timperley, 2007; Haugan & Lysebo, 2015; Nicol, 2014), there are arguments for beneficial learning gains by providing timely feedback to students. The proposed method attempts at achieving this by utilizing both peer-assessments and a digital scaffold. Peer-assessment allows for potentially much speedier feedback than the teacher could achieve alone. This is not always utilized by teachers conducting traditional long-form peer-assessment interventions, as discussed and highlighted in chapter three. Not providing timely feedback does not necessarily nullify the potential benefits from peer-assessments, but it might mean that some potential benefits are lost. The method, with its scaffold, provides an example of how technology can assist with keeping peer-assessments short. The process of issuing tasks and assigning feedback tasks as quickly as the method suggests would probably not be feasible without an automated system that handles the flow of communication.

The term short also had an impact on the nature of the initial tasks given to the students in the proposed method. A short process of the proposed method would require that the students are able to answer the question without being able to study the problem at hand before they attempt to solve the task. Having clear instructions and having the student be able to attempt to solve the task is highlighted as vital for peer-assessment by Hattie and Timperley (2007) and by Nicol (2014). But in the method proposed earlier, this might be an even more significant issue as there is less time for clarification.

According to leading researchers in the field of formative assessment, peer-assessment has a positive effect on students' skills in self-assessment (Black & Wiliam, 2009; Hattie & Timperley, 2007; Nicol, 2014). The model and method proposed in this thesis revolve around conducting short peer-assessments. By conducting short peer-assessments, there can potentially be room to conduct a higher volume of peer-assessments. Having a higher number of perspectives and providing feedback to works of varying quality is one of the more effective ways of increasing gains in student self-assessment skills according to Nicol (2014). By conducting a higher volume of peer-assessments, the students get exposed to a higher number of perspectives and have a higher number of opportunities for self-assessment. The method also allows a student to perceive and work with, at least, 12 different perspectives to an assignment. A more traditional peer-assessment with only one other student, and potentially also a teacher, provides far fewer perspectives in comparison. Having the tasks be short allows for a greater number of feedback possibilities within the feedback-process and also allows for a greater number of feedback interventions. Based on the literature presented in this thesis and on the discussion regarding the model and the method, allowing for a significantly higher volume of feedback opportunities might be the element with the highest possibility of increased student gains in self-assessment. The

potential correlation between a higher number of feedback opportunities, as suggested in this thesis and increased learning gains in self-assessment, will need to be studied further to determine.

The methodology of this thesis has not revealed any real downsides in using a digital scaffolding. Neither was it a goal of the thesis to examine any potential negative effects a digital scaffold might have. Adverse effects might appear more prominently by approaching the research questions in this thesis with a different methodology or by conducting real-world instantiation of the model or the method with feedback from both teachers and students.

Wasson and Hansen (2014) found that some teachers that found it challenging to incorporate formative assessment, and another report (Norwegian Directorate for Education and Training, 2019) could possibly point to low adaptation of formative assessment in Norwegian schools. In exploring a model and a method for shorter peer-assessment interventions, some potential aids for teachers have emerged. The digital scaffold that is suggested is not just a support for the students. It can also potentially provide teachers with the necessary nodes to confidently carry out peer-assessment interventions and inspire them to develop their own peer-assessment interventions. Further study is needed to determine a potential correlation between having a tool for short peer-assessment and an increase in peer-assessment interventions in schools.

Group compositions can be a significant factor in peer-assessment interventions, as argued by Nicol (2014); Wang et al. (2015). The proposed method contained no special considerations for group compositions other than randomly grouping students. It can be argued that the number of possible feedback situations potentially can give students the benefits of different types of group composition. Since the feedback interventions are short, and each student gets feedback from multiple sources in each intervention, it could be argued that group composition possibly is of less importance in this type of peer assessment intervention.

Having a classroom discussion is in no way revolutionary. But the Discussion step does not stand isolated in the proposed model and method. The Discussion step builds upon the students' earlier actions. In the proposed method, each student works on three different sentences before they come to the Discussion step. The discussion is supposed to only revolve around these three sentences. This means that the student potentially has an investment in each of the discussion topics, and that could possibly enhance students' interest in participating in the discussion. Another potential effect from the earlier steps of the process is that the student has seen other students work. This could provide the student with enough confidence so that he would potentially partake in discussions in situations where he usually would not. Further study is necessary to determine if the proposed model or method could have an effect on student participation in a subsequent discussion.

