Doctoral theses at NTNU, 2022:177

Gabrielle Hansen

Use of Feedback in a Higher Education Learning Context

"It's not like it's too late to learn. And this is an important message, you can learn even if you have misunderstood."

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Social and Educational Sciences
Department of Teacher Education

NTNU
Norwegian University of
Science and Technology

Gabrielle Hansen

Use of Feedback in a Higher Education Learning Context

"It's not like it's too late to learn. And this is an important message, you can learn even if you have misunderstood."

Thesis for the Degree of Philosophiae Doctor

Trondheim, June 2022

Norwegian University of Science and Technology
Faculty of Social and Educational Sciences
Department of Teacher Education

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Social and Educational Sciences
Department of Teacher Education

ISBN 978-82-326-5193-1 (printed ver.)
ISBN 978-82-326-6175-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2022:177

Printed by NTNU Grafisk senter

Contents

PART 1: EXTENDED ABSTRACT

Acknowledgements

It has been a few years since I started this process in the spring of 2012. The significance of use of feedback in a higher education learning context, however, has not diminished over the years, perhaps rather the opposite. I hope the findings in my thesis can contribute to teachers creating their own feedback practice in a more formative way.

The road to submission has been long due to both private reasons, such as becoming a mother, and professional choices I have made along the way, such as accepting wonderful opportunities to both develop and run other key projects at NTNU. The downside, of course, is that it has taken me a long time to finish. So, to put it mildly, I am incredibly relieved and truly happy to finally submit my work.

This thesis has been created in meetings between people. I would therefore like to take this opportunity to thank several people now that I have reached completion.

First and foremost, I would like to thank the teachers who were involved in this research work with me. Your efforts, courage and insight are still a great inspiration to me.

Thanks also to my supervisors Alex Strømme (NTNU) and Thomas Dahl (NTNU). Thank you for putting up with me through all these years and for being my supervisors.

I would also like to thank my "statistical partner in crime", Regine Ringdal. Because of you, the work on Sub-study 3, mapping the students' achievement-goal patterns was a fantastic journey. Thank you so much for your hard work on the statistical analysis, for all the wonderful and constructive conversations, and of course, I must thank you for all the chocolate and laughter.

My dear friend Cicilie also deserves a big thank you. Your support over the years has meant more than words can say.

Finally, I would like to thank my supportive family. My dear children, Mona, Even and Iver (and the fourth one on the way), who have eagerly counted down to the submission date and cheered me on, and my better half, Audun. This submission would simply not have been possible without you. Thank you.

Summary

The overall theme addressed in this thesis is use of feedback in a higher education learning context. The main research question to be addressed is *how can educators develop their own feedback practice in a more formative way?*

The background for this objective is relatively simple and straightforward: feedback represents a major dilemma within higher education today as there is a relatively large gap between its theoretical potential and the actual practices (O`Donovan et al., 2016), and it is necessary to scale up educational change (Carless, 2017). Two important consequences are the repeated failure to engage students in their own feedback processes (Broadbent et al., 2018), and thus undermine the importance of increasing student involvement and understanding through interactive dialogues (Carless, 2017), and the second, yet just as important, an increase in motivational patterns and self-beliefs that undermines students' learning (Dweck and Master, 2012; Forsythe and Johnson, 2017). However, quite the opposite to tertiary education research, school research vividly demonstrates that with support, teachers can transform research findings into new and effective assessment and feedback practices (Black et al., 2003; Pedder and James, 2012).

This thesis has been based on a two-and-a-half-year intervention study using a qualitative approach within an interactive action research framework. Through a mutual collaborative process, inspired by several learning practices from school teaching, up to five mathematics teachers and the author of this thesis, used one of the most influential accounts of feedback in higher education, namely the model created by Nicol and MacFarlane-Dick (2006), which posits *"seven principles of good feedback practice"*, as an intervention tool to create a formative assessment practice in mathematics that is characterised by active student participation, self-feedback, reflection, peer dialogue and student-teacher dialogue. Through this intervention process the following items were examined:

- how the teachers experienced changing their own practice (Sub-study 1),
- the degree of compliance between the teachers' intentions in using the "seven principles of good feedback practice" and the students' experience of them in practice (Sub-study 2), and
- the student's achievement-goal patterns within a formative assessment practice in mathematics (Sub-study 3).

The results from the three studies shows that development and change in practice are directly connected to one's personal development as a teacher (Sub-study 1). The findings

also reveal that compliance between student and teacher understandings, requires dialogue and active efforts and participation by both parties (Sub-study 2). Finally, the results argue for the importance of teachers` effort in relation to the development of students' achievement-goal patterns, and for maintaining achievement-goal stability (Sub-study 3).

This thesis shows that work methods for educators and a concrete feedback design can help the present practice in higher education to come closer to the current empirical and theoretical understanding of feedback and formative assessment. The conclusion is that educators can develop their own practice in a more formative direction by being responsible for creating a social learning environment where feedback is created through active student participation in such mastering-oriented and dialogic learning activities as dialogues between students, teacher-student dialogues, and last but not least inner dialogues through reflection.

List of articles

Article 1

Hansen, G. (manuscript). Formative assessment from paper to practice: teachers` experience of change. Empirical lessons from school literature practiced in higher education.

Article 2

Hansen, G. (2020). Formative assessment as a collaborative act. Teachers intention and students experience: Two sides of the same coin, or? *Studies in Educational Evaluation,* 66, 1-10.

Article 3

Hansen, G., & Ringdal, R. (2018). Formative assessment as a future step in maintaining the mastery-approach and performance-avoidance goal stability. *Studies in Educational Evaluation,* 56, 59-70.

List of tables

List of figures

1. INTRODUCTION TO THE THESIS

The overarching theme of this thesis is use of feedback in a higher education learning context. The background for this is relatively simple and straightforward: feedback still represents a major dilemma within higher education today as there is a relatively large gap between its theoretical potential and the actual practices (O`Donovan et al., 2016). Even though much effort is being invested in creating better practices, this repeatedly results in very little practical difference (Sadler, 2016). Draper and Nicol (2013) put this more succinctly:

> In education, new research findings or even best practice tips spread slowly, if at all (p. 194).

On the other hand, it is important to point out that a wide range of research activity in higher education has been influenced by both feedback and formative assessment principles over the last twenty years or so, and formative feedback has become well entrenched as part of higher education pedagogy (Carless, 2017). Consistent with the assumptions underpinning the second Bologna decade up to 2020, where student-centred learning has been identified as a higher education priority area (Leuven/Louvain-la-Neuve Communiqué, 2009, Bucharest Communiqué, 2012; Yerevan Communiqué, 2015), effective assessment feedback design has been established and student-centred feedback has become an increasingly important aspect of higher education learning and teaching strategies (Brown, 2010; Hoidn, 2016). The main challenge, however, is the implementation of such designs and areas, which has been demonstrably more problematic (Evans, 2013; Hoidn, 2016; Broadbent et al., 2018). In other words, increased interest in feedback does not necessarily lead to widespread implementation at course levels, and it is necessary to scale up educational change (Carless, 2017; Boud et al., 2018; Dawson et al., 2019).

To that end, this thesis is based on findings from a two-and-a-half-year intervention study using a qualitative approach within an interactive action research framework. In this study, a collaboration team consisting of up to five mathematics teachers and the educational researcher, referred to in the following as "author", was established. The team worked deliberately with empirical findings as an evidence base for introducing various changes into the classroom, and used learning practices from school literature, such as teacher support and inquiry-based collaboration, to implement potential and sustainable changes. As an evidence base for developing a student-centred feedback practice, the team used one of the most influential accounts of feedback in higher education, the model created by Nicol and MacFarlane-Dick (2006) positing *"seven principles of good feedback practice"* as a pedagogical framework for creating a dialogic and learning-oriented feedback and assessment

context characterised by active student participation, self-feedback, reflection, peer dialogue and student-teacher dialogue.

1.1 Initial Empirical Review

The early 2010s marked a shift in how feedback was positioned within the literature, with understandings of feedback moving from something "given" to students towards feedback as a process in which students have an active role to play (Dawson et al., 2019). Self-regulative capacities that promote sustainable and formative feedback practices are considered key aspects of the development of quality student learning (Hattie and Timperley, 2007; Carless et al., 2011; Boud et al., 2018). This means that learning from feedback is not about transferring knowledge from the teacher to the student, rather it is constructed in a process of social interaction (Dunworth and Sanchez, 2016; Carless, 2019). In other, and more theoretical, words, it is *dialogic* (Nicol, 2010), wherein meaning and understanding, and thus learning, are created through interaction (Vygotsky, 1978; Bakhtin, 1981; Dysthe, 1996, 2008).

Unfortunately, there is a fundamental failing in current practices: the social constructivist processes are generally not applied within the dominant everyday discourse of testing and marks, and there continues to be little emphasis on the agency and activity of students in feedback processes (O'Donovan et al., 2016; Carless, 2017; Boud et al., 2018). Two important consequences are the repeated failure to engage students in their own feedback processes (Broadbent et al., 2018), and second, yet just as important, in increase in an achievement-goal pattern that undermines students learning (Senko and Harackiewicz, 2005; Dweck and Master, 2012). The literature thus has moved forward in how it understands feedback and formative assessment, but it is not so clear if those involved have been brought along with it (Dawson et al., 2019).

1.1.1 The crucial premise: teachers' professional learning

Looking into some of the barriers to realising the theoretical benefits shown in the research literature can help to shed light on this rather modest influence of educational research at the tertiary level. For starters, misconceptions of the meaning of formative feedback, conceptions of its value and time to carry it out, lack of incentives and motivation to engage in such practices, as well as beliefs about teaching and learning can impede integration of formative feedback into classroom practice (DeLuca et al., 2012; Carless, 2013). Moreover, even if many teachers might read literature on teaching and learning in

general, or the available literature in their discipline, they often have little experience in translating educational ideas into effective teaching practices (Nicol and Draper, 2009). This means that a key challenge in higher education is how to support teachers so they can make informed changes to their own practices (Draper and Nicol, 2013).

Quite the opposite to tertiary education research, primary and secondary school research vividly demonstrates that with support, teachers can transform research findings into new and effective feedback and assessment practices (Black et al., 2003; Dekker and Feijs, 2005; Pedder et al., 2005; Leahy and Wiliam, 2012). It is therefore important to ask what the researcher and academic staffs in higher education can learn from the lower educational level when it comes to feedback and the facilitation of actual changes.

One of the most important lessons from school research is that teachers' *professional learning* is a crucial factor (Thompson and Wiliam, 2008). School research provides powerful evidence that giving teachers opportunities to engage in collaborative classroom-focused inquiry is an important instrument that can encourage them to work with principles of formative feedback and develop new practices (Pedder and James, 2012). Importantly, however, applying certain principles on their own is not necessarily an adequate step for facilitating actual changes in the classroom. The teachers also need to actually believe in them if they are to bring about changes in practice (James et al., 2007).

All in all, what we can gain from lower educational research, which is often lacking in higher education literature, is concrete learning practices that transform research findings into new practices beyond the focus on best practice tips.

1.2 The Research Questions of this Thesis

The theme of this thesis is use of feedback in a higher education learning context. The studies aim to highlight key aspects of feedback practices which today is well established within the research literature, but where there is a way to go before it is established as a regular part of the wider formative assessment practices in higher education, including interactive dialogues, motivational patterns that promote learning and professional learning of educators. The overarching research question addressed in this thesis is:

How can educators develop their own feedback practice in a more formative way?

This main question has been further divided into three sub-questions, with related research questions, corresponding to three sub-studies that will be presented as separate articles in this thesis. The three sub-studies examine:

1. How teachers experience their own change process towards creating a formative assessment practice
    a. What are the teachers' foremost experiences of their own change process?
2. The degree of alignment between the teachers' intentions in using the "seven principles of good feedback practice" and the students' perceptions of them in practice
    a. To what extent are the teachers' beliefs about what they are doing in a dialogic formative assessment concordant with how the students experience this?
3. The achievement goal patterns pursued by the students in a formative assessment practice in mathematics
    a. Which achievement goal patterns do the students pursue in a formative assessment practice in mathematics?
    b. How do students perceive a formative assessment practice in mathematics?

The two-and-a-half-year intervention study comprises three sub-studies. These will be referred to as Sub-study 1 (2012-2013), Sub-study 2 (2012-2013) and Sub-study 3 (2013-2014). Table 1 provides an overview of the three sub-studies that are used to answer the main research question addressed in this thesis.

1.2.1 A matter of relevance and legitimacy

As can be read from the table below, the research work associated with this thesis was carried out from 2012 to 2014, seven years ago. Much has happened in higher education since then, and as mentioned above, principles of feedback have been given increasing priority. This raises the question of whether the question formulation of this thesis, and the findings from the included studies, are just as relevant today as they presumably once were.

Both the national and European contexts (which will be described in detail in the next chapter) surrounding the main research question being addressed in this thesis, and the empirical landscape in which it operates (which was briefly presented above and will be further elaborated on in Chapter 3), show that in spite of important advances in feedback theory and practice over the past decades, it is apparent that change does not necessarily come easily (Boud et al., 2016). This can be interpreted to mean that the main question of this thesis is not outdated. On the contrary, the main question, the research work and results related to this thesis are relevant and point to key challenges in today's higher education, and can hopefully contribute constructively to an overall understanding of use of feedback in a higher education learning context.

Table 1. Overview of the three sub-studies.

	Research questions	**Methods**	**Participants**	**Timeline**
Sub-study 1 (presented in Article 1)	What are the teachers' foremost experiences of their own change process?	Semi-structured interviews	Two mathematics teachers	Spring 2012-spring 2013
Sub-study 2 (presented in Article 2)	To what extent are the teachers' beliefs about what they are doing in a dialogic formative assessment concordant with how the students experience it?	Semi-structured interviews Focus-group interviews	Two mathematics teachers Students from two preparatory engineering classes	Fall 2012 - spring 2013 Spring 2013
Sub-study 3 (presented in Article 3)	Which achievement goal patterns do the students pursue in a formative assessment practice in mathematics?	Questionnaire Focus-group interviews	Students from five preparatory engineering classes	Fall 2013 - spring 2014 Spring 2014

1.3 Structure of the Thesis

Chapter two, the *National Context*, presents a contextual and historical review of the place feedback and formative assessment has held and currently holds in Norwegian higher education, and aims to frame the research presented in this thesis in a broader national and European context.

Chapter three, *Literature Review*, provides an overview of the concept of feedback in higher education. The chapter provides insight into how the current research literature understands feedback and its potential for learning and development. It also looks into important trends in the literature, concrete formative feedback practices, which challenges we are facing in higher education today, how formative feedback practices can be connected to other important research areas, such as motivational mindframes, and lastly, what we can learn from the lower educational levels in terms of facilitating actual changes in practice.

Chapter four, *Theoretical Groundings*, places the research work in a wider learning-theory landscape. More concretely, this chapter presents a theory of formative assessment as regulation of learning, a sociocultural view of learning that comprises the expanded dialogue concept, reflection and mediating tools, and finally, a social-cognitive approach to motivation through theories on student intelligence.

Chapter five presents the *Context of the Research* and provides a brief overview of the context of the thesis, covering the mutual collaborative process between the teachers and

"author", the point of departure for the work, the work methods and a comprehensive description of how the seven principles were implemented in practice.

In Chapter six, *Research Methods,* the methodologies employed in this intervention study, using a qualitative approach within an interactive action research framework, are presented. This chapter provides a complete overview of the intervention and research process, including the research methods, data collection and analytical approaches, including the preliminary statistical examinations and their considerations and results. Finally, this chapter concludes with a discussion of quality criteria in research and a critical reflection on these.

Chapter seven, *Findings,* provides a brief description of the appended papers, the research questions and the findings from each.

The eighth and final chapter, *Discussion and Conclusions*, focuses on how the three articles presented in this thesis can contribute to a broader and more comprehensive understanding of feedback in a higher education learning context. The findings are discussed in terms of the main research question.

2. THE NATIONAL CONTEXT

Before moving on to a deeper empirical and theoretical examination of the concept of feedback in higher education, it is in its place to give a brief contextual and historical review of the place feedback and formative assessment has held and currently holds in Norwegian higher education, especially when it comes to mathematics and the natural sciences, which constitute the academic context for the research in this thesis. Below is an overview of some of the educational changes, and current evaluations, in relation to the development of feedback and assessment practices in Norwegian higher education. These will also be considered in relation to a broader European higher education context.

This chapter is both important and necessary as it presents the national situation surrounding this thesis that is very much the same as what the international research literature has to say on feedback, and thus legitimizes the empirical choices and use of the literature in the subsequent chapters. In the next chapter, key aspects of this review will be illuminated in relation to the body of research literature, elaborating on important aspects in relation to previous research.

2.1 Formative Assessment in a Norwegian Context: from Control and Efficiency to Common Culture

The development of feedback and assessment practices in the Norwegian education system must be considered in light of a number of major modifications and reforms that have changed the approach to education significantly over the last 15 to 20 years (Damsgaard, 2019). Important national, and not least international trends, such as the Bologna process, have had great impact on the development of Norwegian higher education and led to changes. In 1999, 29 European ministers of education, Norway included, met in Bologna, Italy, to formulate a declaration on higher education. The original goal was to create a common European area for higher education within 2010. But the work in this area, *the European Higher Education Area* (EHEA), has been continued into 2020 (Norwegian *Ministry of Education and Research*, 2017). For Norway, and also of interest for this thesis, these efforts resulted in the introduction of the Norwegian quality reform, the Report to the Storting no. 27 (2000-2001).

2.1.1 The Norwegian quality reform: focused on feedback and guidance

The goals of the quality reform were established by the Mjøs Committee in 2000 (NOU 2000:14) and the reform was introduced in higher education in 2003. The overarching

purpose of the reform was to strengthen the quality of all education so that more students could complete their studies (the Office of the Auditor General of Norway, 2015), and included higher intensity in the teaching, a structure for internal control, a new form of funding, and the introduction of study points and a new grading system using letter grades (Damsgaard, 2019). Beyond this, the reform also included clear expectations about changes, where ways of teaching were to be introduced that would require students to be more active, with closer follow-up, more guidance, and last but not least, the introduction of new forms of examinations and assessment (*Ministry of Education and Research*, 2008).

The expectations for the new forms of assessment may be considered in the light of the quite harsh criticism of Norwegian higher education, as expressed by the OECD, where Norwegian universities were described as "… research institutions conducting exams" (OECD, 1997). The report criticised Norway for relying too much on final examinations and using them as the only assessment of student work. The OECD report concludes that there is a need to focus more on learning where the students receive feedback throughout their studies, and that closer ties should be forged between teachers and students.

2.1.2 Educational consequences and changes in the assessment form

A comprehensive evaluation of the quality reform was published early in 2007. The Report to the Storting no. 7 (2007–2008) stated that it was difficult to measure whether the reform actually had increased quality, and that it was therefore impossible to ascertain whether the overriding objective of the reform had been satisfied. However, the report stated that the quality reform had led to closer follow-up of students, and closer follow-up in this context generally meant giving feedback and guidance relating to written tasks, and that the students were mostly satisfied with this (*Ministry of Education and Research*, 2008). What changed most in relation to assessment was the introduction of portfolio assessment and an increase in the number of small tests throughout the studies that would count toward the final grade. However, the engineering and natural science studies had a noticeably lower number of changes (*Ministry of Education and Research*, 2008).

10 years later: are we there yet? The quality reform was evaluated again ten years after it was implemented. The Office of the Auditor General of Norway then stated that the reform in the course of these years had not successfully increased completions or reduced the number of non-completions (Office of the Auditor General of Norway, 2015). The report points out, furthermore, that the need to take a critical look at teaching and assessment forms in Norwegian higher education remains.

2.1.3 Students exercise their voice: Studiebarometeret [the study barometer], SHoT-undersøkelsen [the Students' Health and well-being survey] and Kandidatundersøkelsen [the graduates survey]:

The Quality Reform also led to the establishment of NOKUT (Norwegian Agency for Quality Assurance in Education), as proposed by the Mjøs Committee, and UHR (Universities Norway), to improve the quality assurance and external control of education in Norway (Ministry of Church, Education and Research, 2001). For this thesis it is most interesting to focus on the national student survey, *Studiebarometeret*, which is conducted annually by the Norwegian Agency for Quality Assurance in Education on assignment for the *Ministry of Education and Research*. Before the Study Barometer was introduced, there was little empirical material on how Norwegian students in higher education actually perceived the quality of feedback, assessments and follow-up (Hamberg et al., 2016).

The finding that may have attracted the most attention, from the initial survey in 2013, is that of all the surveyed areas, the students show the lowest satisfaction level for the area of feedback and follow-up from the academic staff (Hamberg et al., 2016; Wiggen et al., 2020). The survey from 2019, comprising almost 32 000 students, shows that only approximately half the students are satisfied with the ability of their teachers and professors to give constructive feedback on their work and with academic discussions providing guidance from instructors and professors, while 30 per cent are little satisfied or dissatisfied with the amount of feedback they receive on their work from professors and teachers (Wiggen et al., 2020).

These findings do not, on the other hand, mean that students in Norway are generally unhappy with the overall quality of their studies, rather the opposite is true (Wiggen et al., 2020). Another study conducted by the Norwegian Agency for Quality Assurance in Education in the spring of 2015 (Hamberg et al., 2016) concluded that the students were dissatisfied with the feedback and guidance they received because they received so little of both. Moreover, the students found that the feedback and guidance they received did not contribute significantly to their learning outcome. On the other hand, the students also had low expectations when it came to receiving constructive feedback and guidance, and also lacked understanding of the importance of feedback and guidance, so these factors were not considered when they assessed their overall level of satisfaction.

The fact that assessment and feedback may be a challenge for higher education has also been pointed out by the Student Health and Well-being Survey. In 2018, of the more than 50 000 students who answered the survey, only 29 per cent of the students stated they were very or quite satisfied with the feedback they receive on their learning (Knapstad et al, 2018).

The survey of graduates undertaken by the Nordic Institute for Studies in Innovation, Research and Education, where Master's degree students were questioned six months after graduation had a similar finding. In the survey from 2017, the graduates assessed the overall academic content, teaching quality and feedback from their teachers as generally positive, but their evaluation of the quality of the feedback and guidance was low (Nesje and Støren 2018).

2.1.4 What about the scientific staff?

The Norwegian Agency for Quality Assurance in Education expanded the Study Barometer by conducting a similar national survey among higher-education teachers in Norway (Lid et al., 2018). Similar to the Study Barometer, this survey is a stage in the efforts to raise quality in higher education. Conducted initially in 2017, the intention is to hold the survey every third year, and it will be aimed at all those who have taught on the Bachelor or Master degree levels over the last two years (Damsgaard, 2019).

The findings of the Teaching Staff Survey (Lid et al., 2018) show that as many as 77 per cent of the educators believe that they attach importance to giving the students comments and feedback. Among the students, as described above, slightly different perceptions prevail about feedback and guidance. According to Lid et al., (2018), it is useful to bear in mind that while the students are asked to assess these aspects of the education based on their experiences, the teaching staff have been asked about what they focus on, which is not necessarily understood as what they actually do.

2.1.5 Today's status: Knowledge evolves in academic environments

The last report to be included in this review, Report to the Storting no. 16, *Kultur for kvalitet i høyere utdanning* [Quality Culture in Higher Education], was published in 2016. Even if this report was not available during the data-collection period and initial analyses of the data in this thesis, it is still important to include it to obtain a more concrete, overarching and valid image of expectations when it comes to assessment and feedback in Norwegian higher education today.

Already in the preface to the Report, the tone is set under the heading "Knowledge evolves in academic environments". Knowledge is thus presented as something that evolves and develops in the interaction between educators and students, through dialogue, discussion and feedback (*Ministry of Education and Research*, 2017). The focus is, in other words, shifted from seeing quality as structure, numbers and quantities, to developing a common culture for raising the quality of education (Damsgaard, 2019).

Even if the report underlines that there are many good practices in Norwegian higher education, it also clearly points to areas where improvements could be made. The report finds, for example, that there is too little focus on developing the educational competence of the staff, and that the staff therefore are not receiving sufficient feedback on their own teaching and assessment practice, which means that the work on raising quality is not stimulated (Damsgaard, 2019). It is also pointed out that the academic faculties should give more and better feedback to the students throughout their studies so they have a better idea of how they are doing. It is argued that assessment and feedback are vital for raising student awareness, stimulating learning and development and helping students develop the ability to follow up, evaluate and regulate their own learning (*Ministry of Education and Research,* 2017).

Second Bologna decade. The quality report's focus on more active student participation is consequently also in accordance with the broader European guidelines. In the context of the second Bologna Decade up to 2020, student-centred learning (SCL) has been defined as an approach that replaces purely transmissive models of education (Hoidn, 2016; EUA, 2019). SCL conveys the notion of students as constructivist learners and active participants with shared responsibilities for outcomes (EUA, 2010). The central function of student-centred learning for the development of high-quality education is also highlighted by Standard 1.3 in the 2015 Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG), according to which universities "should ensure that the programmes are delivered in a way that encourages students to take an active role in creating the learning process, and that the assessment of students reflects this approach" (ESG, 2015; EUA, 2019).

This review can be concluded by coming back to the start, where it was found that this overview is important for framing the research here in a broader national and European context. By describing the national context for this thesis, the need for and significance of the research focus and findings are elucidated and clarified. Accordingly, this contextual review will also serve as a backdrop for the further empirical exploration in the following chapter.

3. LITERATURE REVIEW

Grades often tell the student 'the work is over'. We must not confuse grading with feedback. (Hattie and Clark, 2019, p. 2)

3.1 The Concept of Feedback in Higher Education

Feedback is a prominent topic in current higher education research (Molloy et al., 2020). Whilst feedback is recognised as a core component of a learning process (Hattie and Timperley, 2007), several national surveys across several countries consistently rate feedback at the bottom of the student-satisfaction scale (Carless, 2017; Pitt and Norton, 2017; Wiggen et al., 2020). The natural response to this predicament has been to attempt to improve the quality of the feedback information provided by teachers, in particular, its promptness, level of detail, clarity, structure and relevance (Nicol et al., 2014). Well-meaning as these interventions are, there is little evidence that they have had any effect (ibid), and furthermore, they continue to amplify the dominant view of feedback as information, a teacher-controlled process, with the accompanying assumption that when delivered well, it will automatically be absorbed by the learner (Molloy et al., 2019).

To elaborate, today marking or grading student work and then providing feedback comments are still two of the most common undertakings within universities around the world (Winstone and Boud, 2020). These practices create further challenges in terms of low student satisfaction (students wanting more feedback that is less generic and more personalised, constructive and interactive), student engagement (many students do not use the feedback they receive) and staff workload (providing feedback requires time, individual capacity and scalability of feedback practices) (Henderson et al., 2019). The latter point is especially problematic given the rising student numbers in higher education (Nicol et al., 2014). This means that for both educators and students feedback processes as they are commonly enacted in higher education are often neither productive nor satisfying. Finally, as seen, for example, in theories on intelligence and achievement goals, traditional performance feedback can lead to less favourable motivational patterns that actually undermine students' learning (Senko and Harackiewicz, 2005; Dweck and Master, 2012).