The most significant changes to the model came from insight into how the scaffold manifested itself in the method. This information was supplemented with the literature on the potential benefits of digital tools in formative assessment. The digital scaffold provided

several possible avenues of encouraging an atmosphere of trust and respect for the students. An introductory video in the initial step could help assure the students of how the process works, and demonstrate that no one would be able to determine who had written what. Keeping the students anonymous is something that can potentially be a great asset in maintaining an atmosphere of trust and respect in further iterations of the method and instantiations. Allowing the teacher a full view of all student activity could also help deter unwanted student actions during the peer-assessment process. Conducting peer-assessment with the same amount of communication control would probably be difficult without a digital scaffold. As argued earlier in the thesis, threats to student's perception of safety have shown to diminish learning gains in peer-assessment situations (Hattie & Timperley, 2007; Nicol, 2014). A digital scaffold could, therefore, potentially be beneficial for shorter peer-assessment interventions by aiding in encouraging an atmosphere of trust and respect. The amount of effect this would have on peer and self-assessment situations will require further study to determine. It is important to note, however, that no model, method, or scaffold is inherently formative. Formative assessment occurs between people (Pachler et al., 2009). It is the social actions taken by the participants that can create formative possibilities for the students. So even a near-perfect tool for conducting peer-assessment will fail to yield positive learning outcomes if the participants actively sabotage the process. This is possibly also true for creating an atmosphere of trust and safety. The scaffold can not alone ensure a safe atmosphere.

An atmosphere of trust and safety is not the only potential benefit of having the students be anonymous. Anonymity can also require extra consideration from the students when they read and use feedback. They do not know the source of the feedback provided to them and are forced to evaluate the feedback critically. If the students knew the skill level of the student providing feedback to them, it could potentially lead to a situation where they trust the feedback because a specific student provided it.

In this chapter, there are avenues of discussion that point to ways short formative peer-assessment interventions, using a scaffolded approach, contribute to self-assessment for students. All of the avenues are followed by suggestions for further study. This is a result of the methodology chosen for this thesis. An exploratory design-method will not be able to determine to what degree each of the individual avenues might affect a student's skills in self-assessment. The study does, however, provide some valuable insights into different ways short formative peer-assessment interventions, using a scaffolded approach, contribute to self-assessment for students.

## 9. Literature

- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan Magazine*, 86(1), 8-21. doi:10.1177/003172170408600105
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. doi:10.1080/0969595980050102
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5-31. doi:10.1007/s11092-008-9068-5
- Burner, T. (2016). Formative assessment of writing in English as a foreign language. *Scandinavian Journal of Educational Research*, 60(6), 626-648. doi:10.1080/00313831.2015.1066430
- Campus-Inkrement. (2020). Campus Inkrement. Retrieved from <https://campus.inkrement.no/>
- Chang, C.-C., Tseng, K.-H., & Lou, S.-J. (2012). A Comparative Analysis of the Consistency and Difference among Teacher-Assessment, Student Self-Assessment and Peer-Assessment in a Web-Based Portfolio Assessment Environment for High School Students. *Computers & Education*, 58(1), 303-320. doi:10.1016/j.compedu.2011.08.005
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, 48(3), 409-426. doi:10.1016/j.compedu.2005.02.004
- Gamlem, S. M., & Smith, K. (2013). Student Perceptions of Classroom Feedback. *Assessment in Education: Principles, Policy & Practice*, 20(2), 150-169. doi:10.1080/0969594X.2012.749212
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. doi:10.3102/003465430298487
- Haugan, J., & Lysebo, M. (2015). Medstudentvurdering i matematikk og fysikk. *Uniped(04)*, 327-335.
- Haugan, J., & Lysebo, M. (2018). Hvorfor antall arbeidskrav bør reduseres ; om formativ vurdering i ingeniørutdanning. *Uniped [elektronisk ressurs] : tidsskrift for universitets- og høgskolepedagogikk*, 41(3), 347-360.
- Heidi, L. A. (2019). A Critical Review of Research on Student Self-Assessment. *Frontiers in Education*, 4. doi:10.3389/educ.2019.00087
- Hovardas, T., Tsivitanidou, O. E., & Xacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers and Education*, 71, 133-152.

Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2020). Does peer assessment promote student learning? A meta-analysis. *Assessment & Evaluation in Higher Education*, 45(2), 193-211. doi:10.1080/02602938.2019.1620679

Liang, M.-Y. (2010). Using synchronous online peer response groups in EFL writing: revision-related discourse.(Report). *Language, Learning & Technology*, 14(1), 45.

Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment: feedback for students with various thinking-styles. *Journal of Computer Assisted Learning*, 17(4), 420-432. doi:10.1046/j.0266-4909.2001.00198.x

Lise Vikan, S., & Trond, B. (2014). Vurdering i skolen. Utvikling av kompetanse og felleskap Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS). In: Unpublished.

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251-266. doi:10.1016/0167-9236(94)00041-2

Nicol, D. (2014). *Guiding Principles for Peer Review:Unlocking Learners' Evaluative Skills*: Edinburgh University Press.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. doi:10.1080/03075070600572090

Node.js. (2020). About Node.js. Retrieved from <https://nodejs.org/en/about/>

Norwegian Directorate for Education and Training. (2019). Erfaringer fra nasjonal satsing på vurdering for læring (2010-2018). Retrieved from <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/erfaringer-fra-nasjonal-satsing-pa-vurdering-for-laring-2010-2018/>

Norwegian Directorate for Education and Training. (2020). Engelsk (ENG01-04). Retrieved from <https://www.udir.no/lk20/eng01-04/kompetansemaal-og-vurdering/kv4>

Pachler, N., Mellar, H., Daly, C., Mor, Y., & William, D. (2009). Scoping a vision for formative e-assessment: a project report for JISC. In: WLE Centre and JISC.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74-98. doi:10.1016/j.edurev.2017.08.004

Reinholz, D. (2016). The assessment cycle: a model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, 41(2), 301-315. doi:10.1080/02602938.2015.1008982

Schleicher, A. (2018). OECD Learning Framework 2030 Retrieved from [https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf)

Seufert, S., Guggemos, J., & Sonderegger, S. (2019). *Learning Analytics in Higher Education using Peer-feedback and Self-assessment: Use Case of an Academic Writing Course*.

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*.

St.meld. nr. 28 (2015-2016). Fag – Fordypning – Forståelse — En fornyelse av Kunnskapsløftet. *Kunnskapsdepartementet*. Retrieved from <https://www.regjeringen.no/no/dokumenter/meld.-st.-28-20152016/id2483955/>

Tenório, T., Bittencourt, I. I., Isotani, S., Pedro, A., & Ospina, P. (2016). A gamified peer assessment model for on-line learning environments in a competitive context. *Computers in Human Behavior*, 64, 247-263. doi:10.1016/j.chb.2016.06.049

Walker, J. (2006). Å lære å gi og motta konstruktiv kritikk gjennom medstudentvurdering. In: Program for læringsforskning, Universitetet i Bergen.

Wang, J. H., Chen, S. Y., Chang, B., & Chan, T. W. (2015). From integrative to game-based integrative peer response: high ability versus low ability. n/a-n/a. doi:10.1111/jcal.12125

Wasson, B., & Hansen, C. (2014). Making Use of ICT: Glimpses from Norwegian Teacher Practices. *Nordic Journal of Digital Literacy*(01), 44-65.

Wisniewski, B., Zierer, K., Hattie, J., & Wisniewski, B. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in psychology*, 10, 3087-3087. doi:10.3389/fpsyg.2019.03087

Yurdabakan, İ., Gelbal, S., Olkun, S., & Aydoğan, M. (2012). The Effect of Co- and Peer Assessment Training on Self-Assessment Skills of Teacher Trainees. *Eğitim ve Bilim*, 37(163).

## 10. Figures and Tables

Figure 1- A single task with input .....	30
Figure 2 - Start screen - Teachers perspective .....	36
Figure 3 - Student overview - Teachers perspective .....	37
Figure 4 - Setting goals - Class perspective .....	38
Figure 5 - First task - Student perspective.....	38
Figure 6 - Exposure to different tasks if all tasks are similar.....	39
Figure 7- Exposure to different tasks if all tasks are unique .....	39
Figure 8 - Workflow of three students .....	40
Figure 9 - Giving Feedback - Student perspective.....	40
Figure 10 - Feedback on feedback - Student perspective .....	41
Figure 11 - All contributions - Class perspective .....	41
Figure 12 - Final review of the translation - Student perspective.....	42
Table 1 - Patterns for e-assessment (Pachler et al., 2009) .....	14
Table 2 - Proposed model .....	23
Table 3 - Revised model .....	43