This state of affairs has now stimulated scholars and researchers to re-examine feedback in higher education when it comes to how it is conceptualised, its consequences and how this translates into actual classroom practices (Sadler, 2010). Underpinning this re-examination is the important recognition that if feedback processes are to improve learning outcome, we must move away from an exclusive focus on a teacher-controlled process and

what teachers do to initiate feedback (input), and rather move towards a learning-focused process where students are the main actors (Boud and Molloy, 2013). Thus, this shift in how feedback is understood places emphasis on many more features of feedback than just the provision of 'hopefully useful' comments from teachers to students (Henderson et al., 2019). The conceptualisations of feedback currently prominent in the research literature see the entire feedback process as controlled by the student rather than the teacher, involving a multitude of players and necessarily involving the student utilising information to bring about change (Dawson et al., 2019).

3.2 A New Feedback Paradigm: Feedback Is a Social Practice

According to Winstone and Carless (2020), the different ways of thinking about feedback in higher education, the old (focused on input, the provision of information or comments given to students), and the new (adopting a more learning-oriented approach), represent two paradigms. Rather than the teachers providing information and the students being positioned somewhat passively, the new feedback paradigm aims for more of a *partnership* between teachers and students. This is because feedback as a delivered message overlooks what mediates the exchange: the interaction between sender and receiver, and the network of social relations which shape the interpretation processes and open for the development of shared understanding (Price et al., 2013). Moreover, this partnership envisages a key teacher role of designing feedback practices to facilitate student participation and support them so they understand how to engage in productive feedback interactions (Winstone and Carless, 2020).

3.2.1 The notion of dialogic feedback, sustainability and self-regulating abilities

In an attempt to encourage more interaction with feedback, *dialogue* is proposed as the core of new ways of thinking about feedback in higher education (Winstone and Carless, 2020). It is important to understand that bringing dialogue into the feedback process means much more than having individual discussions in a face-to-face setting; this process is based on student-generated dialogues, peer-to-peer feedback, student self-feedback and the development of assessment literacy (e.g. Carless and Boud, 2018). It can occur during plenary interactive teaching or through generic feedback, and it can occur through the medium of technology (Winstone and Carless, 2020).

Dialogic feedback: a definition. Bearing the above and the work of Carless et al. (2011) in mind, this thesis therefore sees feedback as "all dialogue that supports learning in

both formal and informal situations" (Askew and Logde, 2000). More specifically, dialogic feedback promotes interactive exchanges where interpretations are shared, meaning negotiated and expectations clarified, and it aims to provide opportunities for students to interact on notions of quality and standards in the discipline (Carless et al., 2011). These types of dialogue are important because they allow students to make sense of new knowledge they encounter and help them to develop new conceptual understandings.

Given that students often seem to have a rather different view on feedback than their teachers, and that teachers often perceive their feedback as more useful than their students do (Maclellan, 2001; Carless, 2006; Mulliner and Tucker, 2017), an important role of dialogue is to narrow the gap between the teaching staff's and students' expectations and perceptions of feedback through communication and negotiation (Winstone and Carless, 2020).

Nevertheless, the core argument behind dialogic feedback is the need to let go of the 'transference' feedback model (Blair et al., 2014) in order to facilitate the *sustainability* of the feedback (Johnson and Molloy, 2018). Dialogue is a key to sustainable feedback in that it emphasises the students' role in making sense of feedback and using it to develop their own *self-regulating capacities*, including identifying learning goals, selecting effective learning strategies, monitoring progress and refining strategies accordingly (ibid). Of special importance is the practice whereby students judge the quality of their own and others' work (Sadler, 2010), thus developing their capacity for evaluative judgement (Ajjawi et al., 2018) and making them less reliant on external feedback. This will then better equip them to generate feedback for themselves (Winstone and Carless, 2020). Sustainable feedback is thus co-constructed by teachers and students, with an emphasis on the students' engagement with feedback as part of their development as self-regulative learners (Carless, 2013).

3.2.2 Seven principles of good feedback practice

Of key relevance to the notion of dialogue and sustainable feedback is the influential model of Nicol and Macfarlane-Dick (2006), where the aim is to shift the focus to see students as having a proactive rather than a passive role in generating and using feedback. Based on an analysis of extensive research material on formative assessment and feedback, Nicol and Macfarlane-Dick identified the following seven principles of good feedback practice:

Good feedback practice:

1. helps clarify what good performance is;
2. facilitates the development of self-assessment in learning;

3. delivers high-quality information to students about their learning;
4. encourages teacher-and-peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape teaching.

These are familiar principles where their underlying value is supported by a substantial amount of research and they are all defined in terms of their contribution to the development of self-regulatory learning (Nicol and MacFarlane-Dick, 2006; Panadero et al., 2018).

3.3 A Formative Framework

Finally, the growing focus on notions of feedback in which students are positioned as active players rather than recipients of information complies with the very essence of a broader assessment context, formative assessment. In other words, a key element of an effective feedback practice is that it has a formative assessment framework (Hattie and Clark, 2019).

Formative assessment has gained in prominence in recent years as it focuses on student learning rather than merely judging levels of performance (Panadero et al., 2018), of which the latter resembles the more dominant assessment traditions we still see today, the summative measurement tradition (Boud et al., 2018). More specifically, formative assessment is a student-centred measurement model that is associated with meaningful feedback that is used for guiding instruction, enhancing student learning and developing self-regulated learning practices (Black and Wiliam, 1998; Broadbent et al., 2018).

Formative assessment: a definition. This thesis chooses to conceptualise formative assessment in accordance with the definition provided by Black and Wiliam (2009), drawing both on their earlier definitions (Black and Wiliam 1998) and the definition from the Assessment Reform Group (ARG 2002):

> 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 (p. 9).

According to Black and Wiliam (2009), this means that formative assessment is, in essence, concerned with the creation of, and capitalisation on '*moments of contingency*' in the

instruction for the purpose of the regulation of learning processes (2009, p. 10). This means that the teacher's task moves away from delivering learning to the student and towards the *creation of situations* in which students can learn (Wiliam and Thompson, 2007). This definition and its theoretical groundings and implications will be presented and elaborated on in the theoretical review in Chapter 4.

Wiliam and Thompson (2007) see formative assessment in relation to the function it actually serves (as opposed to a particular assessment or even the purpose of an assessment). Assessment is thus formative to the extent that information from it is fed back within the system and actually used to improve performance in some way. Therefore, in order to be *formative*, feedback needs to contain an implicit or explicit plan for future action. Formative feedback answers three major questions: where am I; where am I going next and how do I get there? (Hattie and Timperley, 2007; Black and Wiliam, 2009).

Overall, a considerable amount of research evidence shows that feedback in the context of formative assessment has a strong impact on student learning (Hattie and Timperley, 2007; Evans, 2013). Black and Wiliam's substantial review of the research literature, from 1998, revealed that formative assessment 'works' because it effectively promotes student learning across a wide range of educational settings. Formative feedback can enable students to judge the quality of what they are producing and also monitor themselves during the act of production (Sadler, 1998), and it can empower students to evolve from being dependent on teacher-led feedback to being able to generate their own feedback on learning and progression, and become more self-regulated in their own learning process (Boud et al., 2018).

3.4 Formative Feedback Practices: Examples and Research Evidence

Among the practices of formative assessment that emphasise sustainable feedback, dialogue and increased student involvement, self-feedback and peer feedback play an important role. Even though these practices could be summative, research evidence strongly suggests that they are most beneficial for achievement, student perceptions and self-regulated learning when they are used formatively and supported by training (Andrade, 2019; Panadero et al., 2019).

3.4.1 Self-feedback and reflection

Self-feedback can be defined as the implementation of self-assessment in ways that generate feedback information and processes for the students' own purposes (Panadero et al.,

2019). In other words, the purpose of self-feedback is to inform adjustments to processes and products that improve learning (Andrade, 2019). Instead of using self-assessment for purely grading purposes, as the former practice has been (Panadero et al., 2016a), this learning-oriented approach emphasises the importance of using self-assessment for formative purposes.

One technical challenge that arises, however, involves establishing valid criteria to judge whether or not student self-feedback is accurate (Falchikov and Boud, 1989). However, to get closer to effective self-feedback for student learning, content accuracy is argued to be more important than scoring accuracy (Panadero et al., 2019). Thus, from a pedagogical perspective, the benefits of self-feedback come from active engagement in the learning process through the deep *reflection* that can accompany it (Harris and Brown, 2013; Andrade, 2019). In other words, reflection is the step that elicits the benefits of self-feedback; reflective thinking can help students to explore and elaborate on their understanding of problems encountered during learning (Yan and Brown, 2017), and this enhances students' self-regulating skills, which in turn leads to improved learning (Yan, 2016).

However, although, self-feedback differs from other externally generated feedback processes in that it, first and foremost, is an internal practice that is conducted by and within the student, self-feedback might be more powerful as an instructional and learning activity if it includes external sources of feedback, such as teachers or peers (Boud, 1999). Butler and Winne's (1995) review pointed out that students have their own internal path to feedback that occurs regardless of the reception of explicit and direct external feedback from teachers or peers. This means that having students engage in an effective cycle of self-feedback benefits from the implementation of such scaffolds as modelling, formulating explicit criteria and using examples and other instructional tools (Panadero et al., 2019).

3.4.2 Peer feedback: receiving and creating reviews

Another way of engaging students actively with feedback processes that is beginning to receive more attention in higher education is the implementation of peer feedback, where students comment on each other's work and thereby develop their abilities to make academic judgements (Winstone and Carless, 2020).

Research shows that students often perceive the feedback they receive from peers as more understandable and helpful than teacher feedback because it is written in a more accessible language (Falchikov 2005). Where multiple peers are involved, the quantity and variety of feedback that students receive naturally increases (Topping 1998). Moreover, a key research finding concerning peer feedback is that students benefit, not only from receiving

feedback from their peers, but also from producing feedback reviews for their peers (Cho and Cho, 2011; Cho and MacArthur, 2011). Findings show that producing feedback reviews engages students in multiple acts of evaluative judgement, both about the work of peers, and, through a reflective process, about their own work (Nicol et al., 2014).

Peer assessment involves similar processes as peer feedback but has recently been distinguished from peer feedback because it is argued that the former includes a grading element, whereas the later focuses on only providing comments (Winstone and Carless, 2020). An important reason behind this distinction is that peer assessment usually generates more student concern than peer feedback (Liu and Carless, 2006; Harris and Brown, 2013). The potential learning benefits of peer feedback might also be undermined, at least for some students, when peer marking is involved (Sadler and Good 2006; Kaufman and Schunn 2011). Hence, making peer feedback count, for instance on the final grade, may be counterproductive to learning and constitute a reason to resist rather than implement it (Panadero and Brown, 2017).

Finally, and in line with effective self-feedback practices, students need teachers to model the processes of peer feedback for them and to coach them in doing it successfully (Winstone and Carless, 2020).

3.5 Essential Barriers and Decisive Methods

3.5.1 The difficulty in effecting change

Despite a growing body of literature and considerable investment on the part of universities, research still shows, unfortunately, that feedback in higher education continues to be poorly understood and enacted by both educators and students (Carless and Boud, 2018; Dawson et al., 2019). Winstone and Boud (2020) refer to this:

> Feedback practices in higher education have remained stubbornly similar and habitual for far too long, and things do not have to be this way (p. 10).

According to Molloy et al. (2020), the main problem that needs to be collectively addressed is that even with professional development of university teachers, and institutional cultures that value facilitation of learning rather than 'telling', learners and teachers may still have an expectation that feedback is part of the teachers' domain and is judged according to the information they generate.

To illustrate this, an important challenge with the implementation of more student-centred feedback practices, such as self- and peer feedback, is that, when given the choice, many teachers prefer not to promote students' active involvement (Jonsson et al., 2015;

Panadero et al., 2015). Research findings suggest several limiting factors that can help to explain this, for example, cultures and traditions, beliefs about teaching and learning, lack of incentives to engage in such practices, as well as time pressure and workload (Nicol and Draper, 2009; DeLuca et al., 2012; Carless, 2013; Draper and Nicol, 2013), even though many productive feedback practices require less time (Ajjawi et al., 2018). Beyond this, however, research also suggests that a possible explanation may be that not all students are comfortable with or ready for student-centred practices requiring their active participation in the feedback process (Jonsson et al., 2015).

Research points out that students are aware that the teacher is the most expert person in the classroom and some have grave doubts as to the necessity of relying on anyone's judgment other than that of the teacher (Peterson and Irving, 2008; Gao, 2009). Thus, some students resist the idea of assessing themselves, or others, preferring that the teacher is the one who does this (Panadero et al., 2016a). Students' unwillingness to participate in more student-centred practices ultimately affects the teachers and their workload, as, in order to progress, the students require more timely, detailed and individualised feedback on their work (Jonsson et al., 2015; Henderson et al., 2019). Furthermore, when teachers take all the responsibility for the assessment, they run the risk of making the students passive recipients of feedback, which counteracts precisely what formative feedback practices are aiming to achieve.

Moreover, in addition to the preference for teacher-controlled feedback information, students also report an interest in grades, and when these two are combined, which they often are in higher education, students tend to focus more on the grades than the feedback information, and teachers focus more on defending the grade rather than facilitating for more constructive feedback practices (Winstone and Boud, 2020). According to Winstone and Boud (2020), the entanglement of assessment (grades) and feedback (development) in higher education can thus impede the shift towards more student-centred feedback practices. Finally, it is also evident that students might not recognise dialogic interactions as feedback, especially when given informally (O'Donovan et al., 2016), meaning that an important dilemma to address is how to encourage students to become proactively involved in formative activities (Broadbent et al., 2018).

Another empirical way to explore the unwillingness, or perhaps unreadiness, of some students to participate in more student-centred practices is to use Carol Dweck's fixed and growth mindset research, which highlights the importance of students becoming 'learning ready'.

3.5.2 Becoming learning ready: mindsets and mind-frames

Carol Dweck's fixed and growth mindset research is well-known amongst educators, itself the culmination of 30 years of research on motivation (Hattie and Clark, 2019). Dweck's research shows that there are two core mindsets, or beliefs, that people have about themselves which shape how we approach challenges:

A fixed mindset: the belief that one's abilities cannot be changed, and

A growth mindset: the belief that one's intelligence, skills and qualities can be developed through effort, input and a range of learning experiences

These mindsets have an important impact on students by giving different meaning to achievement situations (Molden and Dweck, 2006). In this way, they shape students' achievement goals and their values, change the meaning of failure and guide their responses to difficulty and their use of self-regulating learning strategies (Dweck and Master, 2012). A more comprehensive theoretical description of students' motivational beliefs and achievement goals will be presented in Chapter 4.

In general, traditional teacher-led feedback, such as rewards, grades or points, enhances performance rather than mastery involvement – that is, it focuses students' attention on their abilities rather than encouraging them to think about the work itself, the importance of effort and how they can improve (ibid). When assessment is used for summative purposes, carrying high stakes for students, it creates a strong reason to make an effort. But this effort, for the vast majority of students, seriously undermines the scope and depth of learning (Harlen, 2012). On the other hand, feedback which focuses on what needs to be done can encourage everyone to believe that they can improve. Such feedback can enhance learning, both directly through the ensuing effort and indirectly by supporting the motivation to invest in such an effort (Schunk et al., 2010). Thus, when designing a feedback practice, care should be taken to differentiate between what is 'mastery' and 'performance' oriented (Daniel and Poth, 2017), and what creates a growth mindset culture relating to feedback, in which students want to actively challenge themselves, are not afraid of failure or making mistakes and know that they can 'grow' with their learning (Hattie and Clark, 2019). Finally, good classroom *relationships* between teacher, students and their peers is paramount (Black and Wiliam, 1998). Without a classroom philosophy that views mistakes as an opportunity for learning and that encourages honest reflection, the learning benefits of more student-centred feedback practices are likely to be compromised (Yan, 2016; Molloy et al., 2019).

3.5.3 Teacher support and teacher learning communities

Another important barrier that is important to consider is a lack of understanding about formative feedback, or perhaps more importantly, that fact that even though many educators may understand what formative feedback is, this does not always translate into actual practice (Nicol and Draper, 2009; DeLuca et al., 2012; Carless, 2013; Boud et al., 2016). It is thus important to develop recommendations based on the research literature on how to support teachers in their implementation of student-centred feedback practices as part of an overall learning-focused approach to assessment (Panadero et al., 2016b).

Even though a strong research base on how to effectively help teachers to implement a high-quality formative assessment practice is lacking (Anderson and Palm, 2018), the research literature from lower education has several and clear recommendations in relation to facilitating changes and developing formative assessment and feedback practices.

Initially, what comes out of the research literature on the school level is an acknowledgment that what teachers do in 'taking on' research is not a more or less passive adoption of some good ideas from someone else, but an active process of knowledge creation (Wiliam and Thompson, 2007), an acknowledgement, in other words, that adult learning is fundamentally similar to that of school students (Bransford et al., 2000). Important examples that have had significant impact on teaching and assessment practices in school are the work of Black et al., (2003) in the King's Medway Oxfordshire Formative Assessment Project (KMOFAP), and the studies in the 'Learning How to Learn' project (LHTL) (James et al., 2007), to mention a few.

More precisely, analysis of the LHTL survey provides very strong evidence relating to the importance of 'inquiry' in teachers' learning practices (Pedder et al., 2005; Pedder and James, 2012). This reflects a number of research-informed, classroom-based approaches to collaborative teacher learning, and these are the most directly and powerfully approaches associated with the promotion of self-regulation in classroom assessment (Pedder, 2006).

However, the most important source of support for ensuring that changes in formative assessment and feedback practices are sustained is frequent personal contact with colleagues (Dekker and Feijs, 2005). This means that opportunities for teachers to work collaboratively in an atmosphere of trust and mutual respect can build the social capital needed for teachers to learn, share, reflect on and develop their ideas and practices. Forming learning teams, also known as *teacher learning communities* (TLCs), for professional development in formative assessment is therefore a clear recommendation (Stiggins, 1999; Thompson and Wiliam, 2008; Leahy and Wiliam, 2012). For collaboration to be effective, however, it needs to be

purposeful and must be structured so that it supports the goal of improving teaching and learning (Schneider and Randel, 2010).

However, if changes in practice are to be lasting, they must be integrated into teachers' existing *routines*, and this takes time (Leahy and Wiliam, 2012). In practical terms this means that one of the greatest threats to a successful intervention is that teachers' time and attention is disrupted and split (Thompson and Wiliam, 2008). However, it is not only difficult to make many of the changes in practice associated with implementing a formative feedback practice because the practice is habituated – but also because there are, for example, widely distributed and strongly held beliefs about the value of grades in motivating students (ibid).

Timperley and Alton-Lee (2008), maintain that teachers do not approach professional learning or teaching situations as empty vessels but rather as practitioners who have rich theories about how students learn, how best to teach them and what constitutes desired content and outcomes. Practitioners who intend to build professional knowledge by prescribing particular teaching behaviours, an 'outside expert developed recipe', so to speak, without engaging existing beliefs or understanding the constraints of their practice situation, and those who claim that providing teachers with time and resources, alone, is sufficient for promoting professional learning, without developing teachers' current knowledge and practice or challenging problematic attitudes, typically fail to take this complexity into account, and the interventions are much less effective. For teachers to be able to implement a new formative classroom practice, they also need the motivation to commit to learning how to carry out such practice, as well as the commitment to implement it in the classroom (Anderson and Palm, 2018).

Thus, not only must the support be sustained over time, this support must embed teachers' learning within the realities of their practices, in their own institutions and classrooms. The support must also allow for a process of trial and error where it is allowed to make mistakes, and where it opens for repeated cycles of learning, practice, reflection and adjustment within their native contexts (Thompson and Wiliam, 2008; Schneider and Randal, 2010).

An illustration: Keeping Learning on Track. Keeping Learning on Track (KLT) is a sustained professional development programme pioneered by Dylan Wiliam and his colleagues at the Educational Testing Service (ETS). The programme was designed to support teachers in using formative assessment in their everyday teaching. The model will not be reviewed in detail here (see Thompson and Wiliam, 2008, for a comprehensive and exhaustive review), but it is important to point out some essential elements here.

A key component of this approach is the idea that teachers should have complete freedom to choose which of the KLT strategies and techniques they want to use. This freedom ensures that teachers' professional judgment is a major component in the successful implementation of strategies and techniques. As a primary process to effect changes in practice, embedded in the programme's theory of action, the KLT programme recommends TLCs, where opportunities for practice in real settings, followed by reflection, are structured. As teachers are accountable for taking charge of their own learning, the idea is to create a climate in the meetings that make expectations clear. Essential parts of a TLC meeting are reflective self-reports, where all participating teachers share their most recent efforts (the How's-It-Going developmental segment), supportive accountability shown by colleagues, a written commitment to take specific next steps (the Personal-Action-Planning developmental segment) and the *storytelling/feedback/planning cycle*. However, it is not necessary to meet in a KLT-sanctioned learning community that follows a KLT module, according to Thompson and Wiliam (2008), as any structure that supplies these processes will do, as long as these essential ingredients are present.

4. THEORETICAL GROUNDING

This chapter will place the research work in a wider learning theory landscape. Based on key concepts from the introduction, empirical review and not least the overriding research question in this thesis, *how can educators develop their own feedback practice in a more formative way*, the learning theory landscape will include both a sociocultural understanding of learning and a social-cognitive approach to motivation. Both these directions can contribute significantly to an understanding of developing feedback in a more formative way, each with their clear strengths. To clarify this further, this chapter will start by looking into the choice of learning theory landscapes, and how these can be used to strengthen a comprehensive understanding of the findings.

Formative assessment as regulation of learning. Black and Wiliam's theory of formative assessment as regulation of learning will be used here as the overriding framework for discussing results in terms of what it may mean and what it means in practice to create a more formative feedback practice. Since this thesis conceptualises formative assessment in accordance with the definition by Black and Wiliam, which is built on their theory, it is natural to examine the theoretical framework in an overarching understanding of formative assessment.

Sociocultural view of learning. Considering the strong dialogical focus in the research literature's new ways of thinking about feedback in higher education, it is also natural, interesting and important to incorporate a sociocultural view of learning, which places much emphasis on knowledge being constructed through interaction and in a social context. Even though Black and Wiliam's approach has been to treat social-individual interaction as an important feature of learning, which in turn confirms the importance of interactive *dialogues* as the core in their attempt to theorise formative assessment, important elements of the socio-culturally oriented view of learning will be included here, comprising the expanded dialogue concept, and thus also reflection and mediating tools. This will be used to understand the importance of the context that was created and explore how the new learning design in the form of both facilitation and implementation of different student activities may be understood as dialogues for learning.

A social-cognitive approach to motivation. The third and final part of this chapter is based on an increasing focus within the research literature today where the development of productive feedback practices is understood as entailing more than quality assessment and instructional decision making that promotes learning. Careful management of the motivational and emotional aspects of the experience from the students' point of view is also required. To

elaborate on these aspects, a social-cognitive approach to motivation, theories on student intelligence, will be examined here. This will have an important position in the discussions on the students' observed achievement goal patterns in a new formative learning design and will be assessed and applied as an essential theoretical model for the development of more formative feedback practices in higher education.

4.1 A New Theory of Formative Assessment

To provide a better theoretical grounding for formative assessment, Wiliam and Thompson (2007), drawing on both Ramaprasad (1983) and Sadler (1989), matched the three most important processes involved in teaching and learning, establishing where the learners are in their learning, where they are going and what needs to be done to get them there, with the different agents involved in assessment (teachers, peers and learners), and provided a unifying basis for diverse practices that are said to be formative, which indicates that formative assessment can be conceptualised as consisting of five key strategies (as shown in the framework in Table 2).

Table 2. Framework Relating Strategies of Formative Assessment to Instructional Processes (from Wiliam and Thompson, 2007, p. 63).

	Where the Learner Is Going	Where the Learner Is Right Now	How to Get There
Teacher	1. Clarifying learning intentions and criteria for success	2. Engineering effective classroom discussions and tasks that elicit evidence of learning	3. Providing feedback that moves learners forward
Peer	Understanding and sharing learning intentions and criteria for success	4. Activating students as instructional resources for one another	
Learner	Understanding learning intentions and criteria for success	5. Activating students as the owners of their own learning	

While each of these five "key strategies" has generated a substantial research base individually (see Wiliam, 2007, for a summary), the idea behind Black and Wiliam`s seminal article in 2009 was that they can also be viewed collectively as strategies for the *regulation of learning*. Black and Wiliam (2009) thus presented a general overview of *interactive*

dialogues, or formative interactions, which constitute the grounds for merging all the five strategies, rather than serving any of them directly.

4.1.1 Precision in definition

To provide a more comprehensive definition of formative assessment, based on this theoretical grounding, Black and Wiliam (2009) proposed, as presented in the literature review, that assessment 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 (p. 9).

To gain a better understanding of the definition, Black and Wiliam elaborated on five key points. First, they clarified that the term "*instruction*" is used in the sense in which it is used in the United States – the design of *learning environments* – and is used to describe any planned activity intended to create learning. Second, the "next step in instruction" can be taken by teachers, learners or their peers, or any combination of these three, implying that any of them – *teacher, learner* or *peer* – can be the agent of formative assessment. Third, the focus of the definition is on *decisions*. Black and Wiliam noted that the focus of the definition could have been on the intentions of those involved in instruction, in collecting the evidence, but that would mean that a situation in which evidence was collected but not used would be formative. Such a definition would thus, in that sense, be too open. The focus could also have been on the outcome and thus required that the assessment did in fact lead to better learning. This, however, as they understood it, would appear as a rather stringent criterion, given the unpredictable nature of learning. Fourth, the idea "*that the decisions are likely to be better*" reflects the fact that even the best designed interventions will not always result in better learning for all students. What this definition requires, on the other hand, is that the evidence collected *improves* the likelihood that the intended learning takes place. Finally, *the assessment does not need to change the planned instruction*. The evidence elicited by the assessment may indicate that what the teacher had originally planned to do was, in fact, the best course of action. Thus, this would not be a better decision (since it was the same decision that the teacher was planning to make without the evidence), but it would be a better-*founded* decision.

Moments of contingency. From this definition, Black and Wiliam, proposed that formative assessment is, in essence, concerned with "the creation of, and capitalization upon,

'*moments of contingency*' in instruction for the purpose of the regulation of learning processes" (2009, p. 6). This means that the strategies, as presented in Table 1, can be integrated within the more general theoretical framework of the regulation of learning processes as originally suggested by Perrenoud (1998).

4.1.2 Regulation of effective learning environments

Within this framework only the learner can create learning, but the teacher has a clear responsibility to design the learning environment so that the learner can achieve the intended learning (Black and Wiliam, 2009). This means that the responsibility for learning in a formative mode rests with *both* the teacher and the learner (ibid). This means that the teacher is regarded as responsible for "engineering" a learning environment, both in its *design* and its *operation* (Wiliam and Thompson, 2007). A more thorough elaboration of various regulations of learning environments, as conceptualised and operationalised by Wiliam and Thompson (2007) and Wiliam (2011), follows below.

Proactive regulation. Proactive regulation is achieved "upstream" of the lesson itself (before the lesson begins), by setting up "didactical situations" (Brousseau, 1984) where the teacher "does not intervene in person, but puts in place a "metacognitive culture". The idea of shared responsibility is essential in this setting. Sharing of responsibility means that teachers must give increased priority to the need to equip students with the cognitive strategies required for them to be able to live up to their responsibility and thus develop new understanding (Black and Wiliam, 2012). In other words, the teacher's responsibility is to engineer situations in which the opportunities for the students to learn, and to develop learning autonomy, are maximised.

Interactive regulation. On the other hand, the didactical situation may be set up so that the regulation is achieved through the mediation of the teacher, interactive regulation, where the teacher, in planning the lesson, creates questions, prompts, or activities that evoke responses from the students that the teacher can use to determine the progress of the learning, and if necessary, to make adjustments, also called a *synchronous* moment of contingency. Examples of such moments include teachers' "real-time" adjustments during one-on-one teaching or plenary class discussion (Black and Wiliam, 2009).

Retroactive regulation. When teachers reflect on instructional sequences after they have been completed, it provides a third form of regulation, namely retroactive regulation (Wiliam, 2011), also called *asynchronous* moments (Black and Wiliam, 2009). For example, ideally, from examining the students' responses to the task, the teacher will be able to judge a)

how to help the learners learn better and b) what can be done to improve the teaching of this topic to a future class. In this way, the assessment could be formative for the students, through the feedback the teacher provides, and formative for the teacher herself, in that appropriate analysis of the students' responses might suggest how the lesson could be improved for other students (Wiliam and Thompson, 2007).

These "Moments of contingency" thus point to the instructional sequence when the instruction can proceed in different directions according to the responses of the students, and these are the *heart* of the regulation of learning (ibid). One of the features that make a lesson formative then is that the lesson can change its course in the light of evidence about the progress in learning, facilitating for interactive dialogues.

Interactive dialogues. In sum, within this theoretical framework the fundamental core activity of formative work is the enrichment of dialogical interactions; a formative interaction in which an interactive situation influences cognition, i.e., an interaction between external stimuli and feedback, and internal production by the individual learner (Black and Wiliam, 2009; 2012). This dialogue thus involves looking at three aspects, the external, the internal and their interactions, as illustrated in Figure 1.

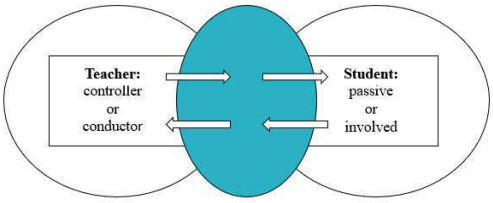

Figure 1. The interacting domains in an interactive dialogue (from Black and Wiliam, 2009, p. 11).

4.1.3 Contingent interactions: the need for a model

Thus, in a formative situation the teachers' attention must be focused on what they can learn about the students' thinking from their responses. However, what the students actually hear and interpret is not necessarily what the teacher intended to convey and rehearsed, and what the teacher hears and interprets is not necessarily what the students intended to convey (Black and Wiliam, 2012). Perrenoud (1998), in his response to Black and Wiliams seminal review of 1998, identified a similar challenge:

> Without a theoretical model of the mediations through which an interactive situation influences cognition, and in particular the learning process, we can observe thousands of situations without being able to draw any conclusion. (1998: 95)

All in all, this complexity presents the teachers with quite a challenge since their feedback needs to be constructed in the light of some insight into the mental life that lies behind their students' utterances (Black and Wiliam, 2012). A theoretical model that can be relevant to the management of dialogic interactions between teachers and learners are theories concerning students' motivational *self-beliefs*. Research shows that feedback both regulates and is regulated by motivational beliefs. External feedback has been shown to influence how students feel about themselves (positively or negatively), which achievement goals they pursue, and what and not least how they learn (Dweck, 1999). The way feedback interacts with achievement motivation and beliefs can thus be an important guide to teachers' contingent actions.

But before moving on to a theoretical understanding of theories of student intelligence, it is important to elaborate further on the concept of dialogue to gain a fuller understanding of dialogue as a theoretical concept. Below a sociocultural view of learning is presented, wherein dialogue is seen as being fundamental to learning.

4.2 A Sociocultural Understanding of Learning

Dialogue as critically important for all sense-making. As mentioned in the introduction to this chapter, the sociocultural view of learning places much focus on how knowledge is constructed through interaction and in a *context* (Bråten, 2002). There are a number of differences between theorists in this tradition, but they share a fundamental understanding that the social group and the community individuals belong to are the point of departure for learning (Dysthe, 2001; 2008).

More concrete, from a sociocultural stance, dialogue and interaction are basic elements of learning (Bakhtin, 1981), not merely a positive element in the learning environment (Dysthe, 2001). Russian language philosopher and literary theorist Mikhail Bakhtin (1981) used the concept of dialogue in three contexts, dialogue in an ontological macro-perspective (as an overarching view of human existence), dialogue in a micro-perspective (meaning and understanding are created in interaction) and dialogue as the counterpart of monologue. Bakhtin thus rejected the traditional communication model where a message is transferred from a sender to a recipient. Learning always occurs in interaction, whether in dialogue with live voices, or in dialogue with texts, which also are an expression of voices. From a

theoretical perspective, an *extended* dialogue concept is thus applied which includes both an external and an internal dialogue where the former involves participants present there and then, while the latter is an internal dialogue and may be described as an interaction between the individual student and the academic content (Bakhtin, 1981; Dysthe, 1996).

In a sociocultural view of learning meaning is thus not created by the individual, but rather in the interaction between the interlocutors (Bakhtin,1984). According to Dysthe (1996; 2008), such a sociocultural and interactive view of learning is completely fundamental for all teaching and assessment: "we" are the creators of meaning. Meaning and understanding is generated by the communication situation itself, in the interaction between those who participate (Rommetveit, 1974).

Situated learning: the notion of mediation and language. More concrete, in a sociocultural view of learning knowledge cannot be detached from the context it is developed in, it is situated (Dysthe, 2001). Therefore, the surroundings and tools involved in a learning situation are important, what is learnt, how much and how it is learnt will therefore be determined by the applicable sociocultural conditions (Säljö, 2001). Interaction and dialogue are not only something that takes place between people, but also between people and various cultural artefacts (Dysthe and Igland, 2008). An important part of the sociocultural context is thus the tools used in the different activities.

Mediation hence is a key concept in sociocultural theory (Dysthe and Igland, 2001). The concept, one of the fundamental topics derived from Vygotsky (1978), sees that cognitive functions can be traced back to social processes: the individual's cognitive functions and skills originate in social activity and interaction. These functions are on different levels. The first one is a social or inter-cognitive level, which concerns what takes place between people in a social context. Next is an intra-cognitive level, where tools that are introduced and presented in a social context are transformed into thinking and action on the individual level. Vygotsky calls this development internalisation (Dysthe and Igland, 2001). This means that all learning and knowledge initially occurs *externally*, in interaction with the surroundings, thereafter it occurs on the *internal* level, in our thoughts (Dysthe, 2001). More concrete, the internalisation process means that mental functions are presented, transferred and supported or controlled – i.e. mediated – by means of physical and intellectual tools that are used in various forms of social activity (Dysthe and Igland, 2008).

The most important mediating tool for cognitive development and learning is, according to Vygotsky, *language* and the use of language. Cognitive abilities are structured and developed through the acquisition of language and concepts as they are used in social

interaction. Dysthe (2001) also sees language as the most important mediating tool, claiming that "how language can function as a cultural mediating tool is a particularly important topic in sociocultural learning theory, and here not the least Bakhtin has furnished us with new insight" (ibid:47). For Bakhtin, words do not merely belong to an individual person or persons taking part in a dialogue. Words are always part of a chain of utterances. An utterance is a response to previous utterances, which means that an utterance depends on other utterances to become meaningful. This creates a dynamic language where words change their meaning according to who utters them and in which context (Bakhtin 1981, 1986). Therefore, the feedback from the second person, the response, becomes the activating principle that creates the basis for understanding (Bakhtin, 1981). The responses may be explicit answers or unarticulated responses: "The fact is that when the listener perceives and understands the meaning (the language meaning) of speech, he simultaneously takes an active, responsive attitude towards it. He either agrees or disagrees with it (completely or partially), augments it, applies it, prepares for its execution, and so on" (Bakhtin 1986:68). Meaning and understanding occur in this encounter.

Interaction and dialogue with other and more competent persons are therefore quite important in cognitive development and learning (Vygotsky, 1978). According to Bakhtin (1984), the tensions and confrontations between the voices of the participants are where new insight and understanding occur. Thus, not just any exchange or dialogue creates understanding and meaning, something more than reproduction occurs when different voices participate in the dialogue and interact in the sense that they build on each other and/or are in conflict (Dysthe, 2001). According to Dysthe, this perspective is completely fundamental for what she calls learning dialogues.

To elaborate, dialogue in a sociocultural understanding can also be a counterpart to discussion, in the sense of arguing for a personal point of view and aiming to convince others to find a common solution. Dysthe (1996) considers dialogue and discussion to be, respectively, learning dialogue and argumentative dialogue. The aim of the former is to progress further in insight and understanding through common exploration and by using the thoughts of others to develop one's own thoughts, while the argumentative dialogue often functions as critical testing of arguments and stances, which may lead to the discovery of weak aspects of one's own argumentation.

Symmetry and asymmetry. Another important theoretical element in terms of dialogue is the level of symmetry and asymmetry in dialogue situations (Dysthe, 2008). Basically a teacher-student relationship is always asymmetric in the sense that the basis for one being a

teacher is that this person has more knowledge about the subject or has more knowledge about how learning situations may be structured and carried out. Asymmetry is, however, not necessarily a barrier to dialogue, rather the opposite. The fact that we know and can do things in different fields renders dialogue necessary and useful. It is important to bear in mind, however, that asymmetry is also often connected to different kinds of status, such as between student and teacher, and that there is accordingly a power aspect in the asymmetry. An important element here is what Rommetveit (1991) calls "distribution of epistemic responsibility". This means that one of the partners in a dialogue may have complete control of the conversation through deciding what is worth talking about, and through mastering the "correct" language. This may in practice function as an effective barrier to having a constructive dialogue between a teacher and a student.

Reflection as internal dialogue. Finally, and in line with the extended dialogue concept, internal dialogue is also an essential part of a student's learning process. Dysthe (1996), describes it, in brief, as a part of a three-part division of learning sequences, where phase 1 is input, phase 2 is processing and phase 3 consolidation. The role of the teacher is traditionally prominent in phase 1, for example presenting information. It is also common to organise phase 2 with room for interaction and dialogue (for example group work). The third phase, consolidation of knowledge through *reflection* (an internal dialogue), is, however, often lacking after a dialogic sequence, and is generally not a widely used practice in higher education (Dysthe, 1996; Boud et al., 2018; Carless, 2019). According to Dysthe, teachers must actively implement this phase as an essential final instance in a learning sequence, as this phase is entirely necessary if students are not to be left with a number of fragmented impressions. The students must have the opportunity to sum up and reflect on what they actually are left with of learning. More specifically, they must be allowed the opportunity to sort different perspectives, reflect on choices and implications and to link new understanding to what they already know, and in this way use the internal dialogue as a tool in their own development of knowledge.

4.3 Students' Motivational Beliefs: Two Frameworks for Understanding Intelligence and Achievement

With a deeper sociocultural description of the concept of dialogue in place, the last part of this chapter will return to the management of dialogic interactions and focus on theories concerning students' motivational *self-beliefs*. The emotional aspect of assessment and feedback has often been underrated in higher education, but it is essential to a student's

motivation and thus participation in more student-centred learning practices (Harlen, 2012; Bryan and Clegg, 2019).

A theory that can help shed better light on this important issue is found in the extensive research of Carol Dweck and colleagues on theories of student intelligence (and Dweck's model of achievement goals). Evolving within a *social-cognitive framework*, this is based on the idea that people develop beliefs that organise their world and give meaning to their experiences. Social cognitive theory emphasises the importance of social influence on human behavior and assumes a triadic reciprocality between personal factors, behaviors, and environmental influences as they interact with and affect one another (Bandura, 1997; 2001).

4.3.1 Self-theories and the meaning system approach

The self-theories Dweck and her colleagues focus on are people's beliefs about the fixedness or malleability of their personal qualities, such as their intelligence: Do people believe that intelligence is s fixed trait (the entity theory) or a malleable quality that they can cultivate through learning and effort (the incremental theory)? The validity of this theory has been confirmed in higher education (Grant and Dweck, 2003; Yorke and Knight, 2004), where recent research shows a majority of students with a fixed mindset (Forsythe and Johnson, 2017). A more comprehensive theoretical description follows below.

As seen in Figure 2, the entity theory creates a meaning system focused on the goal of measuring and validating competence and is thus associated with ability-oriented *performance goals*, ability attributions for setbacks and the belief that effort indicates low ability (Dweck and Molden, 2005). Due to the stable conception of ability, these students will be most concerned about how their performances are evaluated, how they compare with those of others and will try to be best (Schunk et al., 2010). In other words, when students pursue performance goals they are concerned about their level of intelligence: they want to look smart (to themselves or others) and avoid looking dumb (Dweck, 1999). These goals and beliefs lead, in turn, to helpless or defensive reactions to difficulty and to lower self-esteem, intrinsic motivation and learning in the face of difficulty.

An incremental theory, on the other hand, creates a meaning system built around the acquisition of competence and is thus linked to *learning goals*, effort and strategy attributions for setbacks, and the belief that effort increases ability. This reflects a desire to learn new skills, master new tasks and understand new things. These students will try to increase their competence and judge the success at reaching their goal by using criteria focused on self-improvement, and not social comparison with others. These students have little or no fear of

failure. Instead, their goals and beliefs promote mastery-oriented strategies in the face of challenge, and this leads to enhanced self-esteem, intrinsic motivation and learning (Dweck, 1999; Dweck and Molden, 2005).

This means that students' active use of learning strategies – and their continued use of them in the face of challenges and difficulties – is based on the *belief* that these strategies are necessary for learning, and that they are effective ways of overcoming obstacles (Dweck and Master, 2012). Students' self-theories can thus have a strong impact on how they self-regulate and how effectively they learn (Zimmermann and Schunk, 2012).

Theory of Intelligence	**Goal Orientation**	**Confidence in Intelligence**		**Behaviour Pattern**
Entity theory (intelligence is fixed)	Performance goal (to gain positive judgement of competence)	If High	→	Mastery oriented Seeks challenge High persistence
		But		
		If low	→	Helpless Avoids challenge Low persistence
Incremental theory (intelligence is malleable)	Learning goal (to increase competence)	If High or Low	→	Mastery oriented Seeks challenge (fosters learning) High persistence

Figure 2. Dweck's model of self-theories and achievement goals.

Beyond the sections presented in Figure 2, research has also developed an important distinction in performance goals (Elliot and Harackiewicz, 1996; Harackiewicz et al.,1998; Elliot, 1999). Elliot and Harackiewicz and their colleagues differentiated approach performance goals (students who strive to appear competent and demonstrate high ability) from avoidance performance goals (students who strive to conceal their relative incompetence and avoid negative judgments). Although Dweck did not formally separate two distinct performance goals, her theory includes concerns over avoiding judgement of incompetence in the conceptualisation of performance goals, similar to the avoidance-performance orientation. The benefit from attending to the approach-avoidance distinction in the social-cognitive model of achievement motivation has been clearly documented, and research demonstrates

that incremental theories of ability are direct predictors of performance attainment and intrinsic motivation, thus supporting Dweck's initial theory. Entity theory increases the adoption of performance-approach and performance-avoidance goals and decreases the adoption of mastery approach goals (Cury et al., 2006).

4.3.2 A dynamic view: strategies to enhance learning and motivation

According to Dweck, the two theories of intelligence are usually operationalised as opposite sides of a dichotomy and people hold to one or the other implicit theory (Schunk et al., 2010). But the theory actually predicts that there may be a continuum between entity and incremental beliefs and that individuals show mixed theories of intelligence (Dweck, 1999; Dweck and Master, 2012).

> It is often difficult for people to believe that simply changing a belief will have much impact given the many things that affect students' learning. However, as this belief is at the heart of student motivation, it can have more impact than one would expect. (Dweck and Molden, 2005, p. 136)

This means that although self-theories and achievement goals can be relatively stable over long periods of time (Robins and Pals, 2002), they are knowledge structures and, as such, their accessibility can be changed by powerful situations and targeted interventions with striking effects (Aronson et al., 2002; Nussbaum and Dweck, 2006; Blackwell et al., 2007). The fact that self-theories can be induced experimentally and altered through targeted interventions suggests a dynamic view of these theory-based motivational systems. In other words, these theories are useful to teaching and assessment because they assume that students' beliefs are not just stable personal characteristics of individuals but can be shaped by teachers' actions and by how learning environments are organised and structured. Thus, the more teachers develop, and regulate, supportive learning environments that help students to see failure as a natural part of learning and not an indictment of their worth as individuals, and allow students to get things wrong and learn from their mistakes, the less likely the students are to adopt an entity view and mastery goals (Dweck 1999; Bryan and Clegg 2019).

Finally, learning an incremental theory puts students in charge of their own mind and how they develop. It thus *motivates* them to put their repertoire of learning skills into practice. In other words, having learning skills is not the same as using them. Entity theorists may have learning skills in their repertoire but may not use them. This is why it is so important to ensure that students do not only have the learning strategies they need but also the motivation to apply them (Dweck and Master, 2012).

5. THE CONTEXT OF THE RESEARCH

The overarching theme of this thesis is use of feedback in a higher education learning context. The thesis has been based on an intervention study using a qualitative approach within an interactive action research framework (see Chapter 6 for a more detailed description). Through a mutual collaborative process, inspired by learning practices from school teaching, two mathematics teachers and me as the researcher, used the "seven principles of good feedback practice" as an intervention tool to create a formative assessment practice in mathematics. Below is a short overview of the current collaborative process, its work methods and a comprehensive description of how the seven principles were implemented in practice.

5.1 The Point of Departure: "a relentless hunt to find errors"

Let us start by taking a brief look back in time. To understand the changes the teachers made in their practice it may be useful to return to the start to summarise and elaborate on the original practices of the teachers and how they perceived their role as teachers at that time. This can help to highlight any contrasts and to obtain deeper insight into the teachers' changes.

The way it was. The earlier assessment practice of the teachers were focused on short tests (taking approximately 45 minutes), which were given every two weeks, and which the teachers then spent two weeks correcting, then returning them to the students, followed by a review of the test tasks on the board. The teachers described the review as relatively simple and traditional where they basically presented an answer key to the students, which the students could then copy. In this way the students would see how they should have calculated, and the solutions to the tasks were explained, which they then could compare to their own performance. The teachers were clear that the review was teacher-controlled, and the teachers were the most active party. They found that the students' focus in general was on copying from the board.

What is also important to note about this process in this intervention study is that the teachers had a genuine desire to change. They all experienced a form of meaninglessness in connection with their own practice, not only on their own behalf, but also their students'. The teachers said that the focus was mainly on the level of correctness, perhaps not that surprising in a subject such as mathematics. They corrected the students' papers by assessing the number of correct and incorrect answers, returned the tests and presented an answer key, based on right or wrong. In the words of one of the teachers: "It felt like an endless hunt for errors".

Another way of describing this, using terms from theory and motivation research, is that it was a performance-oriented practice, where the assessment criteria, the teachers' corrections and the post-test review were focused on the students' achievements in terms of the number of correct answers. The teachers wanted to create a practice where the students participated more actively, and to move testing away from the too one-sided focus on errors, as they saw it. The main intention of the teachers was that the students would perceive the assessments in mathematics as an arena for learning, which led to a theory about student intelligence and achievement-goals becoming important elements in the team's evidence base.

5.2 Part 1. Inquiry-based Collaboration and Teacher Support

The initial phase of this study, which lasted for one-and-a-half school years, consisted of comprehensive *interactive action research* work (see chapter 6 for more detailed methodological descriptions). The research was initiated by establishing a cooperative relationship between the author of this thesis and two mathematics teachers from two preparatory engineering courses at the Norwegian University of Science and Technology (NTNU). Together "author" and the teachers formed a small team with the shared intention of establishing and applying a formative assessment practice in mathematics. The "seven principles of good feedback practice", recommended by Nicol and Macfarlane-Dick (2006), comprised the pedagogical framework for this objective. To summarise, the framework states that good feedback practice:

1. helps clarify what good performance is;
2. facilitates the development of self-assessment in learning;
3. delivers high-quality information to students about their learning;
4. encourages teacher and peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape teaching.

"Author", occupying the role of partner, collaborator and educational counsellor, then guided the teachers, both individually and as a team. In this way the teachers were introduced to the various principles, and through dialogue with the teachers, collaboration within the team, training in practice, observation and reflection, "author" and the teachers agreed together on how the principles should be applied in practice (this process is described in more detail in Chapter 6, section 6.1).

The teachers' experiences of this collaborative process and how they experienced changing their own practice is the main theme of Article 1. The teachers' foremost experiences of their own change process are thus described in this article.

5.3 Formative assessment inspired by the "seven principles of good feedback practice"

5.3.1 New assessment requirements

Inspired by, and in line with previous important school research (Black et al., 2003), the team decided to use the aftermath of minor tests as an opportunity for formative work. The assessment practice in mathematics involves 14 mathematics assessments during a school year: six mathematics tests (where students individually solve various mathematics tasks, calculated either with or without a calculator), six multiple-choice tests and two mock exams, followed by a final summative exam. The mathematics and multiple-choice tests are given at approximately 14-day intervals, whilst the mock exams are held at the end of the fall and spring semesters, with the final exam at the end of the spring semester.

The team worked mainly with the mathematics and multiple-choice tests, thus a total of 12 assessments, and examined how these situations could be applied in a more formative way.

The aim was to promote and establish a strong focus on learning, and thereby reduce the focus on performance, errors and correction. To achieve this, it was agreed that the entire assessment culture had to be challenged and changed. Grounded on the "seven principles of good feedback practice" (Nicol & Macfarlane-Dick, 2006), various learning objectives, such as self-feedback, making an effort, reflection and dialogue, replaced the more traditional performance objectives. Furthermore, it was decided that the assessments were not to be completed once the performance feedback, such as a teacher-led grading process, had been given. To gain approval in the various mathematics assessments, the students were measured against specific learning rather than performance goals. This meant that it was not their academic performance (number of correct answers on a test) that determined whether their assessment performance was approved, it was rather whether they had focused on and put an effort into reaching the *learning objectives* of the assessments. The main focus was on the students and their efforts to learn through reflection, self-feedback and peer dialogue. To achieve an overriding focus on mastery and learning, teacher-led grading was replaced by *self-feedback* through *reflection.*

5.3.2 The aftermath of minor tests as an opportunity for interactive dialogic reviews

In stark contrast to the teachers' previous practice, and as an important stage in creating a formative practice, the students' performance changed from being the main focus of the assessment to becoming an introduction to the core of the new assessments, namely the *interactive dialogic reviews*.

More concretely, all the mathematics assessments included in this study were given in classrooms and lasted about two teaching periods (45 minutes each). During the first 45 minutes, the students worked individually (on about 6 to 10 questions), followed by a 10-minute break. After the break the main part of the "formative aftermath" began, the interactive dialogic review. This was developed as a joint plenary review of the test questions and their possible solutions, with a strong emphasis on immediate performance feedback using response technology, dialogue, active student participation, effort, reflection and self-feedback.

The students assessed themselves during this interactive dialogic review. However, it is important to clarify that this did not involve them writing "correct" or "incorrect" on various tasks, but rather *reflecting* on their mathematical solutions, choice of strategies, level of effort, learning outcomes and further progress. In sum, the students participated in a joint interactive dialogic review for an elaborate review of the test questions immediately after completing the test and focused on unravelling their misconceptions and creating further learning opportunities.

Below is a more detailed description of how the teachers implemented "the seven principles of good feedback practice" in a total of six mathematics tests and six multiple-choice tests over one school year. The procedures in these two types of assessment have clear differences and will therefore be presented separately.

The teachers' intentions behind the various choices when they implemented the seven principles, their experiences of the principles in practice, and finally, the students' experiences, are presented and thoroughly described in Article 2.

5.3.3 Assessment procedure: mathematics test

A mathematics test was initiated with a short teacher introduction (*principle 1*). Here the following were presented to the students: the academic goals of the test (e.g. which part of the curriculum the students would be questioned on), followed by an emphasis on the main purpose of the test, namely to create an opportunity for students to learn from their misconceptions, to make an effort and experience mastery (*principle 5*). The teacher pointed

out two important "learning tools" that the students were to use: self-feedback and reflection (*principle 2*). As mentioned above, the students' academic performances were "moderated" in the sense that it was not the number of correct answers on a test that the teacher endorsed as important, but rather how the students dealt with their performance during the interactive dialogic review (*principle 3*).

After the introduction, the students started working on the test questions individually. When the test was finished, they were given a short break before they returned to the classroom and a joint plenary interactive dialogic review began.

During the interactive dialogic review, all the questions and their solutions were presented by the teacher and discussed with the entire class. The students were repeatedly encouraged to participate and thus be in dialogue with the teacher (*principle 4*). The students' main task in the interactive dialogic review was to evaluate and assess their own performance (*principle 2*). As part of this self-assessment, they were regularly given time to individually *reflect* on their own achievements (*principle 2*), and were encouraged to consider the following questions during reflection:

- How did you solve this task?
- Which method did you use?
- Are there other methods or strategies you could have used?
- How can you improve?
- Where should you put further effort?
- How did you cope?

The students wrote down their reflections and submitted them to the teacher after the interactive dialogic review. Their reflections served as a *criterion* for approval of the assessment. This criterion also applied to the multiple-choice tests. Below are a few examples of how the students reflected during the interactive dialogic review.

- "I see now that I'm struggling a bit with integration. I find it difficult to understand when to use substitution or integration by parts, for example, and how to use these methods to solve the problem. I managed to determine both volume and area all right, except I forgot that area can't be a negative quantity. I have no problems at all working with geometric and arithmetic series. That's good to know."
- "Now the time has come to connect the dots. Especially in probability theory, I tend to do things without thinking them through – I need to start identifying what I'm supposed to calculate, using a Venn diagram. I also need to develop a better understanding of what the problem is about, and what the result of various calculations

should look like so I'll know how to proceed. Next time I'll draw a Venn diagram so I can see it for myself! It's very logical when you see it. Because then I know what to include and not include. As for integration, which is a subject I know I'm struggling with, I didn't look through the various integration methods, and didn't reflect on what I need to do. So, my calculations were sloppy!"

5.3.4 Assessment procedure: multiple-choice tests

As with the mathematics tests, the multiple-choice tests also started with a teacher introduction (*principle 1*). The content of the introduction was fairly similar, typically giving a presentation of the academic goals and the main purpose of the test to create an opportunity for the students to learn from their misconceptions, make an effort and experience mastery (*principle 5*). The difference from the former test was the presence of learning tools. In addition to self-feedback and reflection, the multiple-choice tests also explored the potential for learning through peer dialogue (*principle 4*). Prior to the test, the students were therefore placed in smaller groups of approximately four members.

After the introduction, the students started working on the test questions individually. Towards the end of the test, they received the various alternatives for the multiple-choice questions and responded by using mobile-phone technology: the student response system (SRS), also known as electronic voting systems and classroom communication systems (Boyle & Nicol, 2003; Draper & Brown, 2004). This technology enables teachers to gain immediate access to student performances. In other words, the teacher gained an overview of the students' understanding (what questions they have answered more or less correctly) and used this as a compass during the interactive dialogic review (*principle 3*).

During the interactive dialogic review, all the questions and their subsequent solutions were presented by the teacher in class and discussed in a plenary session. As with the mathematics tests, the students evaluated and assessed their own performance (*principle 2*) and were repeatedly encouraged to participate in this process (*principle 4*). Moreover, the students also discussed a number of questions with each other in small groups (*principle 4*). They were advised to use the following tools and methodologies during the discussions:

- Share with the rest of the group how you responded to the question and the reasons and arguments behind your response
- Try to get everyone in the group to participate
- Listen to your peers
- Evaluate all responses

- Try to reveal misunderstandings and learn from them together

The discussions were often completed with a second SRS round, where the students anonymously answered a test question once more. In doing this, the students had what can be called a "second chance", an opportunity to learn from their misconceptions in-group and respond again (*principle 6*). At the end of the interactive dialogic review, each student wrote down a final individual reflection (*principle 2*) on their performance (e.g. learning outcome, misconceptions, further effort and peer dialogue) and submitted it to the teacher.

5.4 Part 2. From Two to Five: A Continuation of Established Changes

After the first one-and-a-half school years, the team was expanded, as three more mathematics teachers were added. The collaboration lasted yet another school year. The extension was suggested by one of the original teachers in the team because he wanted to include all the mathematics teachers working in the preparatory engineering courses.

To elaborate, the research conducted in this thesis is based on data collected from preparatory engineering courses at NTNU. There are approximately 50 students in each preparatory course, each lasting for a year. NTNU Trondheim has *five* preparatory courses and each course has its own mathematics teacher. The curriculum, instructions, number of teaching hours (ten hours per week) and assessment practice are the same in all five courses.

The three remaining teachers were positive about joining the team and thus willing to make changes in their own practice. Their point of departure was the same as the initial team teachers, and they were interested and motivated to implement changes in their practice.

The expanded team, now consisting of five mathematics teachers and "author", continued working with the "seven principles of good feedback practice" as a pedagogical framework for applying a formative practice in mathematics, and continued the practice developed by the initial team through the established work methodologies focusing on teacher support, observation, reflection and collaboration. This means that as a team "author" and the teachers met regularly (at 14-day intervals) to discuss and evaluate the assessment sessions and learn from each other's experiences. "Author" also met the teachers for individual reflection conversations. However, due to limited time and resources, the newest team members were prioritised for these conversations.

Finally, the second year's context, with a total of five teachers and their involved students, also constituted the grounds and opportunity for observing the students' achievement goal patterns. These results are presented and discussed in Article 3.

6. METHODS AND METHODOLOGY

As briefly mentioned above, the research in this thesis is conducted as an *intervention study* using a qualitative approach within an *interactive action research* framework, which has further been divided into three sub-studies. In this type of research, the researcher and research participants are equal partners researching practice to develop it (Postholm, 2007). To elaborate:

> If the research work is simply described in a text, there is no guarantee that the text will be read. If it is read, there is no guarantee that changes will take place in the field of practice. This therefore suggests that the practice should be changed while the research is taking place, if the conditions and expectations facilitate this (Postholm 2007, p 15).

Using this framework, the research associated with this thesis can to a great extent be defined within the *constructivist* paradigm. Within this tradition, people are perceived as actively acting and responsible, and knowledge is perceived as a construction of understanding and meaning created in a meeting between people in social interaction. Knowledge is thus not something that is given once and for all, and which is to be transferred, redeemed or revealed; it is constantly changing and renewing (ibid).

Before going into more detail about the phenomenological approach to qualitative research, the choice of research instruments, the data collection, the participants, analytical procedures, and last but not least, the ethical and critical considerations related to this intervention study, this chapter will start with a description of interactive action research. This to provide an overall framework for the way this study has been conducted and the methodology this has entailed in practice.

6.1 Interactive Action Research

When the goal of a study is to change something, the common term for many research areas with different overarching traditions or motives is action research (Johannessen et al., 2016). This means that there is no one way of undertaking action research. Rather, it is described as a "family of approaches", where there will never be one "correct way" to carry out this type of research (Reason and Bradbury, 2008). An important characteristic in all action research, however, is that the focus is on both actions and research on these actions so that thoughts and ideas are united with the practice field both in reflections on practice and testing in practice (Postholm, 2007). One important aim thus is to give understanding through research to create action for change (Johannessen et al., 2016).

Action research has earlier been divided into "technical" and "practical" approaches (Carr and Kemmis, 1986). In the technical approach, research participants are selected by researchers who function as experts in the action research. In research on education, the goal is to increase the efficiency of the education practice. The practical approach also aims to increase efficiency, as in the technical approach, but the development of the practitioners' understanding is also important. Moreover, in the practical approach the researcher functions more as a consultant than an external expert. Beyond this initial division, however, recent researchers (see for example Postholm, 2007 and Svensson et al., 2007) also refers to a third approach for research within the organisation's framework, namely interactive action research where researchers and research participants cooperate on equal grounds to develop the field of practice, while at the time researching this practice. The term "interactive" refers precisely to the relationship between researcher and research participant, and it is this form of action research that has inspired the intervention study here.

Interactive action research has been chosen as the framework for this study because of the necessity in higher education to bring about educational change. The focus on reciprocity, cooperation and important work methods within this framework, such as observation and reflection, was also found useful for implementing important empirical evidence from school research in practice. Procedures from school education, such as the KLT programme, highlight the importance of teacher support, building social capital, critical and responsive learning and inquiry-based collaboration, including using empirical evidence and working together to plan, test, evaluate and implement new ideas and changes. Thus, in accordance with learning practices from school research, an interactive action research framework was considered appropriate for this intervention study that has focused on mutual collaboration and practical procedures as instruments to encourage teachers to engage with and implement principles of formative assessment in practice.

6.1.1 Intervention and research process

Persuasion and initial start-up. The point of departure for interactive action research can be the problem areas the practitioners want to examine, as well as what the researchers want to explore, and the researchers can cooperate on different levels in the school context. A team of teachers, as in this intervention study, is a large enough unit for both research and development activities (Postholm 2007). The research activity in this study was initiated by the researcher, "author", and the teachers participated on a voluntary basis in the cooperation team.

Since the collaborative relationship between the researcher and research participants in an interactive action study is very close, this type of research requires that the researcher has good social competence, and if she is to acquire trust and hence be listened to, she must also have general competence in relation to teaching and learning (Svensson et al., 2007; Postholm and Madsen, 2006). Moreover, successful studies of change in higher education, such as the REAP (Re-engineering assessment practices in higher education) project (Nicol and Draper, 2009), show that persuasion is one of the most important ingredients in any change stage in a development process, and that a constructive change process must be based on a concrete "evidence base". "Author" was aware of these elements from the very start and actively included them in the introductory dialogue and start-up with the teachers.

The most important and initial task was to persuade the teachers that a common development process towards a change in the practice could become both meaningful and constructive for all the parties involved. The teachers could not be told in advance what the outcome of a shared research process would be. To the contrary, the main message was that the real value would be in creating development and change together in a community. Thus, it was important that the teachers felt that any change process would start from below and within, even if the initiative came from "author". However, although the end result could not be foreseen, the teachers could at least be presented with an overriding theoretical and empirical umbrella based on formative assessment and theories on student intelligence and achievement goal stability and change, as well as a concrete educational framework: "the seven principles of good feedback practice" (Nicol and Macfarlane-Dick, 2006), which could be implemented and used as an active intervention tool for educational change.

Finally, it was pointed out that the cooperation and social community would be characterised by different roles (Svensson et al., 2007), where "author" would have responsibility for drawing on theory and educational frameworks, which in turn could open for other and new thoughts and hence development of practice (Postholm, 2007), while the teachers with their many years of experience from teaching and assessment, which "author" did not have at all, would be responsible for applying their professional judgement about what would work for them in their practice, with its variability and complexity, and their teaching style, and finally, implementing the common planned changes in their practices. The hope behind this approach was that these two teachers would be interested in developing their assessment practice together with "author". Both the teachers immediately accepted. They appreciated the idea of a concrete framework and work methods based on cooperation, observation and reflection. It was agreed that the teachers and "author" would be equal

partners cooperating and conducting research on practice, while the goal would be to change and hence create a formative assessment practice for the teachers' assessments in mathematics.

Development and implementation process. An introductory implementation phase in an interactive action research framework may be described as follows:

> In this part of the development activity the dialogue between the researchers and research participants is very important. In the implementation phase of the solution or model, observation of practice and joint reflection between the researchers and teachers based on the observations will go hand in hand and supplement each other. The observations will form the point of departure for dialogue, and the dialogues will in turn give direction and focus for the observations. In this implementation phase the interactive action research is thus process driven (Postholm 2007, p. 28).

Inspired by the interactive action research framework described above, the initial research activity in this study consisted of observations and dialogues that focused on the point of departure, i.e. the existing assessment practice the teachers then had, where there was to be close collaboration in the team and training in practice and the teachers' assessment practice was to focus on how the "seven principles of good feedback practice" could be implemented and contribute to creating a formative assessment practice. The contours of a work process began to form, which established into a permanent work process in the next phase. The initial planning process lasted approximately six months before the implementation process itself commenced.

In the autumn of 2012, the teachers were ready to perform the planned changes and implement the seven principles in practice. The autumn was the start of a new semester with new students. Based on experiences from the planning phase, knowledge from school research, the KLT programme, and important action research tools, the team established the following continuous work process (see Figure 3). (1) The teachers conducted an assessment inspired by the "seven principles of good feedback practice" in their practice together with their students. As a mutual partner, collaborator and pedagogical support, "author" was present during and an observer of all class assessments. (2) Immediately after each assessment, the teachers reflected individually and wrote down their personal experiences of the classroom. (3) Then the teachers had individual reflection conversations with "author" shortly after the various assessments. For each conversation, the teachers brought their own reflection notes, and "author" brought her classroom observation notes. These were used as tools to facilitate reflection. The conversations served as an in-depth summary and evaluation

of the completed assessment, as an arena for discussion of relevant empirical findings and, furthermore, as a planning session for the next assessment. (4) Afterwards, the entire team met to discuss and evaluate the assessment sessions, learn from each other's experiences and perhaps adjust their plans.

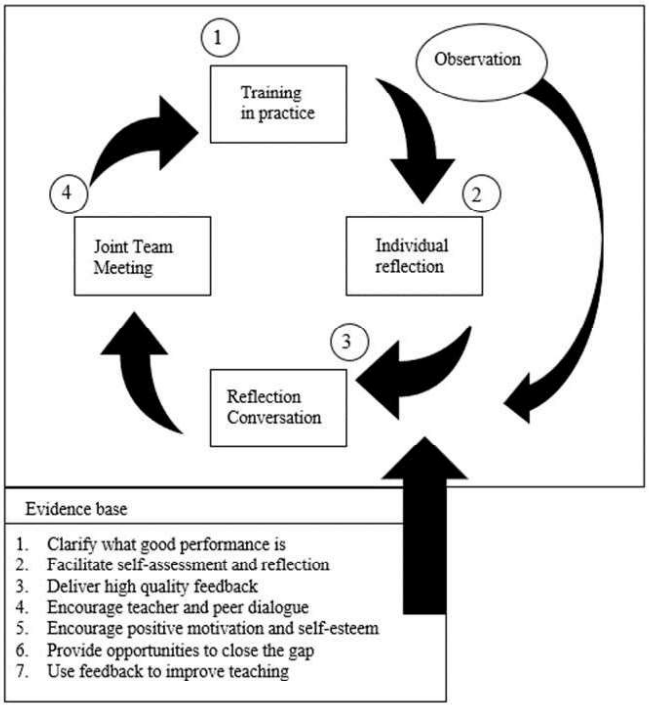

Figure 3. The collaboration team`s continuous work process.

Both the individual conversations and the joint team meetings focused on the teachers' preparation ahead of the assessments, their main objective of creating a formative assessment context, their various choices relating to the implementation of certain principles, their experiences relating to their practice and their experiences of the students. The pedagogical principles were used as a concrete framework the team could use in its work and when comparing their experiences. The principles and key empirical findings relating to them thus functioned as pigeonholes where the teachers could place their own experiences. Finally, the reflections and the subsequent conversations were used as the starting point for the next

assessment in practice, and for observation, which then served as the foundation for the next reflection and conversation. In this way, the implementation phase was process driven.

Towards the end of the implementation phase (the spring of 2013) the work process was shared with other members of the teaching staff on the initiative of one of the teachers in the original team. This resulted in an expansion of the team, adding three more teachers and further cooperation the next school year, using the same continuous and established work process.

6.1.2 Collection of information within an interactive action research framework

It is recommended that both the researchers and research participants use various data collection strategies in this type of process so they will be able to retain experiences, thoughts, preliminary analyses and interpretations, both for their conversations and for the writing of their texts (Postholm, 2007). During the implementation phase in question the various action processes were therefore captured through observation notes, reflection notes, transcriptions and preliminary analyses. As mentioned above, the teachers also kept their own reflection notes. All the information was used as the foundation for further reflections and conversations. Coghlan and Brannick (2005) refer to this type of process as a "primary circle" for the development of teaching. The aim is that the practice, observations and conversations will supplement each other in a continuing change process. A more detailed description of the data collection with its methods and analytical approaches is described in section 6.3.

6.2 A Brief Overview of the Research Work

In sum, through a collaborative process within an interactive action research framework, inspired by learning practices from school education, the team used the "seven principles of good feedback practice" as an intervention tool to facilitate changes in practice and create a formative assessment practice in mathematics that was characterised by active student participation, self-feedback, reflection, peer dialogue and student-teacher dialogue. Through this intervention process the following items were examined:

- how the teachers experienced changing their own practice (Sub-study 1),
- the degree of compliance between the teachers' intentions in using the "seven principles of good feedback practice" and the students' experiences of them in practice (Sub-study 2), and
- the student's achievement-goal patterns within a formative assessment practice in mathematics (Sub-study 3).

The first studies, conducted during the first one-and-a-half years of the two-and-a-half-year intervention study, comprised two mathematics teachers and their students. The students' achievement goal patterns were examined in the second year, where this part of the study comprised five mathematics teachers and their students from a total of five mathematics courses. A more general overview of the three sub-studies, and the different methods and analytical procedures associated with these can be found in Table 3.

Table 3. Overview of the research work.

Studies:	**Research questions:**	**Participants:**	**Data collection methods:**	**Conducted:**	**Analytical procedures**
Sub-study 1 (presented in article 1)	What are the teachers' foremost experiences of their own change process?	Two teachers	17 Semi-structured interviews with two teachers (34 interviews in total).	Spring 2012 Fall 2012 Spring 2013	Grounded theory
Sub-study 2 (presented in article 2)	To what extent are the teachers' beliefs about what they are doing in a dialogic formative assessment concordant with how the students experience it?	Two teachers 16 Students from two preparatory engineering classes of 2012/2013	17 Semi-structured interviews with two teachers (34 interviews in total). Four Focus-group interviews	Fall 2012 Spring 2013 Spring 2013	Grounded theory Grounded theory
Sub-study 3 (presented in article 3)	Which achievement goal patterns do students pursue in a formative assessment practice in mathematics?	Students from five preparatory engineering classes 16 Students from five preparatory engineering classes of 2013/2014	Questionnaire Four focus-group interviews	Fall 2013 Spring 2014 Spring 2014	Preliminary statistical examinations Principal component analysis Differential continuity Mean level change Individual level change Grounded theory

Below, the research instruments, data collection and analytical procedures related to the *qualitative research* conducted in this intervention study (associated with all three sub-studies) will be described. Then the quantitative method in relation to the use of mixed

method design (used in Sub-study 3), and the data collection, analytical procedures and initial statistical results associated with this research work will be presented.

6.3 Qualitative Research Methodology

Interactive research – like action research – does not represent a particular method; it is more a question of an approach that can comprise several different methods (Svensson et al., 2007). To gain an in-depth understanding of how the teachers experienced changing their own practice, their intentions and understandings, and not least, the student views and perspectives, a *qualitative methodology* with a phenomenological approach to research was adopted in the form of two qualitative case studies using semi-structured interviews and focus-group interviews. The aim of this research was thus to understand the research participants' perspectives (Schutz, 1972; Postholm, 2005; Kvale and Brinkmann, 2009).

A special feature of the qualitative method is that the researcher remains close to the informants and is more subjective in relation to the perceptions of the informants. This means that qualitative research creates a better basis for in-depth understanding of the phenomena in question (Lund and Haugen, 2006) and for a holistic understanding (Johannessen et al., 2016). The researcher wants to form a comprehensive picture of the participants' perspectives when it refers to a particular research focus (Postholm, 2010). The sample will therefore consist of a limited number of informants, and the focus will be on the meaning of and understanding the phenomenon (Robson, 2011). The researcher is presented as the most important research instrument in qualitative research, as the researcher must maintain an interpretative role throughout the research process (Postholm, 2010).

6.3.1 A phenomenological perspective

According to Kvale and Brinkmann (2009), a *phenomenological perspective* is based on the individual's experiences: How the person experiences a situation, and what he/she feels and thinks is important in the current situation. This means that the phenomenological perspective is based on the understanding that reality is the way the informant perceives it. The aim is therefore to understand and highlight social phenomena from the informant's perspective, to illuminate the informant's exact description and to obtain the core meaning (ibid).

This can be done by the researcher focusing on the participants in their natural context (Johannessen et al., 2016). The research will nevertheless be colored by her own theoretical point of view and experiences, and this must be shown in the research (Postholm, 2005). In

this study, this has been done through, first and foremost, presenting the researcher's scientific point of view, which was done at the start of this chapter, and further pointing out that theoretical choices and selections from the data material are the researcher's subjective choices about what should be focused on in the research.

6.3.2 Teacher experiences: semi-structured interviews

Semi-structured interviews were found to be a suitable in-depth method for obtaining knowledge about the teachers' own experiences, intentions and opinions. This is also the most commonly used data collection strategy within phenomenological research (Postholm, 2005).

Kvale and Brinkmann (2009) define a semi-structured interview as "a planned and flexible conversation aiming to obtain descriptions of the interviewee's lifeworld with the intention of interpreting the meaning of the phenomena described" (p. 325). Thus, a semi-structured interview is used if the aim is to obtain descriptions of how people understand their lifeworld. It is not unlike an everyday conversation, but the professional interview aims to collect data where special approaches and techniques are necessary (Kvale and Brinkmann, 2009). For this reason, this interview method was a natural choice for the study here. By choosing this type of interview the aim was to have an unbiased and equal relationship, in that the interview would typically be conversations with the informants. The intention was to create trust and confidence in the researcher, "author".

A semi-structured interview has an overall interview guide as the point of departure for the interview, while the questions, themes and their order can be varied, and the interviews can move back and forth between the topics (Johannessen et al., 2016). In this way the teachers were asked approximately the same questions, while it was also possible to ask questions based on their descriptions and statements for further elaboration and clarification.

6.3.3 Student experiences: focus-group interviews

To illuminate the students' experiences of the formative assessment practice in their mathematics classes, focus-group interview was conducted. Focus-group interviews were chosen primarily because this type of interview lends itself well as an independent method for acquiring the views of informants on various topics (Johannessen et al., 2016). Furthermore, a focus-group interview is a type of group interview where conversation and discussion processes are key elements. One of the advantages of focus-group interviews is that when handled properly, they may become extremely dynamic (Berg, 2007). Such a group dynamic is often described as a synergistic group effect, an effect which allows the participants to

continue building on what others have stated or to take part in collective brainstorming with the other participants. The goal behind choosing this type of interview was thus to exploit the meaning potential that might be inherent in the student group.

6.3.4 Participants

To protect the identity of the teachers participating in the research they will not be described in more detail here, except to mention that they are Norwegian nationals and have many years' experience of teaching mathematics. The students, in turn, were part of a larger student group and can thus be described in more detail.

Interviews were conducted with students, both the first (for Sub-study 2) and the second (for Sub-study 3) year of this intervention study. Given that preparatory engineering is a one-year study, it was not the same group of students who participated in both years, but the students who studied the current school year. Between 15 and 20 students volunteered as participants from each class, both years, meaning that the sample comprised students from the preparatory courses of the teachers involved. The only selection criterion established was that the students had attended the formative assessments in mathematics. The starting point for the selection of student participants was thus convenience, also called strategic selection (Johannessen et al., 2016).

Furthermore, the mean age of the students was about 24 years within the 19 to 33 range. They were evenly represented in terms of gender, and all were Norwegian nationals. Moreover, they had a great variety in backgrounds (including theatre, photography, literature and sociology; some of them had previous education, such as electricians and carpenters, while others came directly from high school) and a common goal of becoming an engineer (within various disciplines).

6.3.5 Data collection procedures

The qualitative data collection process can be summarised as follows:

- Semi-structured interviews with teachers were carried out regularly throughout the planning and implementation phase (spring 2012- spring 2013) within an interactive action research framework. All in all, the data collection process for the teacher viewpoints amounts to 17 semi-structured interviews with two mathematics teachers (34 interviews in all), and 17 classroom observations (34 observations in all).
- Four semi-structured focus-group interviews, with four students in each group, were conducted with the students at the end of both the first (2013) and second (2014)

spring semesters. All in all, the data collection process for the student viewpoints amounts to eight focus-group interviews.

All qualitative data material was digitally recorded, subsequently transcribed verbatim, and both the teachers and the students were given pseudonyms to protect their identity. More detailed descriptions follow below.

Semi-structured interviews. The semi-structured interviews with the teachers were completed shortly after the various formative assessments and served as interactive reflective conversations in which the teachers described and reflected on their assessment practice and responded to various probing questions. Each interview lasted about one, to one and a half hours. For each interview, the teachers brought their reflection notes which they wrote immediately after the formative assessments, while "author" brought her classroom observation notes. These were used as a tool to facilitate reflection on the various interview questions.

Because these were semi-structured interviews, the same structure was used in all the conversations, but with ample room for flexibility about what the teachers highlighted as important and wished to share. For example, each conversation started in the same way, with an open question about their last assessment in the classroom with their students. The aim was that the teachers would be able to initiate the conversation with what they had reflected on, found important and wanted to talk about. This would place their "lifeworld" at the forefront of the interview. The perceptions raised by the teachers were elaborated on and would often be present throughout the interview. An interview guide was also developed (see Appendix 3), where the main points were the teachers' preparations ahead of the formative assessments, their main objectives, their various choices relating to the implementation of particular principles, and their experiences relating to their practice, their role and their students, and lastly, their plans for the next formative assessment. After this part of the conversation was finished, and the more or less fixed focuses had been discussed, "author" shared her observation notes which were then discussed and reflected on together. Each conversation was concluded with a plan for the next assessment.

In sum, the interviews served as an in-depth summary and evaluation of the completed formative assessment, and furthermore, as a planning session for the next formative assessment.

Focus-group interviews. The students were randomly distributed into four groups. The students were informed prior to the interview that their participation was entirely voluntary, and they could withdraw at any time, or if they wished, at a later date they could

have their data withdrawn from the study. All the interviews were held at the then Sør-Trøndelag University College, now NTNU, which was the place of study for the informants in question. All the interviews were tape recorded and lasted from one and a half to two and a half hours.

As the facilitator of the focus-group interviews, "author" used a semi-structured question guide (see Appendix 4). The same guide was used with the students both years, as the assessment practice for 2013/2014 was a continuation of the practice introduced the year before (2012/2013). It is important to point out that the students were not aware that the teachers were using the "seven principles of good feedback practice" as pedagogical inspiration for their own practice. The guide was therefore not related to these seven principles and the students were instead asked open questions about how the assessments in mathematics were undertaken and how they experienced them. The aim was that what the students highlighted as interesting and important would be discussed in the focus groups.

6.3.6 Analytical approaches

It is possible to be inspired by various analytical approaches to find the approach that is best suited according to its unique data material (Postholm, 2005). As this part of the research builds on a qualitative method with a phenomenological approach, analytic strategies were applied that followed this thinking and allowed "author" to be inspired by the "constant comparative method", where coding and categorisation have a key place (ibid).

More concretely, all the qualitative data material, both the conversations (semi-structured interviews) with the teachers and the focus-group interviews with the students were analysed by means of the constant comparative method, also called *grounded theory*. A method developed by Strauss and Corbin (1998), it involves structuring and seeing patterns in the data material to collect information that belongs together. Grounded theory is an ideal method for analysis of qualitative data as it enables researchers to generate theories stemming from experience, i.e. theories that build on empirical findings (Charmaz 2001). Grounded theory is a suitable analytical approach for several topics, such as personal experiences, feelings and attitudes (Charmaz, 2003). As an analytical approach, it offers both proximity and structure, two factors that were deemed essential in relation to the goal of getting close to and exploring the intentions and experiences of the participants with respect to the development and application of formative assessment.

The interviews were analysed according to an approach suggested by Kathy Charmaz (Charmaz 2001, 2003). This means that the analytical steps used were coding line by line,

focused coding and categorisation. It is important to add that grounded theory was used as a *method of analysis* to carry out a thorough analysis and develop a rich description of the perceptions of the participants. More precisely: grounded theory was used as a method, in contrast to grounded theory as a product of the method. The analysis has thus *not* aimed for concurrent involvement, theoretical sampling or theoretical saturation.

Line-by-line coding is the initial coding used in the analysis. This means that each line of data is examined and events that stand out or are representative are then defined. In-vivo codes were regularly developed in the analysis. This means that the words and formulations of the participants were used to develop codes rather than interpreting the statements made by the participants using their own concepts. Below are two examples from the initial line-by-line coding. These are taken from the analysis of the student interviews in Sub-study 2:

Example 1:

Emma: I may not know how much I, I learn, really. But anyway I'm becoming aware of my own situation in precisely this subject (Line-by-line coding: *aware of my own situation in the subject*). How good am I? (Line-by-line coding: *aware of how good I am*). Where am I making mistakes? (Line-by-line *coding: aware of why I make mistakes*). Well, that's what I'm learning, then (Line-by-line coding: *learning to become aware*). I'm learning what I'm good at and not good at, what I need to do the next time (*learning what I'm good at and not good at*). I may not learn so much right there and then about that there (Line-by-line coding: *not necessarily learning from the curriculum*). It's more like learning about my own situation, really (Line-by-line coding: *learning about her own situation*).

Example 2:

Eva: You know, if we had been given the test back after a week, then there wouldn't have been any reflection at all, I believe. At least not by me (Line-by-line coding: *late feedback - no reflection*)
Mona: No, me neither (Line-by-line coding: *late feedback - no reflection*). It's imperative to do it right afterwards. You know, you get much more involved in your own mistakes, in a way, or what you did right, if you reflect on it right away (Line-by-line coding: *immediate reflection – more engaged in own performance*). A week later you barely remember what the test was about (Line-by-line coding: *late feedback – barely remember the test content*). Then there's nothing to reflect on (Line-by-line coding: *late feedback - weak memory → nothing to reflect on*).

Focused coding is the next step in the analysis. The aim of this phase is to arrive at a smaller number of codes which represent a larger portion of the data material (Charmaz, 2001). In brief this part of the analysis consists of reviewing the data material again, based on the line-by-line coding, and determining whether some of the codes recur and represent a larger portion of the data material. They are then "elevated" up to focused codes and used as the point of departure for the final phase of the coding, categorisation. To show how the

analysis progressed from line-by-line coding to focused coding, the same examples are used that were presented above:

Example 1:

Emma: I may not know how much I, I learn, really. But anyway I'm becoming aware of my own situation in precisely this subject (Line-by-line coding: *aware of my own situation in the subject*). How good am I? (Line-by-line coding: *aware of how good I am*). Where am I making mistakes? (Line-by-line *coding: aware of why I make mistakes*). Well, that's what I'm learning, then (Line-by-line coding: *learning to become aware*). I'm learning what I'm good at and not good at, what I need to do the next time (*learning what I'm good at and not good at*). I may not learn so much right there and then about that there (Line-by-line coding: *not necessarily learning from the curriculum*). It's more like learning about my own situation, really (Line-by-line coding: *learning about her own situation*).

Focused coding: *awareness as learning* (learning to become aware of her own situation in the subject).

Example 2:

Eva: You know, if we had been given the test back after a week, then there wouldn't have been any reflection at all, I believe. At least not by me (Line-by-line coding: *late assessment - no reflection*)
Mona: No, me either (Line-by-line coding: *late assessment - no reflection*). It's imperative to do it right afterwards. You know, you get much more involved in your own mistakes, in a way, or what you did right, if you reflect on it right away (Line-by-line coding: *immediate reflection – more engaged in own performance*). A week later you barely remember what the test was about (Line-by-line coding: *late assessment – barely remember the test content*). Then there's nothing to reflect on (Line-by-line coding: *late assessment - weak memory → nothing to reflect on*).

Focused coding: *immediate reflection*

More specifically, in these examples the different line-by-line codes are summarised, and "author" has endeavoured to point out whether some of them recur, thus describing in a better way what the paragraph is about. "Author" finds that the codes "immediate reflection" and "awareness as learning" (learning to be aware of one's own situation in the subject) describe the texts well. Instead of dealing with several line-by-line codes, the paragraph is summarised in one code. The reason for choosing, for example, "awareness as learning" (learning to be aware of one's own situation in the subject) as the focused code is that "author" interprets that the paragraph basically refers to the student's own learning as awareness. The student explains that what she is learning is not necessarily about learning the mathematics subject, the way we traditionally think of learning as acquiring greater understanding of a particular topic. The learning experience this student had was focused on becoming aware of what she knows or does not know, in her words, "of my own situation". These types of description of the focused codes were outlined in memos.

Categorisation represents the third and final step of the analysis and involves more analytical and conceptual processing (Charmaz (2001). Only the codes which appear to be essential in terms of the perceptions of the participants, and which stand out in relation to other codes, are elevated into categories. These are further defined and examined thoroughly: what do they contain, what do they mean and when and what do they influence, who do they relate to and how, and not least, do they illuminate the perceptions of the participants in relation to development and use of formative assessment?

In the analysis examples above, from Sub-study 2, there were two focused codes which stood out clearly in the analysis of the student interviews and were elevated into categories, "reflection" and "social relations". The category "reflection" is based on the focused codes described above. The categories that are connected to the different qualitative analyses are described in detail in Chapter 7 and in the three attached articles.

All the analytical work was carried out in written memos. Memo writing may simply be described as the step between coding and the first draft of the completed analysis. In accordance with the recommendations by Charmaz (2001), much of the memo writing was focused on comparing the data with the categories, and the categories with each other. This was done to define the characteristics of the categories, and to determine how and when they were possibly linked to each other.

6.4 Sequential Explanatory Mixed-Method Design

To observe the students' achievement-goal patterns within a formative assessment practice, a sequential explanatory mixed-method design, with three complementary data-analytical approaches and focus-group interviews, was used. A sequential explanatory mixed-method design is a procedure for collecting, analysing and "mixing" or integrating both quantitative and qualitative data at some stage of the research process within a single study so a better understanding of the research problem can be gained (Creswell 2003). The sequential explanatory design also involves collecting and analysing quantitative and qualitative data in two consecutive phases within one study (Ivankova et al., 2006). The current study aimed to examine which achievement goal patterns the students pursued in the newly established assessment practice in mathematics, and finally, to gain a broader impression of the assessment practice, focus-group interviews were conducted with the students at the end of the school year.

The research proceeded in the following manner:

- One questionnaire measuring the students' achievement goals was given during two semesters (the fall semester of 2013 and spring semester of 2014). The questionnaire was administered six times during the two semesters (three times per semester).
- As mentioned above, four semi-structured focus-group interviews, with four students in each group, were conducted at the end of the spring semester of 2014.

6.4.1 Achievement goal measures

Mastery-approach goals, performance-approach and performance-avoidance goals were measured through the student version of PALS (Midgley et al., 1998), a 17-item achievement-goal questionnaire. In the scale, six items assessed mastery-approach goals (see item 4, 11, 18, 25, 32 and 40 in Appendix 5), five items assessed performance-approach (item 9, 16, 23, 30 and 35), and six items assessed performance-avoidance goals (item 5, 12, 19, 22, 26 and 37). Participants responded to the items on a scale ranging from one (not at all true for me) to four (very true for me). Each sub-scale was estimated as the mean of the individual item scores.

The items were ranged from 1 to 4 following considerations based on a small pilot study conducted in a preparatory engineering mathematics class prior to the current study. In the pilot study, the scale ranged from 1 to 5, with the majority of the students placing their answer in the middle of the scale. Follow-up interviews with the students revealed that many of them did not read the questions, but automatically placed their answer in the middle to avoid taking a stand. The scale range was therefore converted to four response categories.

Furthermore, the old version of PALS (1998) was found to be more suitable for translation from English into Norwegian than the later version (PALS, 2000). After translating both versions, the 1998 version appeared clearer and more specific. For the purpose of this study, the wording of the items was specified to capture the students' thoughts about themselves in mathematics instead of more general statements. For example, the statement "I like classwork that I'll learn from even if I make a lot of mistakes" was altered to "I like working with mathematics assignments that I learn from, even if I make a lot of mistakes". Moreover, the translation of the items from English into Norwegian also led to small adjustments.

Finally, in addition to the items concerning achievement-goals, four other scales were included in the questionnaire. In the remaining scales four items assessed anxiety (see item 10, 17, 24 and 31 in Appendix 5), six items assessed academic self-concept (see item 6, 18,

20, 27, 33 and 39), six items assessed interest (item 7, 14, 21, 28, 34 and 38), and four items assessed fear of failure (item 8, 15, 29 and 36). These scales were inspired by validated scales from Houston and Kelly; 1987; Herman, 1990 and Harackiewicz et al., 2000, and were included to make the questionnaire more extensive, while also making the scales of achievement goals less obvious for the students. In all, the questionnaire consisted of a total of 40 items.

6.4.2 Measurement procedure

The achievement-goal questionnaire was given three times (T1-T3) during the fall and spring semesters according to a set procedure. Just before the students participated in an assessment, which could be a mathematics test, a multiple-choice test or a mock exam, they were asked to sign a consent form and respond to the attached questionnaire. After completing the questionnaire, the mathematics assessment began. Before the first data collection, the students were informed about giving their written consent.

During the fall semester, the first achievement-goal questionnaire (T1) was given at the beginning of the semester before the students had had any assessments and received subsequent performance feedback. This was given before a mathematics test. The second achievement-goal questionnaire (T2) was given six weeks later, during midterm, before a mathematics test. The third achievement-goal questionnaire (T3) was given at the end of the semester, nine weeks after the second measurement and prior to a multiple-choice test.

During the spring semester, the first achievement-goal questionnaire (T1) was given at the beginning of the semester, just two weeks in, before a mathematics test. The second questionnaire (T2), however, was given eleven weeks later, during midterm, before a mock exam. The third questionnaire (T3) was given at the end of the semester, four weeks after the second measurement and prior to a mathematics test.

All in all, the achievement-goal questionnaire was administered six times over two semesters (three times per semester) in the same student classes, but with a different number of students per semester. The purpose of administering the questionnaire in both the fall and spring semesters was to gain a more comprehensive understanding of the students' achievement-goal patterns. Earlier studies suggest that the first achievement-goal measurements at the beginning of the fall semester are the most important ones in terms of change. The purpose of investigating the fall semester was thus to examine stability and change within a formative assessment practice during the beginning of a school year. However, although the first semester has previously been shown to exhibit the greatest

achievement-goal changes, few studies investigate students' achievement-goal patterns during the spring semester. This study therefore includes both semesters in the school year. Moreover, a different type of assessment was included in the spring semester. A mock exam was included to involve an assessment situation that is closer to the final exam and the assessment structure used in previous studies on achievement goal stability and change. Therefore, in order to observe the students' achievement-goal patterns thoroughly, and through different assessment contexts, it was crucial to include both semesters in the study.

Finally, to gain a broader impression of the assessment practice undertaken in the third Sub-study, focus-group interviews were conducted with the students at the end of the school year.

6.4.3 Analytical approaches

To measure achievement-goal stability and change, Fryer and Elliot's (2007) analytical approach was applied, and three of the four complementary analytical procedures presented in their studies were also used: differential continuity, mean-level change and individual-level change. These procedures were applied in SPSS, version 22.0. The intention behind these analytical procedures was to use the statistics to observe and describe the students' achievement-goal patterns in a formative assessment practice. However, it is important to point out that as the aim was to explore these observations together with the students' interviews and existing research literature in the fields of achievement-goal stability and change, and formative assessment, the analysis presented in this study is thus descriptive.

Differential continuity represents the level of rank-order consistency maintained in a construct over time within a sample and is measured by calculating Pearson product-moment correlations (Fryer and Elliot, 2007). High test-retest reliabilities for differential continuity provide evidence of goal stability (little change in intensity), whereas moderate-to-low-reliabilities signify moderate to high changes in goal intensity (Muis and Edwards, 2009).

Mean-level change represents the degree to which the average amount of a construct changes over time within a sample (Fryer and Elliot, 2007) and is measured with paired-sample t-tests. Significant mean changes in goals suggest changes in goal intensity; the greater the difference between scores, the more change in goal intensity (Muis and Edwards, 2009).

Individual-level change represents the magnitude of increase or decrease in a construct over time exhibited by an individual (Fryer and Elliot, 2007) and is analysed using a *reliable change index* (RCI) (Christensen and Mendoza, 1986; Jacobson and Truax, 1991).

The RCI allows for an assessment of whether an individual shows a significant increase, decrease or no change in scores from one time to the next. RCI values lower than 1.96 or higher than 1.96 are unlikely to occur by chance and are thus considered indicative of reliable change. If change is random, the distribution of RCI values should be normal, with approximately 2.5% of values below -1.96, approximately 2.5% of values above 1.96 and approximately 95% of values between -1.96 and 1.96 (Fryer and Elliot, 2007).

6.5 Preliminary Statistical Examinations, Considerations and Results

6.5.1 Principal component analysis

Principal components analysis (PCA) was used with an oblique rotation to ensure the structural stability of the PALS questionnaire for current samples (from the fall and spring semesters) before checking for mean level change. The possible factors were extracted using Kaiser's criterion and scree plot. Three PCAs were performed during both semesters to assess the goal measurements across the time periods.

Principal components analysis (PCA) and factor analysis (FA) are statistical techniques applied to a single set of variables when researchers are interested in discovering which variables in the set form coherent subsets that are relatively independent of each other (Tabachnick and Fidell, 2007). Variables that are correlated with one another but are largely independent of other subsets of variables are combined into factors. These factors are thus thought to reflect underlying processes that have created the correlations between variables (Tabachnick and Fidell, 2007). Although this study used a PCA, it is common to use the term factor analysis as a collective term for both PCA and FA (Field, 2009; Johannessen, 2009; Tabachnick and Fidell, 2007). The term factor analysis will also be used in this study, unless specification is considered necessary in some places.

6.5.2 Analytical assumptions

Prior to the actual analysis, several analytical assumptions were examined to clarify whether it was possible to perform this kind of analysis on the current data material. Five assumptions were included; sample size, number of variables, normal distribution, linearity and correlation.

Sample size. First and foremost; the reliability of a factor analysis is dependent on sample size (Field, 2009; Johannessen, 2009) as correlation coefficients may be less reliable when estimated from small samples (Tabachnick and Fidell, 2007). According to Tabachnick and Fidell (2007), a sample of 300 can be a satisfactory number for a factor analysis.

However, smaller samples can be accepted when they have high correlations and few distinct factors. The data collection for the fall semester was from a total of 190 participants, with three factors, suggesting that the data material is suitable for a factor analysis, while the data collection for the spring semester was from a total of 96 participants, also with three factors. Another way to consider an appropriate sample size is to look at the communality. A simple rule of thumb is the lower communality, the higher the sample size (Field, 2009). The average communality levels in the analysis for the fall semester are above 0.60 (T1); 0.61 (T2); 0.62 (T3), and the average communality levels for the spring analysis are above 0.6; 0.67 (T1); 0.75 (T2); 0.72 (T3), which are considered sufficient for small samples (Field, 2009).

Number of variables. Another assumption that must be clarified in advance of a factor analysis is the number of variables. An analysis should have a certain number of variables and must also have certain characteristics (Johannessen, 2009). First, there must be at least three variables to conduct a factor analysis, and the variables must be continuous (ordinal, interval, ratio level) with a minimum of four values. Given that this analysis contains a total of 17 variables, with four values measured at the ordinal level, this assumption was also rated as satisfactory.

Normal distribution. In addition to the number of variables, it is also important to consider whether the variables are normally distributed. According to Tabachnick and Fidell (2007), a factor analysis benefits from a normal distribution. This is especially important for generalisation beyond the research project's data sample (Field, 2009). The assumption of normal distribution was initially examined visually using histograms with a normal distribution curve and PP plot (probability-probability plot). Most of the variables, from both semesters, were considered virtually normally distributed. Moreover, a violation of the assumption of normal distribution does not necessarily make results less appropriate when using an exploratory factor analysis (Tabachnick and Fidell, 2007).

To obtain a further measure of the normal distribution beyond the visual aspect the variables were also examined using *skewness* and *kurtosis* measures. Skewness refers to the symmetry of the distribution; a skewed variable is a variable whose mean is not the centre of the distribution (Tabachnick and Fidell, 2007). Kurtosis refers to the peakedness of a distribution; a distribution is either too peaked (with short, thick tails) or too flat (with long, thin tails). When a distribution is normal, the skewness and kurtosis values are zero. Positive values of skewness indicate a pile-up of scores on the left of the distribution, whereas negative values indicate a flat and light-tailed distribution. Positive kurtosis values indicate a pointy and heavy-tailed distribution, whereas negative values indicate a flat and light-tailed

distribution (Field, 2009). In sum, the further the values are from zero, the more likely it is that the data material is not normally distributed (Field, 2009).

In the fall semester, the skewness values ranged from -.879 to 1,677 (T1); -.645 to 1,423 (T2) and -.706 to 1,015 (T3), while the kurtosis values ranged from -1,057 to 2,288 (T1); -989 to 1,144 (2) and -1,211 to .600 (T3). The results showed a moderate negative skewness, indicating an accumulation of scores in the "higher" end of the scale (very true). Furthermore, the results showed negative kurtosis values for most of the variables, which implies that the distribution is flat, with many cases in the tail (Tabachnick and Fidell 2007). It was not considered necessary to transform any of the values, which is recommended for substantial skewness and kurtosis (ibid).

In the spring semester, the skewness values ranged from -.671 to 1,136 (T1); -.850 to .971 (T2) and -.696 to 1,294 (T3), while the kurtosis values ranged from -1,288 to .863 (T1); -1,108 to .812 (2) and -1,193 to 1,248 (T3). The variables displayed both negative and positive skewness values, with negative skewness indicating an accumulation of scores in the "higher" end of the scale (very true), whilst positive skewness can be interpreted as a collection of scores in the lower part of the scale (very untrue). Finally, the results showed negative kurtosis values for most variables, which implies that the distribution is flat, with many cases in the tail. It was not considered necessary to transform any of the values.

Linearity. Normality is also part of the assumption of linearity (Tabachnick and Fidell 2007; Johannessen 2009). An analysis can be undermined when linearity fails (ibid). Thus, the use of factor analysis assumes linear relationships between variables. The assumption of linearity was initially investigated by examining the correlations between the variables, as there must be a statistical correlation between variables to conduct a factor analysis (Johannessen, 2009). Correlation was examined using Pearson r. There are two potential problems here, too high and too low correlations (Field, 2009). It is recommended to exclude variables that have low correlations (less than .03) with other variables in the analysis (Field, 2009). Most of the variables in the fall analysis, and all the variables in the spring analysis had a correlation above .03 with the other variables in the matrix.

Correlation. To obtain a more objective measure of the relationship between the variables, two additional statistical tests were performed: *Kaiser-Meyer-Olkin* (KMO) and *Bartlett's test sphere*. KMO investigates partial correlations ranging from 0 and 1. To implement an adequate factor analysis, the minimum value of KMO is 0.60 (Johannessen, 2009). Bartlett's test examines a null hypothesis that the correlations in the correlation matrix can derive from a population in which all correlations are zero. If the variables are not

correlated, they will be completely independent of each other, and it will therefore not be possible to find some subsets of variables that are correlated. It is therefore desirable to reject H0 to continue with the analysis.

The results from the fall analysis showed KMO values of .77 (T1); .78 (T2); .82 (T3), and a significant Bartlett's test at 0.00 for all three time points, while the results from the spring analysis showed KMO values of .78 (T1); .81 (T2); .81 (T3), and a significant Bartlett's test at 0.00 for all three time points. This can be interpreted as confirmation that the variables are suitable for a factor analysis. It must be noted, however, that Bartlett's test, like all significant tests, depends on sample size (Tabachnick and Fidell, 2007; Field 2009).

However, Pearson's r only reports linear relationships between variables. If there are non-linear relationships between the variables, these are ignored (Tabachnick and Fidell 2007). The assumption of linearity can therefore be further examined visually through a scatter plot. In studies with multiple variables, as in this study, it is recommended to look at the skewness values of the variables to assess which variables should be examined, and to examine the variable pairs that are most likely to deviate from linearity (Tabachnick and Fidell 2007). The skewness values for the variables in this study, however, were initially low, which may be an indication of linearity between the variables.

Based on the results from the various preliminary examinations of the analytical assumptions, including sample size, number of variables, normal distribution, correlation and linearity, the data material from both semesters was assessed as suitable for further use in a factor analysis.

6.5.3 Results fall semester

The main question in the factor analysis (FA) is how many factors are needed to explain the correlations between the variables. The first step in a factor analysis is therefore to extract statistically significant factors. There are two major types of FA: exploratory and confirmatory. The current study performed an exploratory factor analysis with an oblique rotation (so that factors themselves are correlated). The extraction was conducted using Kaiser's criterion and scree plot.

Through all three time points, three factors had an eigenvalue above Kaiser's criterion of 1. In addition to the eigenvalue criteria, a scree-plot was used to visualise the number of possible factors. The scree plot supported a selection of three factors. The curve had a characteristic shape, with three factors before the curve broke and levelled off to the right. Kaiser's criterion and scree plot are believed to act as reliable selection criteria when less than

30 variables are involved in the analysis and the sample size exceeds 250 (Field, 2009). Kaiser's criterion can also be considered to be accurate when the average level of communality is greater or equal to 0.6 after the extraction (Field, 2009), which was confirmed in the preliminary analyses on assumptions.

To obtain a clearer picture of which variables were related to the various factors, the extraction was followed by an oblique factor rotation. The purpose of a factor rotation is to maximise high correlations between variables, and minimise low ones (Tabachnick and Fidell, 2007). Rotation contributes to a clearer pattern in that the variables will have high loadings on one factor and lower loadings on others (Johannessen, 2009). There are basically two ways to perform a rotation: orthogonal or oblique. The former is used if the factors are not correlated, while oblique rotation is appropriate if the factors can be correlated to each other (Tabachnick and Fidell 2007). According to Field (2009), the choice between these rotation methods depends on whether there is a good theoretical reason for assuming whether or not the factors are correlated. Oblique rotation should thus only be used if there are good reasons to assume that the underlying factors can be related to each other in a theoretical sense. In this study there are good theoretical reasons to believe that the factors are related. Moreover, there is also reason to doubt whether orthogonal rotation can be performed on naturalistic data, particularly when the data material refers to people, where most processes are somehow intertwined (Field, 2009).

Three initial analyses were performed to reveal the eigenvalues of potential factors. The analysis resulted in four items being removed from their original scale: two items assessing the mastery-approach goal, one item assessing the performance-approach goal and one item assessing the performance-avoidance goal. These were items that had high factor loadings on more than one factor, and their removal reduced the number of items measuring students' achievement goals from 17 to 13.

Through all three points in time, three factors had an eigenvalue above Kaiser's criterion of 1 and explained a total of 60% (T1); 61% (T2) and 63% (T3) of the total variation in the variables. The first factor was assumed to represent *performance-avoidance goals*, the second factor addressed *mastery-approach goals* and the third factor addressed *performance-approach goals*. To elaborate, table 4, 5 and 6 presents the factor loadings after the oblique rotation for T1, T2 and T3. The tables are based on the pattern matrix, as this is easier to interpret (Tabachnick and Fidell, 2007).

Across all three time points (T1-T3), the factor loadings showed strong relationships between the separate variables and the factors. As a rule of thumb, only variables with

loadings of .32 and above should be interpreted. The greater the loading, the more the variable is a pure measure of the factor (Tabachnick and Fidell, 2007). According to Tabachnick and Fidell (2007), loadings in excess of <.71 can be considered as "excellent", which implies that the majority of the loadings in this analysis can be categorized as excellent. A minority of the variables had loading in excess of .61, which can be regarded as very good (Tabachnick and Fidell, 2007). One variable had a loading in excess of.527, categorized as good, and one of .47, categorized as fair (Tabachnick and Fidell, 2007). It is important to note that the significance of a loading is dependent on sample size (Field, 2009). For large sample, even small loading can be considered as statistically significant. It is therefore recommended to only interpret loadings above 0.4 as significant. All factor loadings in this analysis are above 0.4 and are therefore considered as strong significant loadings.

Table 4. Summary results of PCA with oblique rotation **-** T1.

Variables	Rotated factor loadings		
	PAV	MAP	PAP
Jeg jobber med dette faget for å unngå at de andre i klassen skal tro jeg er mindre smart	.859		
En viktig grunn til at jeg jobber med matematikkfaget er at jeg ikke vil dumme meg ut	.818		
Et av mine viktigste mål i dette faget er å unngå å se ut som jeg har vanskeligheter med å løse matematikkoppgavene.	.778		
Grunnen til at jeg jobber med matematikkfaget er for at læreren min ikke skal tro at jeg kan mindre enn de andre	.635		
En grunn til at jeg ikke deltar aktivt i matematikktimene er at jeg vil unngå å fremstå som mindre smart	.615		
En viktig grunn til at jeg jobber med matematikkfaget er fordi jeg synes det er artig		.856	
Jeg jobber med dette faget fordi jeg er interessert i matematikk		.853	
En viktig grunn til at jeg jobber med matematikkfaget er at jeg liker å lære noe nytt		.659	
Jeg liker å jobbe med matematikkoppgaver jeg lærer av, selv om jeg gjør mange feil		.644	
Jeg føler meg vellykket hvis jeg gjør det bedre enn de fleste andre studentene			-.854
Det er viktig for meg å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.762
Jeg ønsker å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.756
Jeg føler meg veldig bra hvis jeg er den eneste som kan svare på lærerens spørsmål i matematikktimen			-.720
Eigen value	4,89	3,29	1,88
% total variance	26.20	22.32	11.41
α.	.80	.76	.79

Table 5. Summary results of PCA with oblique rotation – T2.

Variables	Rotated factor loadings		
	PAV	MAP	PAP
Jeg jobber med dette faget for å unngå at de andre i klassen skal tro jeg er mindre smart	.763		
En viktig grunn til at jeg jobber med matematikkfaget er at jeg ikke vil dumme meg ut	.768		
Et av mine viktigste mål i dette faget er å unngå å se ut som jeg har vanskeligheter med å løse matematikkoppgavene.	.639		
Grunnen til at jeg jobber med matematikkfaget er for at læreren min ikke skal tro at jeg kan mindre enn de andre	.700		
En grunn til at jeg ikke deltar aktivt i matematikktimene er at jeg vil unngå å fremstå som mindre smart	.799		
En viktig grunn til at jeg jobber med matematikkfaget er fordi jeg synes det er artig		.867	
Jeg jobber med dette faget fordi jeg er interessert i matematikk		.838	
En viktig grunn til at jeg jobber med matematikkfaget er at jeg liker å lære noe nytt		.728	
Jeg liker å jobbe med matematikkoppgaver jeg lærer av, selv om jeg gjør mange feil		.527	
Jeg føler meg vellykket hvis jeg gjør det bedre enn de fleste andre studentene			-.871
Det er viktig for meg å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.784
Jeg ønsker å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.865
Jeg føler meg veldig bra hvis jeg er den eneste som kan svare på lærerens spørsmål i matematikktimen			-.665
Eigen value	4,89	3,29	1,88
% total variance	30.89	18.37	11.57
α.	.81	.74	.81

Table 6. Summary results of PCA with oblique rotation – T3.

Variables	Rotated factor loadings		
	PAV	MAP	PAP
Jeg jobber med dette faget for å unngå at de andre i klassen skal tro jeg er mindre smart	.821		
En viktig grunn til at jeg jobber med matematikkfaget er at jeg ikke vil dumme meg ut	.776		
Et av mine viktigste mål i dette faget er å unngå å se ut som jeg har vanskeligheter med å løse matematikkoppgavene.	.724		
Grunnen til at jeg jobber med matematikkfaget er for at læreren min ikke skal tro at jeg kan mindre enn de andre	.731		
En grunn til at jeg ikke deltar aktivt i matematikktimene er at jeg vil unngå å fremstå som mindre smart	.729		
En viktig grunn til at jeg jobber med matematikkfaget er fordi jeg synes det er artig		.859	
Jeg jobber med dette faget fordi jeg er interessert i matematikk		.826	
En viktig grunn til at jeg jobber med matematikkfaget er at jeg liker å lære noe nytt		.634	
Jeg liker å jobbe med matematikkoppgaver jeg lærer av, selv om jeg gjør mange feil		.602	
Jeg føler meg vellykket hvis jeg gjør det bedre enn de fleste andre studentene			-.877
Det er viktig for meg å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.805
Jeg ønsker å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.863
Jeg føler meg veldig bra hvis jeg er den eneste som kan svare på lærerens spørsmål i matematikktimen			-.678
Eigen value	4,89	3,29	1,88
% total variance	33.66	19.06	10.06
α.	.83	.75	.83

At the bottom of table 4, 5 and 6 the statistical reliability measures of the three factors are presented. The overall alpha value of the four factors is above the criterion of 0.7, indicating satisfactory internal consistency (DeVellis, 2012). In sum, the principal component

analysis resulted in three factor structures. Table 7 presents a brief overview of the means and standard deviations of the three factors across the three time periods.

Table 7. Study 1: Descriptive statistics.

	T1		T2		T3	
Factor structures	*M*	*SD*	*M*	*SD*	*M*	*SD*
MAP	3.31	.51	3.28	.50	3.22	.52
PAP	2.78	.72	2.63	.75	2.56	.79
PAV	1.66	.61	1.68	.61	1.70	.61

Note. T = time; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.

Thus, based on the following: sample size, convergence between two extraction criteria (Kaiser's and scree plot) for selection of three factors and an average communality level of> 0.6, three factors were selected for use in the further analysis.

6.5.4 Results spring semester

Through all three time periods during the spring semester three factors had an eigenvalue above Kaiser's criterion of 1. In addition to the eigenvalue criteria, a scree-plot was used to visualise the number of possible factors. The curve had a characteristic shape, with three factors before the curve broke and leveled off to the right.

To obtain a clearer picture of which variables were related to the various factors, the extraction was followed by an oblique factor rotation. Three initial analyses were performed to reveal the eigenvalues of potential factors. The analysis resulted in five items being removed from the original scale: three items assessing mastery-approach goals, one item assessing performance-approach goals and one item assessing performance-avoidance goals. These were items that had high factor loadings on more than one factor, and their removal reduced the number of items measuring the students' achievement goals from 17 to 12.

Through all three time periods during the spring semester three factors had an eigenvalue above Kaiser's criterion of 1 and explained a total of 67% (T1); 75% (T2) and 72% (T3) of the total variation in the variables. Table 8, 9 and 10 presents the factor loadings after the oblique rotation for T1, T2 and T3. Based on the cluster of variables that loads on the various factors, the first factor is assumed to represent *performance-avoidance goals*. Factor number two addresses *mastery-approach goals*, while the third factor refers to *performance-approach goals.* Furthermore, all factor loadings in all three analyses are above 0.4, and the

overall alpha values of all four factors are far above 0.7. Table 11 provides an overview of the mean and standard deviation for the three factors across the three time periods.

Table 8. Summary results of PCA with oblique rotation – T1.

Variables	Rotated factor loadings		
	PAV	MAP	PAP
Et av mine viktigste mål i dette faget er å unngå å se ut som jeg har vanskeligheter med å løse matematikkoppgavene.	.74		
Grunnen til at jeg jobber med matematikkfaget er for at læreren min ikke skal tro at jeg kan mindre enn de andre	.82		
En grunn til at jeg ikke deltar aktivt i matematikktimene er at jeg vil unngå å fremstå som mindre smart	.74		
Jeg jobber med dette faget for å unngå at de andre i klassen skal tro jeg er mindre smart	.78		
En viktig grunn til at jeg jobber med matematikkfaget er at jeg ikke vil dumme meg ut	.64		
Jeg jobber med dette faget er fordi jeg er interessert i matematikk		.90	
En viktig grunn til at jeg jobber med matematikkfaget er fordi jeg synes det er artig		.86	
En viktig grunn til at jeg jobber med matematikkfaget er at jeg liker å lære noe nytt		.73	
Jeg ønsker å gjøre det bedre i matematikk enn de andre studentene i klassen min			.90
Jeg føler meg vellykket hvis jeg gjør det bedre enn de fleste andre studentene			.89
Det er viktig for meg å gjøre det bedre i matematikk enn de andre studentene i klassen min			.79
Jeg føler meg veldig bra hvis jeg er den eneste som kan svare på lærerens spørsmål i matematikktimen			.74
Eigen value	1,62	4,09	4,23
% total variance	12.51	19.40	34.99
α.	.82	.79	.86

Table 9. Summary results of PCA with oblique rotation – T2.

Variables	Rotated factor loadings		
	PAV	MAP	PAP
Et av mine viktigste mål i dette faget er å unngå å se ut som jeg har vanskeligheter med å løse matematikkoppgavene.	.92		
Grunnen til at jeg jobber med matematikkfaget er for at læreren min ikke skal tro at jeg kan mindre enn de andre	.84		
En grunn til at jeg ikke deltar aktivt i matematikktimene er at jeg vil unngå å fremstå som mindre smart	.77		
Jeg jobber med dette faget for å unngå at de andre i klassen skal tro jeg er mindre smart	.77		
En viktig grunn til at jeg jobber med matematikkfaget er at jeg ikke vil dumme meg ut	.65		
Jeg jobber med dette faget er fordi jeg er interessert i matematikk		.94	
En viktig grunn til at jeg jobber med matematikkfaget er fordi jeg synes det er artig		.93	
En viktig grunn til at jeg jobber med matematikkfaget er at jeg liker å lære noe nytt		.75	
Jeg ønsker å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.93
Jeg føler meg vellykket hvis jeg gjør det bedre enn de fleste andre studentene			-.92
Det er viktig for meg å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.77
Jeg føler meg veldig bra hvis jeg er den eneste som kan svare på lærerens spørsmål i matematikktimen			-.66
Eigen value	5,62	3,62	1,59
% total variance	41.59	20.53	12.84
α.	.87	.87	.88

Table 10. Summary results of PCA with oblique rotation – T3.

Variables	Rotated factor loadings		
	PAV	MAP	PAP
En viktig grunn til at jeg jobber med matematikkfaget er at jeg ikke vil dumme meg ut	.83		
Jeg jobber med dette faget for å unngå at de andre i klassen skal tro jeg er mindre smart	.81		
Et av mine viktigste mål i dette faget er å unngå å se ut som jeg har vanskeligheter med å løse matematikkoppgavene.	.80		
Grunnen til at jeg jobber med matematikkfaget er for at læreren min ikke skal tro at jeg kan mindre enn de andre	.79		
En grunn til at jeg ikke deltar aktivt i matematikktimene er at jeg vil unngå å fremstå som mindre smart	.78		
En viktig grunn til at jeg jobber med matematikkfaget er at jeg synes det er artig		.93	
Jeg jobber med dette faget er fordi jeg er interessert i matematikk		.92	
En viktig grunn til at jeg jobber med matematikkfaget er at jeg liker å lære noe nytt		.79	
Jeg føler meg vellykket hvis jeg gjør det bedre enn de fleste andre studentene			-.86
Jeg ønsker å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.85
Jeg føler meg veldig bra hvis jeg er den eneste som kan svare på lærerens spørsmål i matematikktimen			-.77
Det er viktig for meg å gjøre det bedre i matematikk enn de andre studentene i klassen min			-.76
Eigen value	4,98	3,63	1,89
% total variance	39.51	18.77	13.72
α.	.87	.86	.85

Table 11. Study 2: Descriptive statistics.

	T1		T2		T3	
Factor structures	*M*	*SD*	*M*	*SD*	*M*	*SD*
MAP	3.09	.66	3.09	.72	3.17	.68
PAP	2.55	.82	2.59	.82	2.56	.81
PAV	1.66	.60	1.79	.68	1.73	.65

Note. T = time; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.

In sum, based on a total of three factor analyses, factorial invariance was confirmed. Stability was documented, Cronbach alpha values were high and changes in scores over time can therefore be interpreted as true change in the students' achievement-goal endorsement across the three time periods.

6.6 Quality Research and Critical Reflections

Since the quality criteria for quantitative research were reviewed above, this section will deal with the critical considerations relating to the quality in the qualitative research work. Qualitative research must be subordinate to theoretical principles and guidelines, just as with other research. According to Ringdal (2018), one of the most important research-ethic rules for data collection is voluntary participation expressed in the concept of *informed consent*. This means that the informants are the ones to decide over their own participation. The informants must be informed that their participation is voluntary, and they must be given

sufficient information about the project so they can make an informed decision about participation. When it comes to this study, the teachers and students were all informed about its background and intentions, they were told their participation was voluntary and that they had the right to withdraw at any time. They were also informed that all interviews and conversations would be recorded on tape and that everything would be made *anonymous* by replacing their names with pseudonyms during the transcription phase.

According to Johannessen et al., (2016), anonymity may be difficult to ensure in qualitative studies compared to quantitative studies as in qualitative research it is common to give detailed descriptions of individuals, pointing out that changing the names of the informants is a possible way of guaranteeing anonymity. The teachers and the students were also informed that the data material would only be used for the study in question, and that the transcriptions would be stored safely and destroyed after the study was completed. Bearing all this in mind, it is safe to say that the informants involved in this study were not *deceived*, which according to Ringdal (2018) is another research-ethic principle that researchers must comply with.

Reliability. An important topic for most research is quality assurance. Familiar quality criteria are reliability, validity and generalisability. *Reliability* refers to the dependability of the data, often critically important in quantitative studies, where there are different ways of testing data reliability, such as test-retest reliability and inter-rater reliability. According to Johannesen et al., (2016), this type of requirement is not very useful in qualitative research. First, non-structured data collection techniques are normally used, where the conversation controls the data collection. It is therefore very difficult for researchers to duplicate each other's work.

Second, researchers use themselves as an instrument (Postholm, 2005), nobody else has the same experiential background, and they are therefore unable to fully understand the interpretation process of others. When it comes to structure, this study uses a number of interview forms, through semi-structured teacher interviews to focus-group interviews with students, where the participants could make many contributions to form and content, in contrast to a more structured form of interview. This may have led to a lower level of reliability. In more structured interview forms, the interview is controlled by the interviewer and his/her interview guide, which may give the interviewer more control. In this study an interview guide was used in all the interviews, but with a high level of flexibility.

In the teacher interviews the guide was used as a tool for reflection, and with the students, some questions could, for example, be skipped over as they had already touched on

the topic, where threads they were discussing which had not been included in the guide could be pursued further. But this does not mean that for this reason the interviews were conducted in completely different ways. Both the teachers and all the student groups, distributed over two studies, received the same questions, even if the questions were not all asked in the same order because the participants touched on different topics at differing times. All the topics were covered in all the interviews. What according to Johannesen et al., (2016) may strengthen reliability, and which was done in this study, is to form a detailed presentation of the procedure used throughout the entire process.

This raises an important challenge with the use of grounded theory, in that researchers can, like with other qualitative fields of analysis, influence the result of the analysis, especially in relation to the categories to be developed. For this reason, it is often recommended that several researchers should code the data material together (Charmaz, 2001). This was not done in this analysis. Grounded theory also highlights the importance of having an open mind without preconceived notions and opinions, which of course is difficult to guarantee. This does not mean that grounded theory rejects all use of theory. The point is to not have a theory as a starting point for interpretation but start the analysis with as open a mind as possible (Johannesen et al., 2016). At the same time, researchers cannot work completely unconditionally. Researchers must have a certain focus before the study. Already with the choice of research questions, certain frameworks are set around the analysis, in that it is decided which reality is to be investigated and which aspects of this are to be focused on. Data is theory-loaded, and researchers can never completely free themselves from their frame of reference and their preconceptions (ibid).

Working in a collaborative team within an interactive action research framework, and as mentioned above, "author" had the responsibility for the empirical and theoretical framework, which may, of course, have colored the analysis. In addition, thorough reading about professional development in relation to lessons from lower educational levels may have influenced what "author" drew from the conversations with the teachers and how the content from these conversations was interpreted, in relation to, for example, how they experienced their own learning process.

This was also one of the reasons why this particular method of analysis was chosen, in other words the analytic steps, and the logic of elevating the codes that are most often repeated, led to fewer coincidences and more structure throughout the analysis. This was also one of the main points of Glaser and Strauss when they developed the method; the analysis does not take place exclusively intuitively and unsystematically (Glaser and Strauss, 1967). It

introduces a significant element of systematics that can help structure the analytic process. The aim of this analysis was to get deeper into the conversations and discussions and the interlocutors experiences. All interviews were analysed in the same way using the same analytic steps. In other words, coincidence did not prevail, a clear structure characterised the entire analysis.

In addition, grounded theory brings a closeness to the data material. A closeness that helps to reveal feelings, thoughts, intentions and experiences that can be difficult to obtain with other methods (Charmaz, 2001). In relation to this analysis, the data material had to be reviewed, several times, and every single sentence had to be fine-tuned. The codes were created during reading, and the further analysis was based on the initial coding. The analysis never left the data material which means that time and time again "author" had to dig into the data material to investigate and compare. Thus, "author" never "left" the data material to work further with the analysis on her own. The data material was actively used to confirm or deny "authors" thoughts as they developed along the way. For example, when some of the focused codes had to be elevated to categories, and questions such as what they contained and how they were related to other categories had to be answered, there was only one place where these answers could be found, and that was in the data material. In other words, the data material was at the center throughout the analysis.

A key concept in grounded theory is *theoretical sensitivity*. That is, precisely, the researcher's ability to identify what is important in the data material and give it meaning, an ability that can contribute to streamlining the coding process (Strauss and Corbin 1998). Sources of such sensitivity are professional and personal experience, and literature available in the field (ibid). Researchers in grounded theory are thus not concerned that they have to conduct research without any form of ballast. It is almost assumed that one is familiar with the literature. The essential point is that the categories from this literature do not control the researcher's data collection and subsequent analysis (Johannesen et al. 2016). Interpretations must include the perspectives and voices of the people being studied. At the same time, researchers must take responsibility for their interpretive roles. It is not enough just to report and highlight the views of those who are studied, what is observed, heard or read must ultimately be interpreted (Strauss and Corbin 1998).

Validity. Validity in qualitative research relates to the interpretation of the data (Thagaard, 2016), and can be assessed according to whether the findings from the investigation represent the reality that has been studied (Silverman, 2011).

A relevant point of departure for discussing validity in qualitative research is to use the concept of *transparency* (Silverman, 2011; Thagaard, 2016). Transparency refers to whether researchers describe clearly and openly what they do in their analytical processes (Brandi and Sprogøe, 2019). This can be accomplished in a number of ways. Researchers may explain in detail every stage of the analysis through a systematic description of their own analysis process, or they can prepare supplementary memos and detailed code descriptions that enable outsiders to follow their thinking and practical and theoretical considerations in the analysis (ibid). This study aimed for transparency by providing a thorough description of the analysis process (see section 6.3.6).

Another way for researchers to strengthen the validity of qualitative research is to critically review the analytical process, whereafter a colleague or co-researcher can then critically examine the researcher's analyses (Thagaard, 2016). As mentioned above, this was not done in this study, which might have undermined the validity to some degree. On the other hand, validity in qualitative data can also refer to the process of generating data. One way for researchers to increase validity is thus to explain the approaches used in the project and relationships in the field (ibid). For this study it is therefore relevant and important to emphasise the data collection method, semi-structured interviews, or the reflection conversations, which functioned as an open conversation where the teachers' reflections and experiences were the main focus. Each conversation started openly with their reflections on the previous assessment, where "author" did not lead them in any particular direction with more leading questions, nor did "author" start the interview sharing her observations, which could also have led the conversation in particular directions. This was a deliberate choice made by "author" to ensure that the data collection was authentic and of high quality, thus reducing the degree of researcher bias, which is a practical approach to validity in qualitative research (Norris, 1997).

Another important validity check in qualitative research is to invite the participants to assess the findings. In this approach, called member-checking, the findings are presented to the informants to ascertain whether they support them and recognise what they have in fact said (Brandi and Sprogøe, 2019). Whether this actually strengthens the validity and affects the level of researcher bias depends, according to Thagaard (2016), on the researcher's context for the interpretation. "If the aim is to present the understanding people in the field have of their situation, their perception may form the basis for assessing the validity of the interpretation. If the aim is to place the understanding of the participants in a wider academic context, their assessment will not be able to confirm this" (p. 207). In this study the teachers

were given the opportunity to read the findings from the study they were involved in (Sub-study 1 and Sub-study 2), where the aim was to present the findings of the analysis, not to address the wider academic context. The teachers found that they could recognise what they said in the descriptions, and that these were in accordance with their experiences. Given that the aim of the interviews was to highlight their lifeworld, their recognition and further approval of the findings can strengthen the belief that they represent the reality that was studied (cf. Silverman, 2011).

However, the students were not given the same opportunity to carry out this type of member-checking, which in itself may have undermined to some degree the validity of the student findings from Sub-study 2 and Sub-study 3.

Another shortcoming that is important to mention in reference to the intervention study and the students is that the student groups were not consistent throughout the research period. A key difference between the student samples in Sub-study 2 and Sub-study 3, is that the former had students from two classes (as only two teachers were in the study), from the 2013 student group, while Sub-study 3 (where five mathematics teachers were in the study) had students from five introductory classes in the 2014 student group. Even if the student groups were not consistent, which could have increased the validity, it is important to point out that in no way were they completely different. The students from both years were around the same age, with a large spread of backgrounds – as is often the case with preparatory courses – and they were all aiming to be admitted to an engineering study programme. Moreover, the teachers in both studies used the same methodology in their assessment practices in mathematics, and the same interview guide was used for both years.

What is interesting to point out in terms of validity, on the other hand, and what may be claimed to be a strength in the student findings, is that in spite of different years and classes, they expressed highly similar perceptions of the assessment practice in mathematics. The student findings from Sub-study 3 therefore confirm the student experiences from Sub-study 2 with common experiences relating to a strong learning focus, the value of a learning-oriented introduction and a dedicated teacher role, and learning outcomes relating to such activities as discussion and reflection.

Generalisability. Another way of quality assuring research is to assess the level of *generalisability*. This means that the research findings can be transferred to similar phenomena. In quantitative research this is a widely used quality criterion, usually called statistical generalisation. In this study statistical generalisation was not a topic. A more important discussion is the broader distribution of the findings.

7. FINDINGS

The following section presents the main findings from substudies 1, 2 and 3 in more detail. The aim of this section is thus to summarise the work and research findings that is presented in the included articles. The first article introduces and discusses the teachers understanding and experience of change and examines how teachers can explore and utilise important empirical lessons from lower educational levels to facilitate and create changes in their own practice. The second article evaluate and discuss the value of dialogue as a means of facilitating compliance between teacher and student understandings. Finally, the last article focuses on students' achievement goals within a continued formative assessment practice.

7.1 Sub-study 1 (Article 1)

The aim of Sub-study 1 was to examine how teachers experience their own change process towards creating a formative assessment practice, their goals and learning experiences. The research question was: *What are the teachers' foremost experiences of their own change process?*

Throughout the analysis one main code, "*a learning teacher*", with two associated sub-codes, "*a new role*" and "*instruments of consciousness*", stood out, and all three were elevated to a category and two sub-categories as they appeared to be the teachers' most important experiences in relation to their process towards creating a formative assessment practice in mathematics.

"A learning teacher" refers to the teachers' learning process. According to them, all development and change in practice was directly related to their personal learning process and change as teachers. Creating a formative assessment practice thus commences with a change process in their role as teacher.

Through specific interventions, such as the implementation of learning-oriented introductions, self-assessment through reflection (rather than teacher-led grading), group discussion, application of response technology and plenary-session dialogues, the teachers created a learning-oriented assessment practice that allowed for a considerably higher level of student activity. In practice this meant relinquishing control, discarding the familiar and safe presenter role and allowing the students to participate actively with them. In other words, the new assessment context formed the basis for "a new role" (sub-category 1) for the teachers to grow into, i.e. the role of co-actor, or a type of collaboration partner.

But to function as a partner, the teachers quickly learned that they also needed to function as credible role models. They needed to persuade, initiate and not least implement.

To elaborate, a critically important part of their learning process was the students, and the teachers had to persuade them that the choices and actions taken in the classroom were well thought out and worthwhile. According to the teachers, this required dedication and specific actions in the classroom. For example, they learned that the quality of the dialogue they managed to establish with their students, and the students' efforts, were directly related to their best efforts to give good, clear and convincing introductions, presenting important learning tools in the form of reflection questions and discussion questions, and finally, their best efforts to create a good joint dialogue in the interactive reviews. Through such experiences, the teachers became aware of their own ability to influence the situation, which they highlighted as one of the most important learning outcomes in their own change process. The sense of having a real influence also functioned as an important catalyst for their further efforts and the development of confidence in themselves, their new role and assessment context.

The core of the teachers' change process was comprised by the work methods in their team. As members of a team, the teachers made a conscious effort to implement specific principles, collaborate with each other, be observed on a regular basis and reflect actively, both individually and jointly. For the teachers, these methods served as *instruments for increased consciousness* (sub-category 2).

The opportunity for the teachers to work collaboratively in an atmosphere of trust and mutual respect was a key part of the social community within the team. First, they experienced a strong relationship to the researcher in the team, "author", a relationship which developed through repeated observations and reflection dialogues. For the teachers, it was important to receive external input, to have someone who could help them assess their own choices and actions against an empirical framework. However, "author" had her limitations in her support function, which made the other part of the community, the teamworking teachers, even more important. For the teachers it was important to not stand alone, but to have each other as support. They represented a different form of credibility than "author" did. While "author" represented empirical credibility, the teachers represented credibility according to the experiences they developed through the interventions in their practice, which gave them a practical ballast that "author" could not provide in her researcher role. Moreover, they found it very useful that they often had different ideas and that the way they implemented and perceived their own assessments differed between them; this gave them inspiration, tips and tools they could use in their own classrooms.

Another important tool for the teachers in their change process was reflection. In short, reflection as a working method had an impact on the teachers' thought processes in a completely new way. Knowing in advance that they would need to reflect on their own performance, and then participate in a joint refection conversation and discussion on how the different techniques they applied in the classroom functioned, made them more alert and aware of their own choices and actions. Reflecting on how to assess their own role as teachers, how the students responded and the extent to which the choices they made contributed to creating the arena for learning that they envisioned also led the teachers to think differently and more actively when they were in the classroom with their students.

The teachers applied the term "awareness-raising" to the work process they experienced. Using such tools as reflection, observation and collaboration, they found that they developed greater understanding of what took place in their own classroom, and they learned the value of their own efforts. According to the teachers, this was where the real knowledge development took place: the increased awareness about their own practice and role as teachers.

7.2 Sub-study 2 (Article 2)

The objective of Sub-study 2 was to examine the degree of compliance between the teachers' *intentions* in using the "seven principles of good feedback practice" and the students' *experiences* of them in practice. More precisely, this article asked the question: *to what extent are the teachers' beliefs about what they are doing in a formative assessment concordant with how the students experience it?*

All in all, the findings from this study differ from previous research findings on the perception of assessment and feedback; while the latter broadly acknowledge a significant mismatch between student and teacher viewpoints, the findings from the study included in this article reveal a clear majority of common features between the students' and teachers' perceptions. The study concludes that compliance is the result of a dialogue between teacher and student, and active efforts and participation by both parties. In this study the correspondence and dialogues were initiated by the teachers and maintained through continuous and deliberate effort, which in turn convinced and engaged the students in such formative activities as self-assessment, reflection as feedback and dialogue.

The data relating to the teacher viewpoint revealed one main category "*a learning arena*", with three associated sub-categories, "*introduction*", "*reflection*" and "*dialogue*", whereas the data relating to the student viewpoint revealed the categories "*reflection*" and

"*social relations*". The latter has two associated sub-categories, "student relations: the value of group discussion" and "student-teacher relations: the value of the teacher role". The categories represent the teachers' and students' foremost intentions behind and experiences of the formative assessment practice in mathematics.

Let us start with the point of departure for the parties involved in this study; namely a strong sense of meaninglessness related to previous assessment experiences, a premise that is in line with the generally weak formative assessment culture described in the research literature. Furthermore, both parties point to the focus on learning as the main distinction between past and present experiences.

For the teachers, the focus on creating "an arena for learning" (teacher main category), was the very essence and the foremost intention of their assessment practice in mathematics. The decision to replace normative assessment with self-assessment through reflection (teacher sub-category 2) was a key part of this intention. Through reflection, the teachers wanted the students to increase their awareness of their own learning by putting their own understanding, challenges and further solutions into words. For the students, and in accordance with the teachers' intentions, self-assessment and reflection (student category 1) represented a genuine opportunity to think actively and become more aware of their own learning process. According to the students, the reflection process functioned as important feedback on their own learning.

Beyond encouraging reflection, the teachers also aimed to create a common arena for learning, where students, and the teacher and students, could develop a good dialogue between each other (teacher sub-category 3). Both the teachers and students felt that the group discussions represented a very clear switching of roles in the classroom, where the students assumed the role as the most active party. The teachers introduced several measures to initiate dialogues about the subject and shared learning in the groups, including changing the positions of the desks to physically create groups, presenting specific tools and hints and using response technology to round off and summarise the discussions. In accordance with the teachers' intentions, the students highlight all these new ideas as important for both starting up and rounding off the discussions, and describe peer dialogue (student sub-category 2.1) as a golden opportunity to learn from each other.

The parties were not quite so harmonious when it came to what they thought about the dialogues in the plenary sessions. While the teachers pointed to the reduction in physical distance, regular involvement and attempts to create a cooperative "we" environment as important for initiating dialogues with the students in the plenary sessions (teacher sub-

category 3), the students found it more natural and simpler to explain what they were thinking and experiencing to a co-student (student sub-category 2.1), finding it rather more difficult to express this to a teacher in a plenary session. The reason is basically the level the students are on in the subject, and the language they use amongst themselves. However, they found the repeated invitations by the teachers to join a common plenary dialogue to be a characteristic of an engaged and interested teacher, which they felt increased their own engagement (student sub-category 2.2). They also pointed out that everyone participated in the plenary sessions using response technology.

There was more harmony on the issue of a calm and stress-free assessment context. The teachers entered the classroom with a clear intention, where they used their introductions (teacher sub-category 1) in a deliberate way to explain and exemplify the focus on learning, which they found had a calming effect on the students. The students confirm this, also pointing out the importance of actions, and that the teachers' introductions were not only "fancy words" used to embellish the opening of an assessment session, but that the teachers' focus on learning and their efforts permeated the entire interactive review (student sub-category 2.2). The students thus learned that a test may actually be a golden opportunity for learning, which we can then say satisfies the teachers' intention to create an arena for learning.

The students were, however, far less satisfied with what occurred after the assessments, and criticise their teachers for not using the results of the assessments more actively in their continuing teaching. The teachers attempted to build further on valuable information from the assessments in their continuing teaching, but the way the students saw this, the information was not used adequately.

7.3 Sub-study 3 (Article 3)

Finally, the overriding *objective* of Sub-study 3 was to examine which achievement-goal patterns the students pursued within this context. Furthermore, the aim was to explore these observations through interviews with students and through relevant research literature in the fields of achievement-goal stability and change and formative assessment. The research was guided by the following questions:

- Which achievement goal patterns do students pursue in a formative assessment practice in mathematics?
- How did the students perceive the formative assessment practice in mathematics?

7.3.1 Results from the fall semester

The results from the fall semester deviate from previous studies, especially when it comes to changes earlier studies have demonstrated in mastery-approach and performance-avoidance goals. Previous research has revealed a rather large decline in mastery-approach goals, particularly from T1 to T2, and an increase in performance-avoidance goals. However, this study finds mastery-approach goal stability between T1 and T2, followed by a minor decrease between T1 and T3 and performance-avoidance stability. Performance-approach goals underwent the greatest changes in study 1, with a significant decrease from T1 to T2, followed by a subsequent decrease from T1 to T3. Thus, the students became less performance-oriented throughout the semester.

The individual-level change analysis confirmed and expanded the patterns presented from the mean-level change analysis and revealed an overall stability for all three achievement goals. Only minor increases and decreases were detected. Unlike the results from the mean-level change analysis, which showed a significant decrease in mastery-approach goals between T1 and T3, the individual-level change analysis presented a small increase in mastery-approach goals between T2 and T3. The mean-level change analysis presented no significant change during this period, which may have been caused by a similar number of increases and decreases cancelling each other out and giving the appearance of no group change in the scores. Furthermore, the stability of performance-avoidance goals was confirmed and clarified through the individual-level change analysis. The majority of the students remained stable in their endorsement of performance-avoidance goals during all three time periods. Finally, the results from the individual-level change analysis verified the decline in performance-approach goals, with the largest decline between T1 and T3.

7.3.2 Results from the spring semester

In comparison to the results from the fall semester, the results from the spring semester revealed great stability in the students' achievement-goal endorsement over time, whilst confirming minor changes. The results revealed one significant change at the group level during the spring semester: an increase in performance-avoidance goals from T1 to T2. In other words, performance-avoidance goals increased before students participated in a mock-exam assessment. According to the individual-level change analysis, performance-avoidance goals exhibited an overall high stability level.

Similar to the fall semester, results of the mean-level change analysis revealed stability in mastery-approach goals from T1 to T2 during the spring semester. The individual-level

change analysis detected changes in mastery-approach goals between T1 and T2, however, the changes were similar in the number of increases and decreases and have thus cancelled each other out at the group level. The mean-level change analysis also indicates an increase in mastery-approach goals from T1 and T3, but the increase was not statistically significant. However, the increase was confirmed through the individual-level change analysis, which suggests that the students had become more mastery-oriented by the end of the school year.

Based on the mean-level change analysis, performance-approach goals appeared to be the most stable achievement goals at the group level. They increased between T1 and T2, but the increase was not statistically significant. At the individual level, the biggest changes in performance-approach goals occurred between T1 to T3, with a minor decrease.

7.3.3 Results student focus-group interviews

The experience the students appeared to agree on the most was the experience of the purpose of the assessments, i.e. learning. It had not gone unnoticed by the students that the teachers wanted the assessments to function as an arena for learning. This intention was clearly communicated to them at the beginning of each assessment: The assessments were to be an arena where they could address their own misconceptions, and then through reflection and discussion experience mastery and learning. For the students, the purpose appeared to be very clear, and they perceived the assessment context as meaningful in terms of their own learning. They experienced that the awareness of their own skills increased through self-assessment and reflection and felt that their understanding of mathematics improved through the discussions in small groups. Finally, the students perceived their mathematics teachers as critically important. The teachers presented the purpose of the assessments and facilitated and initiated measures for learning through reflection and discussion.

Bearing the students' experiences in mind, the assessment practice in mathematics may thus be described as learning oriented, not performance oriented.

8. DISCUSSION AND CONCLUSION

The discussion here will focus on how the study presented in this thesis, supported by the three articles included, and by applying the theoretical frameworks presented in Chapter 4, can contribute to a broader understanding of use of feedback in a higher education learning context. This discussion will illuminate and answer the main research question of this thesis: *how can educators develop their own feedback practice in a more formative way*?

8.1 Formative Learning Environment: New Roles - Shared Responsibility

One of the foremost results from this intervention study is that creating a formative feedback practice commences with the teacher *role*. More concrete, teachers can turn a feedback practice in a more formative direction by adopting and performing a more active role as collaboration partner; taking on a partner role with a concrete responsibility for *creating a learning environment* for their students.

The centrality of roles in a formative practice is illustrated with the key realisation within Black and Wiliam's theoretical framework that learning is a process of *shared* responsibility (Black and Wiliam, 2009; 2012). The students are responsible for learning, implying that the teacher cannot create learning – only students can, the teachers, in turn, are responsible for creating situations; for "engineering" dialogical moments in which students can assume their role as active thinkers and exercise their responsibilities as learners. This responsibility to achieve the competence aims is exercised both in the design of the teaching and in the steering of the dialogue through which the underlying competence aims can be achieved (Black and Wiliam, 2012).

8.1.1 Pro-active regulation: external and internal dialogues

An important part of the learning design created in this study, was the introduction and use of self-feedback through reflection and discussion in small groups. Using theory on formative assessment, this can be seen as proactive regulation, activities where a teacher does not participate directly, rather planning and facilitating for formative dialogues (Wiliam and Thompson, 2007).

Even if the teachers did not involve themselves directly in the students' group discussions or reflections, they still intervened in many ways to facilitate for and thus guide the proactive activities. In accordance with the overarching intention of creating an arena for learning, the activities were emphasised and described precisely as a learning arena, where the learning benefits were highlighted in the introductions. For example, the value of discussing a

subject and developing a shared understanding together was presented as the main activity in the group discussions, rather than focusing on arguing for one's personal view, convincing others and, hence having the winning argument. The group discussions were thus intended to function as a learning dialogue (cf. Dysthe's distinction between learning and argumentative dialogue). Furthermore, the teachers wanted all the students to take an active part in the activities. Each reflection and discussion was therefore initiated or literally launched by a teacher, and the students were presented with a set of "tools" in the form of instructions and questions they could include in their discussion or reflection process, in addition to concrete hints about the subject matter.

Regulation as mediation. One way of understanding the teacher role, the efforts behind the student activities, is to think of them as mediation, an important point in socio-cultural theory based on Vygotsky's works. Mediation is about giving support in learning processes, whether given by individuals or through tools. We are participants in a cultural community and learn to use the tools at the group's disposal. Tools are understood as, for example, physical objects, but also the intellectual and practical resources one has access to. In such a perspective, the teachers' frameworks, in the form of the tools they created, may be understood as support or an instrument adapted to the situation.

The most important tool for learning in a socio-cultural view of learning is, however, language (Dysthe, 2001). Language is a key for communication as well as for thought processes. Hence, from a socio-cultural understanding of learning, we can say that learning occurs in linguistic interaction. Language, whether spoken (through, for example peer dialogues) or written (through, for example, individual reflection), thus becomes a mediating factor, a bridge builder between an individual and knowledge. Language is, however, not only a means, but also in itself the fundamental condition for learning; knowledge is established through language (ibid). Again we can look to Vygotsky (1978) and one of his other key concepts, *internalisation*. Thus, through mediation, external interaction (as in this context the interaction between students in groups, or the interaction between the student and the subject matter in a reflection) becomes an internal experience: communication becomes thought (Dysthe, 2001). In this way, the students' spoken discussions (external dialogue) and written reflections (internal dialogue) becomes a path to knowledge.

Active work process. The students involved in this intervention study particularly point to the group discussions and reflections as being important for their learning. Reflection was described as a process where they needed to find and actively use their own words to assess their own solutions, choice of methods, level of understanding and further effort. They

experienced a close link between the written word and learning and felt that they developed greater insight into their own understanding, and greater awareness about further effort. The reflections thus did not become passive repetition of the teacher's review, from a socio-cultural view of learning they may be understood as rather the opposite, active processing in a dialogic development of knowledge (cf. Dysthe, 1996).

In other words, the students became more aware of and active in constructing their own learning process, confirming theory on self-regulated learning, which sees self-feedback and reflection as effective methods students can use to acquaint themselves with self-regulating aspects of their own learning (Yan, 2016; Panadero et al., 2019).

A common language: a common arena. The students also connect the group discussions to their own development of knowledge. Being explained something, and explaining something, and then together endeavouring to increase one's own understanding, is "a win-win situation" as one student put it. The essential ingredient of this win-win situation, however, was that all the participants in the group were just students. They thus had a common denominator because they all knew what it meant to not understand. As students they also experienced that they shared a special mutual language, using words and forms of expression they mutually understood.

To examine this win-win situation in greater depth, hence gain better understanding of the importance of language in these learning dialogues, we can turn to Bakhtin (1981) for more insight, as one of his main ideas is that understanding and meaning are created in the response: According to Bakhtin, understanding and meaning are not located in a text or in a sender in a dialogue. Nor is meaning something created in an individual and his/her consciousness. Understanding requires an answer or a response from the listener.

The key idea here is that the students experienced the group discussion as a natural and common arena for participation. If we see this from Bakhtin's perspective, we can say that the students in the groups listened to their co-students (thus hearing the utterances of other students), they responded, both by sharing their opinions and giving feedback on the original utterance, and together they used their utterances and responses in a mutual development of language. In Bakhtin's terms, the feedback from the other person becomes the activating principle that creates the basis for understanding; "Understanding comes to fruition only in the response. Understanding and response are dialectically merged and mutually condition each other; one is impossible without the other" (ibid:282). In this way Bakhtin connects understanding to response.

8.1.2 Interactive regulation: plenary dialogues and response technology

In addition to group discussion and reflection, the new learning design created through this intervention study, also included use of plenary dialogue and response technology. These approaches had in common that teachers did not know how they would proceed in the review until they had seen the students' responses, which can make them an example of an interactive regulation in which teachers use formative assessment in "real time" to make adjustments to their instruction during the act of instruction (Wiliam, 2011).

Commencing with the plenary dialogues, the intentional context was to increase the activity and participation of the students; get them into a plenary session – obtain a response from them – establishing underlying thoughts about what is under review so that the teacher could assess further approaches from that point. This is in accordance with the basic core of formative assessment; namely a focus on the creation of, and capitalisation on, moments of contingency in a whole-class discussion (Black and Wiliam, 2009).

What is perhaps most interesting about this study`s use of interactive regulations is that the parties' experiences contrast to the experience of the more proactive regulation (reflection and group discussion), where the teachers and the students were relatively harmonious in their perceptions, the interactive regulation (plenary dialogue) represents one of the very few exceptions from the parties' more or less steady alignment. As described in article 2, the teachers, for their part, emphasised the importance of personal effort, less physical distance, continuous involvement in the plenary session and the creation of a sense of "us" as effective tools for generating dialogue with the students in plenary sessions. The students, on the other hand, had a quite different perception. They found it difficult to express themselves verbally to the teacher in plenary sessions, which for many meant that they did not respond verbally. The verbal participation that was the foundation for participation in the group discussion no longer felt that simple and natural.

There are several ways of understanding this. Considering it in the light of a socio-cultural view of learning, the students may have experienced more symmetry between each other in the smaller groups of students through experiencing a common language, which in turn made the dialogue more natural for them to participate in, whereas conditions may have become more asymmetrical in the plenary dialogue with the teacher. But it is important to point out that in the expanded dialogue concept, the importance of the socio-cultural background and context is given prominence, and it is further pointed out that the relationship between the teacher and student is asymmetrical in the sense that the teacher's role is to guide the students to accepted norms and values in the social and cultural context they are in

(Dysthe 1996; 2001). Asymmetry or imbalanced distribution of knowledge is thus an important rationale behind the teaching system and the wider education system as such, and this is not in itself an impediment to dialogue. However, external requirements for subject matter content and understanding, as well as the teacher's subject competence may result in asymmetry in the learning situation, yielding an apparent imbalance in the dialogue (ibid).

Ensuring a safe student participation. Even if the lack of student verbal response in plenary sessions consequently may be an important challenge in a classroom discussion between teachers and students in a purely formative mode, in the sense that the teacher finds little to build further decisions on, this does not necessarily mean that the parties are not in dialogic interaction. In the presentation of the extended dialogue concept in socio-cultural learning, symbol-based interaction among the interlocutors is also included (Dysthe, 2001). This understanding also starts with Bakhtin's (1981) complex dialogue concept and emphasises the importance of each utterance being dialogic, whether its form is spoken, gesticulated or written, or what may be reasonable to add here, *electronic*.

For the students involved in this study, the participation was actually more closely related to the use of response technology rather than verbal communication with the teacher. Using response technology was also the second interactive regulation form facilitated in the new learning design. For the students this was highlighted as a tool which was important for their perception of learning. If the responses revealed that the students had misunderstood something, the teachers would go through the task in greater detail in the plenary session – which they felt reinforced their learning. Thus, confirming the value of external sources of feedback in effective formative feedback practices (Boud, 1999; Winstone and Boud, 2020). In this way the response technology may also reasonably be understood as a mediating tool for a formative practice. The technology contributed to getting the students to respond, and the responses were then used to make decisions about the next actions. With access to all the students' responses, the teachers obtained an overview of their current status, a type of overview that may otherwise be challenging to acquire in a plenary session, and an overview which enabled the decision to be made, for example to review a specified topic more thoroughly, to be made according to a better foundation than a decision made without this information (cf. Black and Wiliam's definition of formative assessment).

For the students it may appear that the technology was about finding a "suitable channel" for responding in the plenary sessions. Several elements may have had impact on this. One of the strengths of using response technology is that students may participate without exposing themselves (Draper and Brown, 2004). This may perhaps be particularly the

case here, where the technology allowed the students to respond anonymously. From a motivational aspect this may be worth bearing in mind. As we will discuss further in the second part of this discussion, students may have diverging views of intelligence, and hence very different views of how learning takes place. Students who ascribe to an entity view believe that assessment is an all-encompassing activity that defines them (Dweck, 1999). If they fail the task, they are failures (Forsythe and Johnson, 2017). This places a special responsibility on teachers. Verbal participation in plenary sessions where you raise your hand so everybody else can see you, including and not least the teacher, where you especially reveal your own understanding, which may be on one or the other side of the scale, including complete lack of knowledge, may potentially be humiliating. Then, withdrawing from participation may feel safer for one's self-esteem. Stiggins (2010) is very clear on this and promotes the idea that in an assessment-for-learning environment, the path to productive student decision-making passes through their emotional reactions to assessment and what those emotions lead learners to do in response. He concludes with the following strong statement: "teachers need to understand this" (p. 241).

The teachers' reviews, seen as a whole, obviously focused on processing the students' achievements, which could have produced different emotional reactions. Even though the students, soon to be examined in more detail, point to the assessments in mathematics in total as an opportunity to learn, we can in no way ignore the fact that the same assessments have not included emotional experiences that have influenced the students. Research is clear that students' motivation and self-esteem are likely to be enhanced when the focus on social comparison is toned down, or perhaps, even better, completely eliminated (Dweck and Master, 2012; Harlen, 2012; Bryan and Clegg, 2019). In the oral plenary dialogues as they were arranged in this context, it would have been impossible for the students to remain anonymous, meaning that social comparison can be very much present. This may have been one of the reasons why the use of response technology was both preferred by the students and described as simpler. It did not lead to any emotional burdens.

8.2 Creating Student Engagement and Motivation

Another way of interpreting this study`s creation of a new feedback practice, and the teachers' efforts, through both proactive and interactive regulation, is to see them as targeted interventions that enable students to believe in their own intelligence in mathematics not as a fixed characteristic, but rather as a malleable quality that they can develop through effort and learning.

8.2.1 Change the meaning of failure: "You can learn even if you have misunderstood".

As mentioned above, teachers can turn a feedback practice in a more formative direction by assuming and practising an active role as collaboration partners where they adopt a partner role with specific responsibility for creating a learning environment for their students. However, this is not a simple and concrete role as such. On the contrary, an essential part of this role is its *broadness*, i.e. everything the role has that will enable teachers to really function as collaboration partners. The importance of *collaborating* with the students is absolutely essential in the development of more formative feedback practices, as the activities the teachers plan for, initiate and support are completely dependent on the students' active participation. This means that the main challenge for teachers when shaping their learning environments is to consider the students' perceptions (Fryer and Elliot, 2012).

The results of this study show that having a well thought out, comprehensive and *genuine* focus on learning in the development of more formative assessments, is critically important, as a genuine focus on learning benefits the students, convinces them, engages them, and in the light of theories on student intelligence, it motivates them.

The genuine message. In order to promote their main intention, that the assessments should function as an arena for learning, and thus initiate ongoing collaboration with their students, the teachers involved in this study spent a lot of time and effort on developing constructive introductions, trying through them to convince the students that a focus on learning in an assessment context is both constructive and useful. The teachers tried to point out that mistakes, and thus failures, are a natural and necessary part of learning, as opposed to leading to an indictment of their worth as individuals, as students with an entity view, who foster the idea that achievement should flow naturally from ability with little effort or use of learning activities, might feel (Dweck and Master, 2012). The formative assessments were there so that the students could get things wrong and learn from their mistakes.

From the students' point of view, the introductions, with their message about learning, had a calming effect on them, which allowed them to relax and try to focus on their own skills and understanding. Importantly, though, it was not the message from the introductions alone that convinced the students to participate. Here, the importance of the teacher role as a presenter (of a message about learning), enthusiast, persuader and initiator came into play. The fact that the teachers took the time to talk with the students at the start of each assessment and highlight the focus on learning, and also gave them tools to work with in the learning process, was, according to the students, a sign that the teachers had a genuine intention behind the assessments. However, the students were not truly *convinced* until they experienced that

the teachers' initial focus on learning permeated the entire interactive review, with such activities as individual reflection and discussions in small groups, which they, as we saw in the first part of this discussion, experienced as important and meaningful activities for learning.

Motivational patterns in a learning-oriented practice. According to Dweck (Dweck and Legget, 1988; Dweck, 1999), theories on student intelligence explore the type of achievement goals that students adopt in achievement situations. To summarise, the more students have an incremental theory, the more they want to learn; the more they have an entity theory, the more they are instead concerned with how intelligent they appear (Robin and Pals, 2002; Blackwell et al., 2007; Dweck and Master, 2012).

In Sub-study 3, presented in Article 3, students' achievement-goal patterns within this learning-oriented assessment context were observed and mapped out over an entire school year. However, as this study is descriptive, it cannot be claimed that the use of a learning-oriented assessment practice served as a causal reason for the disclosure of a somewhat different achievement-goal pattern than what has been found in previous studies (Senko and Harackiewicz, 2005; Fryer and Elliot, 2007; Muis and Edwards, 2009). Nevertheless, in light of important theoretical and empirical implications of theories on student intelligence, there are several motives for promoting a learning-oriented assessment practice as a *possible* influencing factor.

Research shows that much of the feedback students receive in higher education is performance oriented, where the information is more about success or failure, or about how students compare with their peers, instead of focusing on learning goals and making an effort to improve and learn (Dweck and Master, 2012; Winstone and Boud, 2020). Research also confirms that students tend to focus more on performance goals as they get older, at the expense of mastery goals (Anderman et al., 2002). Social comparison can lead many students to doubt their capabilities. More specifically, these changes in beliefs about competence can lead to a *decrease* in student motivation. In other words, an overemphasis on performance goals can undermine learning goals (Dweck and Molden, 2005).

In stark contrast to performance-oriented feedback practices, this study focused on specific learning goals with the aim of improving and learning. This means that the traditional focus on achievement, which allows for social comparison, was toned down considerably. There was no ranking, no grades and no passed/failed designation based on the percentage of correct answers. Rather the opposite was true, the students were responsible for assessing themselves through reflection. Feedback, which the students had previously associated with a

grade, a number of points or a written comment by the teacher, now was perceived, in the words of the students, as something they *gave themselves* through reflection. Hence confirming the current empirical understanding of the active role students must play in feedback processes if they are to enhance learning (Panadero et al., 2018; Dawson et al., 2019).

This is not to say that the students' achievements were not important in the new learning-oriented design. On the contrary, they were essential, in that they were the core and the starting point for the students' efforts in the interactive reviews. Unlike more traditional assessments, however, they were not the "end product". They were part of an assessment process where students could use their own achievements to develop further understanding through the use of different interactive dialogues. Thus, it could be that as the teachers presented, persuaded, initiated and supported a somewhat different assessment environment, they also changed *the meaning* of a test in mathematics. They changed the meaning of performing in this context from being a measure of achieved learning, to becoming a further learning opportunity. As one of the students put it, "*It's not like it's too late to learn. And this is an important message, you can learn even if you have misunderstood.*"

Thus, in light of theory on students' intelligence, it could be said that the students learned a malleable view of intelligence, and that within this social context, with the focuses and tools that were present, they experienced the belief that they could learn through effort and participating in such activities as reflection and learning dialogues. This further reflects a critically important point in theory on student intelligence: having learning skills is by no means the same as using them (Dweck and Master, 2012). Students with an entity view may have learning skills in their repertoire but may not use them because they feel that smart people do not need them or because they lose heart in the face of difficulty and do not think learning strategies will help them. This is why it is so important to ensure that students do not only have the learning strategies they need, but also the motivation to apply them (ibid). This intervention study argues that a genuine learning focus, through continuous efforts by teachers, can lay the groundwork for such motivation. This is a key part of developing a feedback practice in a more formative direction, as it determines whether students are motivated to participate in various learning activities and whether they are open to learning from to the feedback that is created.

This can also help explain the stability in performance-avoidance goals, which remained low and stable throughout the fall semester, and not least, the continuing drop in performance-approach goals. By making reflection, effort and dialogue the main activities of

the assessments, an arena for learning was created which gave few reasons for an increase in the performance goals, and this might explain why none of the performance goals increased during the fall semester, despite previous research demonstrating the opposite (Senko and Harackiewicz; 2005; Fryer and Elliot, 2007; Muis and Edwards, 2009). To that end, it is also interesting that performance-avoidance did increase, once, in the spring, just before a mock exam, when an assessment context that resembled the context of the final summative exam and, furthermore, was also more like the traditional assessment structure used in previous studies on achievement goal stability and change, was used.

8.3 Concluding Remarks

Teachers are learners. The Norwegian higher education system, and in a broader context, European higher education, is still to a large degree dominated by teacher-controlled instruction and feedback practices (Damsgaard, 2019), and still bears the signs of being a privatised system where it is perhaps more natural for scientific personnel to see themselves as professionals in the field rather than as professional educators and teachers, with the obligations, need to reflect and willingness to change that are required in teaching and assessment practice (ibid).

One of the most important lessons from the research literature on school practices, regarding the development of a feedback practice in a more formative way, as considered from the assessments and experiences of this intervention study, is that professionalising the teacher role in a formative context is a necessary step for achieving lasting changes. Furthermore, implied in this lesson is the critically important understanding that teachers are *learners* (Thompson and Wiliam, 2008), meaning that they must be treated as learners, and must see themselves as learners.

For the teachers involved in Sub-study 1, being a learner was the very core of their change process; an all-encompassing experience that defined their process towards creating a formative feedback practice. However, the possibility and ability to be a learner did not occur in a vacuum. If we go back to the description of formative assessment, as "moments of contingency", as described by Black and Wiliam (2009), and discussed previously in the this chapter, a teacher's foremost task in a formative mode is to create situations in which students learn, and thus engineer learning environments. Equally important, according to the critical understanding that teachers are learners, as described by professional development programmes, such as the KLT, higher educational institutions' foremost task with respect to formative assessments is to provide suitable learning environments that create situations in

which teachers can learn about formative assessment. Based on the findings from this intervention study, especially those from Sub-study 1, this means recognising the implications the higher education initiatives have for teachers' experiences, and appreciating the support they need to nurture reflection and personal development.

In accordance with interactive action research as a method and empirical lessons from school research, and in line with several of the recommendations from the KLT programme's theory of action, a support context, or a learning environment, was developed around the teachers in this study where opportunities for practice in real settings were *continuously* followed by reflection. This work process, led by reflection, observation and collaboration, was a *circular process* that became a permanent, established and regular work method for the teachers over a total of three semesters.

Although this study had time on its side as such, in that the teachers participated continuously over several semesters, which is essential for the development of sustained changes (Schneider and Randel, 2010; Leahy and Wiliam, 2012), it was still not a simple process. Going through a critical reflection process surrounding one's own practice required honesty and commitment towards oneself as a teacher and towards the collaborative team of which each teacher was a part. This means that in addition to the fact that formative assessment is also an institutional responsibility, teachers have a concrete responsibility to participate actively in their own knowledge development. In line with the responsibilities required of students in a formative assessment.

In this study active and full participation in the team became a natural and implied requirement, in that there was, of course, no reward or punishment, but rather a joint goal of developing, changing and thus creating a formative feedback practice for the teachers' assessments in mathematics, and a common understanding that in order to achieve this, all the members' experiences and competence were needed. Thus, from the very beginning we decided to be equal partners with various responsibilities, which further created various forms of credibility. The fact that everyone had a responsibility, and that each experience was essential to create change, reflect the objectives in KLT's theory of action for professional learning. First, it may have helped to create what Thompson and Wiliam (2008) refer to as a *climate with clear expectations*, a climate for sharing one's own experiences and critical reflections. This is important if reflection cycles are to have an effect, both in relation to acting as a spur to practice, but, more importantly, to acting as a tool for critical reflection on practice. Trying out and implementing techniques that make sense to them, and where they have an intention and are the person in charge in the classroom, is an essential part of learning

and developing. This is a responsibility each teacher must accept if an intervention is to lead to change (Thompson and Wiliam, 2008). This also means that if we, who support teachers in a change process, fail to treat them as responsible professionals, learning will be narrow and short-term, and thus there will be no development in their competence.

Academic achievements and the degree of frequency: how much is enough? For several of the studies linked to the understanding of formative assessment as a regulation of learning, frequency is a key concept in relation to achieving positive learning outcome (Black et al., 2003; Leahy and Wiliam, 2012). The argument is simply that the potential for impact is magnified hugely if the lever for change is used *often* (Thompson and Wiliam, 2008). In other words, long-cycle formative assessment does not provide enough information on enough days to have much of a chance to make a difference in students' learning.

Based on the frequency of the assessments in this study, and that they were *not* included in the teachers' everyday teaching, rather an assessment practice with a 14-day cycle was used, it is difficult to assess the academic effect, when it comes to, for example, better examination performances, as this was not noted in this study, even though the students promoted several of the activities in the new assessment practice as important for their own learning. It could be argued though that a 14-day cycle, over an entire school year, led to the establishment of routines, in the form of new working methods for assessments in mathematics, both for students and teachers. In addition, not including teaching in this study, therefore not "blurring the lines" between teaching and assessment, was also about making the process more manageable, digestible, for the teachers involved, which in itself is also an argument in KLT's action theory.

The students, on their part, actually promoted a need for a closer relationship between assessment and teaching, which a more frequent cycle could have entailed, and thus criticised their teachers for not using the results from the assessments more actively in their further teaching. This is an important criticism, from an important partner in a formative assessment practice, which should be included when teachers develop their own practice in a more formative direction.

On the other hand, it is important to point out that developing an assessment practice in a more formative way in higher education in 2021 should include more than academic achievements per se. It should involve facilitating the acquisition of positive motivational beliefs and self-esteem, developing a belief that learning through effort is relevant and feasible, encouraging students to take an active role in creating the learning process through, for example, engaging in such key learning strategies as reflection and peer dialogue,

generating their own feedback on learning and further effort, and learning to become more aware of and self-regulating in their own learning process. But here, too, the learning effects are probably greater if the assessment cycle is exceeded a couple of times a year.

However, in a formative mode, with a shared responsibility for learning, this means that teachers must understand how to use assessments to inform instructional decisions and must understand how to use them to strengthen their students' belief that successful learning outcome is within reach if they keep trying. In sum, teachers need to be capable of using assessment to both motivate their students and support their learning.

8.4. Strengths, Limitations and Future Research

This last section will reflect on the study's strength, what could have been done differently and finally, make recommendations for further research.

A key strength of this study is that it is highly relevant in today's national and international educational context. The findings related to this study both confirm *and* operationalise prominent trends in higher education over the last 10-15 years, both in a national and European context.

The length of this study is another strength. As the study lasted several years, work methods for the teachers and a concrete assessment design became established routines in their everyday work.

Observing and mapping students' achievement-goal patterns within this design over an entire school year can also be considered a strength, since there is little research on this in higher education. However, as this study is descriptive, which in itself can be characterised as a weakness, it cannot be claimed that the use of a learning-oriented assessment practice served as a causal reason between a learning-oriented assessment context and students' achievement-goal patterns. Clearly, these are important areas for future research.

Another limitation of the study is that the observational data were not included as part of the overall analysis material. This means that the observational data *were* included as an important part of the team's work methodology as such, and were thus part of the change process, but they were not scientifically analysed, as the interview data were. This is especially important in relation to the problem of overassimilation, where teachers believe they are enacting new practices but represent this practice in only superficial ways. The student interviews, for their part, are therefore important as further support, in that they confirm the content of the assessment practice, and the role of the teachers. Future research on the use of feedback in a higher education learning context could thus benefit from

complementary interventions studies, such as the one presented in this thesis, with a deeper and more theoretical analysis of observational data linked to professional development.

Furthermore, this study focused on improving teachers' assessment practice, and this was mainly done without a strong focus on the teachers' content knowledge in mathematics. Thus, the underlying assumption was that the teachers had the necessary content knowledge to successfully implement the various activities. This can also be considered a limitation, as content knowledge is considered an important part of professional development.

The advantage of not linking the study too holistically to mathematics, by applying general principles and having an overall focus on the development of work methods and active participation in formative activities such as self-assessment and reflection, feedback and dialogue, is that the results can be considered more transferable to other subject areas.

8.5. Conclusion

This thesis shows that work methods for educators and a concrete assessment design can help the present practice in higher education to come closer to the current empirical and theoretical understanding of formative assessment and feedback. The learning practices that will help to achieve this have both been established and are available through school education research, and the results of this intervention study support their value. The conclusion is that educators can develop their own feedback practice in a more formative direction by being responsible for creating a social learning environment where feedback is created through active student participation in such mastering-oriented and dialogic learning activities as dialogues between students, teacher-student dialogues, ideally with the support of response technology, and last but not least inner dialogues through reflection.

REFERENCES

Ajjawi, R., Tai, J., Dawson, P., & Boud, D. (2018). Conceptualising evaluative judgement for sustainable assessment in higher education. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.). *Developing Evaluative Judgement in Higher Education* (pp. 7-17). Routledge.

Anderman, E. M., Austin, C. C., & Johnson, D. M. (2002). The development of goal orientation. In A. Wigfield, & J. S. Eccles (Eds.). *A Vol. in the educational psychology series. Development of achievement motivation* (197–220). Academic Press.

Andersson, C., & Palm, T. (2018). Reason for teachers` successful development of a formative assessment practice through professional development – a motivation perspective. *Assessment in Education: Principles, Policy & Practice*, 25:6, 576-597.

Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Front. Educ,* 4:87. DOI: 10.3389/feduc.2019.00087.

ARG (2002) A*ssessment for Learning: 10 Principles*. Available on the Assessment Reform Group. website: www.assessment-reform-group.org.uk.

Aronson, J., Fried, C., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, 38, 113-125.

Askew, S., & Lodge, C. (2000). Gifts, ping-pong and loops – linking feedback and learning. In S. Askew (Ed.). *Feedback for Learning* (pp. 1-17). London: Routledge.

Bakhtin, M. (1981). *The Dialogic Imagination*, Holquist, C. (Ed.) Austin TX: University of Texas Press.

Bakhtin, M. (1984). *Problems of Dostoevsky's poetics*, Emerson, C. (ed. and transl.) University of Minnesota Press.

Bakhtin, M. (1986). *Speech genres & other late essays*. University of Texas press, Austin.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1-26.

Berg, B.L. (2007). Chapter five: Focus group Interviewing. In B.L. Berg (Ed.). *Qualitative Research Methods for the Social Sciences* (pp. 144-170). Pearson Education Inc.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning. Putting it into practice*. Open University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5, 7-74.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation & Accountability*, 21, 5-31.

Black, P., & Wiliam, D. (2012). Developing a theory of formative assessment. In J. Gardner (Eds.). *Assessment and Learning* (206-229). Sage Publications.

Blackwell, L. S., Trzesniewski, K., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78, 246-263.

Blair, A., Wyburn-Powel, A., Goodwin, M., & Shields, S. (2014). Can dialogue help to improve feedback on examinations? *Studies in Higher Education*, 39, 1039-1054.

Boud, D. (1999). Avoiding the traps: Seeking good practice in the use of self-assessment and reflection in professional courses. *Social Work Education*, 18:2, 121-132.

Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (2018). *Developing Evaluative Judgement in Higher Education*. Routledge.

Boud, D., Dawson, P., Bearman, M., Bennet, S., Joughin, G., & Molloy, E. (2016). Reframing assessment research: through a practice perspective. *Studies in Higher Education*, DOI: 10.1080/03075079.2016.1202913.

Boud, D., & Molloy, E. (2013). Changing conceptions of feedback. In D. Boud, & E. Molloy (Eds.). *Feedback in Higher and Professional Education: Understanding It and Doing It Well* (pp. 11-33). Oxon: Routledge.

Boyle, J. T., & Nicol, D. J. (2003). Using classroom communication systems to support interaction and discussion in large class settings. *Association for Learning Technology Journal*, 11, 43-57.

Brandi, U., & Sprogøe, J. (2019). *Det magiske øyeblikk. Kvalitativ analyse skritt for skritt*. Hans Reitzels Forlag.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academies Press.

Broadbent, J., Panadero, E., & Boud, D. (2018). Implementing summative assessment with a formative flavor: a case study in a large class. *Assessment & Evaluation in Higher Education,* 43:2, 307-322.

Brousseau, G. (1984). The crucial role of the didactical contract in the analysis and construction of situations in teaching and learning mathematics. In H. G. Steiner (Ed.). *Theory of mathematics education* (pp. 110-119). Occasionnel Paper 54, Bielefeld, Germany, University of Bielefeld, Institut für Didaktik der Mathematik.

Brown, S. (2010). Afterword. *Assessment and Evaluation in Higher Education*, 35, 347-349.

Bryan, C., & Clegg, K. (2019). *Innovative Assessment in Higher Education* (2th ed.). Routledge.

Bråten, I. (2002). Ulike perspektiver på læring. In Bråten, I. (red.): *Læring i sosialt, kognitivt og sosialt-kognitivt perspektiv* [learning in social, cognitive and social-cognitive perspective]. Oslo: Cappelen Akademisk forlag AS.

Bucharest Communiqué. 2012. *Making the Most of Our Potential: Consolidating the European Higher Education Area.* http://www.ehea.info/media.ehea.info/file/2012_Bucharest/67/3/Bucharest_Communique_2012_610673.pdf

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65:3, 245-281.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in higher education*, 31:2, 219-233.

Carless, D. (2013). Sustainable feedback and the development of student self-evaluative capacities. In S. Merry, M. Price, D. Carless, & M. Taras (Eds.). *Reconceptualising feedback in higher education: developing dialogue with students* (pp. 113-122). Oxon: Routledge.

Carless, D. (2017). Scaling up assessment for learning: progress and prospects. In D. Carless, S. Brigdes, C.K.W. Chan, & R. Glofcheski (Eds.). *Scaling up Assessment for Learning in Higher Education* (pp. 3-17). Singapore: Springer.

Carless, D. (2019). Feedback loops and the longer-term: towards feedback spirals. *Assessment and Evaluation in Higher Education*, 44:5, 705-714.

Carless, D., & Boud, D. (2018). The Development of Student Feedback Literacy: Enabling Uptake of Feedback. *Assessment & Evaluation in Higher Education* 43:8, 1315-1325.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, 36, 395-407.

Carr, W., & Kemmis, S. (1986). *Becoming Critical. Education, Knowledge and Action Research.* London: Falmer Press.

Charmaz, K. (2001). Grounded Theory. In J.A. Smith, R. Harre and L. Langenhove (Eds.). *Rethinking Methods in Psychology* (pp. 27-49). London: Sage Publications.

Charmaz, K. (2003). Grounded Theory. In J.A. Smith (Ed.). *Qualitative Psychology: A practical guide to research methods* (pp. 81-110). London: Sage Publications.

Cho, K., & MacArthur, C. (2011). Learning by Reviewing. *Journal of Educational Psychology*, 103:1, 73-84.

Cho, Y. H., & Cho, K. (2011). Peer Reviewers Learn from Giving Comments. *Instructional Science,* 39:5, 629-643.

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy*, 17, 305-308.

Coghlan, D. & Brannick, T. (2005). *Doing Action Research in Your Organization.* London: Sage Publications.

Creswell, J. (2003). *Research design: Qualitative, Quantitative and Mixed Methods Approaches*. Thousand Oaks CA: Sage Publications.

Cury, F., Elliot, A. J., Da Fonseca, D., & Moller, A. C. (2006). The social-cognitive model of achievement motivation and the 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, 90, 666-679.

Damsgaard, H. L. (2019). *Studielivskvalitet: Studenters erfaringer med og opplevelse av kvalitet i høyere utdanning* [Quality of student life: Students' experiences of quality in higher education]. Universitetsforlaget. DOI: https://doi.org/10.18261/10.9788215031354-2019

Daniel. L. M., & Poth, C. A. (2017). Relationships between pre-service teachers' conception of assessment, approaches to instruction, and assessment: An achievement goal theory perspective. *Educational Psychology*. DOI: 10.1080/01443410.2017.1293800.

Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2019). What makes for effective feedback: staff and student perspectives. *Assessment and Evaluation in Higher Education*, 44:1, 25-36.

Dekker, T., & Feijs, E. (2005). Scaling up strategies for change: change in formative assessment practices. *Assessment in Education*, 12:3, 237-254.

DeLuca, C., Luu, K., Sun, Y., & Klinger, D. A. (2012). Assessment for learning in the classroom: Barriers to implementation and possibilities for teacher professional learning. *Assessment Matters,* 4, 5-29.

DeVellis, R. F. (2012). *Scale development: Theory and applications* (3th ed.). London: Sage Publications.

Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, 20, 81-94.

Draper, S., & Nicol, D. (2013). Achieving transformational or sustainable educational change. In S, Merry, M. Price, D. Carless, & M. Taras (Eds.). *Reconceptualising Feedback in Higher Education: Developing Dialogue with Students* (pp. 190-203). London: Routledge.

Dunworth, K., & Sánchez, H.S. (2016). Perceptions of quality in staff-student written feedback in higher education: a case study. *Teaching in Higher Education*, 21:5, 576-589.

Dweck, C.S. (1999). *Self-theories: Their Role in Motivation, Personality, and Development*. Philadelphia: Psychology Press.

Dweck, C.S., & Legget, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.

Dweck, C.S., & Master, A. (2012). Self-Theories Motivate Self-Regulated Learning. In D. H. Schunk, & B. J. Zimmermann (Eds.). *Motivation and Self-Regulated Learning. Theory, Research and Application* (pp. 31-51). Routledge.

Dweck, C. S., & Molden, D. C. (2005). Self-theories. Their impact on competence motivation and acquisition. In A. Elliot, & C.S. Dweck (Eds.). *Handbook of Competence and Motivation* (pp. 122-140). The Guilford Press.

Dysthe, O. (1996). *Ulike Perspektiv på Læring og Læringsforskning*. [Different Perspectives on Learning and Learning Research] Cappelen Akademiske Forlag AS.

Dysthe, O. (2001). Sosiokulturelle teoriperspektiv på kunnskap og læring. In O. Dysthe (red.). *Dialog, Samspel og Læring*. [Dialogue, Interaction and Learning] Oslo: Abstrakt Forlag AS.

Dysthe, O. (2008). *Dialog, Samspel og Læring*. [Dialogue, Interaction and Learning] Oslo: Abstrakt Forlag AS

Dysthe, O. & Igland, M. (2001). Vygotskij og sosiokulturell teori. In O. Dysthe (red.). *Dialog, Samspel og Læring.* [Dialogue, Interaction and Learning] Oslo: Abstrakt Forlag AS.

Dysthe, O. & Igland, M. (2008). Vygotskij og sosiokulturell teori. In O. Dysthe (red.). *Dialog, Samspel og Læring*. [Dialogue, Interaction and Learning] (pp. 73-90). Oslo: Abstrakt Forlag AS.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34, 169-189.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology,* 51, 1058-1068.

ESG. 2015. *Standards and Guidelines for Quality Assurance in the European Higher Education Area.* https://enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf

European University Association (EUA) (2010). *Trends 2010: A decade of change in European higher education* (Andrée Sursock & Hanne Smidt). http://www.eua.be/Libraries/Publications_homepage_list/Trends2010.sflb.ashx

European University Association (EUA) (2019). Student-centred learning: approaches to quality assurance (Anna Gover, Tia Loukkola & Helene Peterbauer). https://eua.eu/resources/publications/884:the-quality-assurance-of-student-centred-learning-approaches-to-quality-assurance.html

Evans, C. (2013). Making sense of assessment in higher education. *Review of Educational Research*, 83, 70-120.

Falchikov, N. (2005). *Improving Assessment through Student Involvement*. London: Routledge–Falmer.

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59:4, 395-430.

Field, A (2009). *Discovering statistics using SPSS*. London: Sage Publications.

Forsythe, A., & Johnson, S. (2017) Thanks, but no-thanks for the feedback. *Assessment & Evaluation in Higher Education*, 42:6, 850-859.

Fryer, J. W., & Elliot, A. J. (2007). Stability and change in achievement goals. *Journal of Educational Psychology*, 99, 700-714.

Fryer, J. W., & Elliot, A. J. (2012). Self-Regulation of achievement goal pursuit. In D. H. Schunk, & B. J. Zimmerman (Eds.). *Motivation and Self-Regulated Learning. Theory, Research and Application* (pp. 52-68). Oxon: Routledge.

Gao, M. (2009). Students' voices in school-based assessment of Hong Kong: a case study. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.). *Student Perspectives on Assessment: What Students Can Tell Us About Assessment for Learning* (pp. 107-130). Charlotte: Information Age Publishing.

Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Mill Valley, CA: Sociology Press.

Grant, H. & Dweck, C.S. (2003) Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, 85:3, 541-553.

Hamberg, S., Bakken, P., & Damen, M. (2016). *NOKUTs notater. Tilbakemelding og veiledning i høyere utdanning. Hva forklarer studentenes misnøye [Department of Foreign Education's notes. Feedback and guidance in higher education. Explaining students' dissatisfaction].* https://www.nokut.no/contentassets/9989482e51f1473786a8037c4b71b46d/hamberg__bakken_damen_tilbakemelding_og_veiledning_i_hoyere_utdanning_2016.pdf

Harackiewicz, J. M., Barron, K. E., & Elliot, A. J. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist*, 33, 1-21.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-Term and Long-Term Consequences of Achievement Goals: Predicting Interest and Performance Over Time. *Journal of Educational Psychology*, 92:2, 316-330.

Harlen, W. (2012). The role of assessment in developing motivation for learning. In J. Gardner (Eds.). *Assessment and Learning* (pp. 171-183). Sage Publications.

Harris, L. R., & Brown, G, T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education*, 36, 101-111.

Hattie, J., & Clarke, S. (2019). *Visible Learning: Feedback*. Abingdon: Routledge

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research,* 77, 81-112.

Henderson, M., Ryan, T., & Phillips, M. (2019). The Challenges of Feedback in Higher Education. *Assessment and Evaluation in Higher Education,* 44:8, 1237-52.

Herman, W. (1990). Fear of failure as a distinctive personality trait measure of test anxiety. *Journal of Research and Development in Education*, 23, 180-185.

Hoidn, S. (2016). The Pedagogical Concept of Student-Centred Learning in the Context of European Higher Education Reforms. *European Scientific Journal*, 12, 1857-7881.

Houston, B. K., & Kelly, K. E. (1987). Type A behavior in housewives: Relation to work, marital adjustment, stress, tension, health, fear of failure, and self-esteem. *Journal of Psychosomatic Research,* 31, 55-61.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Fields Methods*, 18, 3-20.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.

James, M., McCormick, R., Black, P., D, Carmichaels, P., Drummond, M-J., Fox, A., MacBeath, J., Marshall, B., Pedder, D., Procter, R., Swaffield, S., Swann, J., & Wiliam, D. (2007). *Improving Learning how to Learn, in Classrooms, Schools and Networks*. London: Routledge.

Johannessen, A. (2009*). Introduction to SPSS*. Abstrakt forlag AS.

Johannessen, A., Tufte. P.A., & Kristoffersen, L. (2016*). Introduction to Social Science Methods*. (5th ed.). Abstrakt forlag AS.

Johnson, C., & Molloy, E. (2018). Building evaluative judgement through the process of feedback. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.). *Developing Evaluative Judgement in Higher Education* (pp. 166-175). Routledge.

Jonsson, A., Lundahl, C., & Holmgren, A. (2015). Evaluating a large-scale implementation of Assessment for Learning in Sweden. *Assessment in Education: Principles, Policy & Practice*, 22:1, 104-121.

Kaufman, J. H., & Schunn, C. D. (2011). Students' Perceptions about Peer Assessment for Writing: Their Origin and Impact on Revision Work. *Instructional Science* 39: 387-406.

Knapstad, M., Heradstveit, O., & Sivertsen, B. (2018). *SHoT 2018. Studentenes helse- og Trivselsundersøkelse [Students' health and well-being survey].* https://www.uio.no/studier/om/laringsmiljo/shot/rapportene/ shot-2018-studentenes-helse-og-trivselsundersokelse.pdf

Kvale, S., & Brinkmann, S. (2009). *Det Kvalitative Forskningsintervju* (2th ed.). Oslo Gyldendal Akademiske.

Leahy, S., & Wiliam, D. (2012). From teachers to schools: Scaling up professional development for formative assessment. In J. Gardner (Eds.). *Assessment and Learning* (pp. 49-71). Sage Publications.

Leuven/Louvain-la-Neuve Communiqué (2009). *The Bologna Process 2020 – The European Higher Education Area in the New Decade.* http://www.ond.vlaanderen.be/hogeronderwijs/bologna/conference/documents/leuven_louvain-la-neuve_communiqu é_april_2009.pdf

Lid, S. E., Pedersen, L. F., & Damen, M. L. (2018). *Teacher* Survey *2017: Main Trends*. https://www.nokut.no/globalassets/studiebarometeret/underviserundersokelsen/lid_pedersen_damen_underviserundersokelsen 2017_hovedtendenser_2-2018.pdf

Liu, N., & Carless, D. (2006). Peer Feedback: The Learning Element of Peer Assessment. *Teaching in Higher Education,* 11:3, 279-290.

Lund, T. & Haugen, R. (2006). *Forskningsprosessen.* [The Research Process] Oslo: Unipub forlag.

MacLellan, E. (2001). Assessment for learning: The differing perceptions of tutors and students. *Assessment and Evaluation in Higher Education*, 26:2, 307-318.

Midgley, C., Kaplan, A., Middleton, M., Urdan, T., Maehr. M. L., Hicks, L., & Roeser, R. W. (1998). Development and validation of scales assessing students' achievement goal orientation. *Contemporary Educational Psychology*, 23, 113-131.

Ministry of Education, Research and Church Affairs. (2001). *Gjør din plikt – krev din rett. Kvalitetsreform av høyere utdanning* (Report to Parliament nr. 27 (2000–2001) [Do your duty – demand your rights. Quality reform for higher education]. https://www.regjeringen.no/no/dokumenter/stmeld-nr-27-2000-2001-/id194247/

Ministry of Education and Research (2008). *Statusrapport for Kvalitetsreformen i høgre utdanning* (Report to parliament nr. 7 (2007-2008) [Status report on the Quality Reform in Higher Education]. https://www.regjeringen.no/no/dokumenter/Stmeldnr-7-2007-2008-/id492556/

Ministry of Education and Research. (2017). *Kultur for kvalitet i høyere utdanning* [Culture for Quality in Higher Education]. (Report to parliament St. 16 (2016–2017)). https://www.regjeringen.no/no/dokumenter/meld.-st.-16-20162017/id2536007/

Molden, D. C., & Dweck, C. S. (2006). Finding "meaning" in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist,* 61, 192-203.

Molloy, E., Noble, C., & Ajjawi, R. (2019). Attending to emotions in feedback. In M. Henderson, R. Ajjawi, D. Boud, & E. Molloy (Eds.). *The Impact of Feedback in Higher Education. Improving Assessment Outcomes for Learners* (pp. 83-106). London: Palgrave Macmillan.

Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education*, 45:4, 527-540.

Muis, K. R., & Edwards, O. (2009). Examining the stability of achievement goal orientation. *Contemporary Educational Psychology*, 34, 265-277.

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: perceptions of students and academics. *Assessment and Evaluation in Higher Education*, 42:2, 266-288.

Nicol, D.J. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment and Evaluation in Higher Education*, 35, 501-517.

Nicol, D., & Draper, S. (2009). A blueprint for transformational organisational change in higher education: reap as a case study. In T. Mayes, D. Morrison, H. Mellar, P. Bullen, & M. Oliver (Eds.). *Transforming Higher Education through Technology-Enhanced Learning* (pp. 191-207). The higher Education Academy.

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

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: a peer review perspective. *Assessment & Evaluation in Higher Education*, 39:1, 102-122.

Norris, N. (1997). Error, bias and validity in qualitative research. *Educational Action Research*, 5:1, 172-176.

Nussbaum, D., & Dweck, C. S. (2006). *Self-theories and self-esteem maintenance.* Unpublished manuscript, Stanford University, Stanford, CA.

Nesje, K. & Støren, L. A. (2018). Kandidatundersøkelsen 2017. *Nyutdannede masteres møte med arbeidslivet og vurdering av relevans, studiekvalitet og læringsutbytte* (NIFU rapport 22/2018) [Candidate survey 2017. Newly graduated Master's degree students' encounter with working life and evaluation of relevance, quality of studies and learning outcome]. https://www.nifu.no/publications/1602276/

O`Donovan, B., Rust, C., & Price, M. (2016). A scholarly approach to solving the feedback dilemma in practice. *Assessment and Evaluation in Higher Education*, 41:6, 938-949.

Panadero, E., Broadbent, J., Boud, D., & Lodge, J. M. (2018). Using formative assessment to influence self- and co-regulated learning: The role of evaluative judgement. *European Journal of Psychology of Education.* DOI: 10.1007/s10212-018-0407-8

Panadero, E., Brown, G. L., & J. W. Strijbos. (2015). The Future of Student Self-assessment: A Review of Known Unknowns and Potential Directions. *Educational Psychology Review*. DOI:10.1007/s10648-015-9350-2

Panadero, E., Brown, G. T. L., & Strijbos, J-W. (2016a). The Future of Student Self-Assessment: A Review of Known Unknowns and Potential Directions. *Educ Psychol Rev*, 28, 803–830.

Panadero, E., Jonsson, A., & Strijbos, J-W. (2016b) Scaffolding Self-Regulated Learning Through Self-Assessment and Peer Assessment: Guidelines for Classroom Implementation. In: D, Laveault, & L. Allal (Eds.). *Assessment for Learning: Meeting the Challenge of Implementation.* The Enabling Power of Assessment, vol 4. Springer, Cham.

Panadero, E., & Brown, G. T. L. (2017). Teachers' reasons for using peer assessment: Positive experience predicts use. *European Journal of Psychology of Education*, 32:1, 133-156.

Panadero, E., Lipnevich, A., & Broadbent, J. (2019). Turning Self-Assessment into Self-feedback. In M, Henderson, R. Ajjawi, D. Boud, & E, Molloy (Eds.). *The Impact of Feedback in Higher Education. Improving Assessment Outcomes for Learners* (pp. 147-166). London: Palgrave Macmillan.

Pedder, D. (2006). Organisational conditions that foster successful classroom promotion of learning how to learn. *Research Papers in Education*, 21, 171-200.

Pedder, D., & James, M. (2012). Professional learning as a condition for assessment for learning. In J. Gardner (Eds). *Assessment and Learning: Second Edition* (pp. 33-48). London: Sage Publications.

Pedder, D., James, M., & MacBeath, J. (2005). How teachers' value and practice professional learning. *Research Papers in Education,* 20, 209-243.

Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. *Assessment in Education: Principles, Policy and Practice,* 5:1, 85-102.

Peterson, E. R., & Irving, S. E. (2008). Secondary School Students' Conceptions of Assessment and Feedback. *Learning and Instruction,* 18:3, 238-250.

Pitt, E. & Norton, L. (2017). 'Now that's the feedback I want!' Students' reactions to feedback on graded work and what they do with it. *Assessment and Evaluation in Higher Education.* 42:4. 499-516.

Postholm, M.B. (2005). *Kvalitativ Metode. En Innføring med Fokus på Fenomenologi, Etnografi og Kasusstudier.* [Qualitative methodology. An introduction focusing on phenomenology, ethnography and case studies] Oslo: Universitetsforlaget.

Postholm, M.B. (2007). *Forsk Med. Lærere og Forskere i Læringsarbeid.* [Teachers and researchers in learning work] Damn & Søn AS.

Postholm, M. B. (2010). *Kvalitativ Metode: En Innføring med Fokus på Fenomenologi, Etnografi og Kasusstudier.* [Qualitative method: an introduction to phenomenology, ethnography and case studies] Oslo: Universitetsforlaget.

Postholm, M.B. & Madsen, J. (2006). The researcher's role. An ethical dimension. *Outlines,* 1, 49-60.

Price, M., Handley, K., O`Donovan, B., Rust, C., & Millar, J. (2013). Assessment feedback: an agenda for change. In S. Merry, M. Price, D. Carless, & M. Taras (Eds.). *Reconceptualising Feedback in Higher Education* (pp. 41-53). Routledge.

Reason, P., & Bradbury, H. (2008) *The Sage Handbook of Action Research: Participative Inquiry and Practice.* Sage Publications.

Ringdal, K. (2018). *Enhet og Mangfold. Samfunnskapelig Forskning og Kvantitativ Metode*. [Unity and Diversity. Social research and quantitative methods] Fagbokforlaget.

Robins, R. W., & Pals, J. L. (2002) Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity*, 1, 313-336.

Robson, C. (2011). *Real World Research. A resource for Social Scientist and Practitioner Researcher, second edition.* Oxford: Blackwell Publishers.

Rommetveit, R (1974). *On Message Structure*. London/New York: Wiley.

Rommetveit, R. (1991). Dominance and asymmetries in A Doll's House. In I. Markova, & K. Foppa (Eds.). *Asymmetries in Dialogue*. Hemel Hempstead: Harvester/Wheatsheaf.

Sadler, D.R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5, 77-84.

Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35, 535-550.

Sadler, D. R. (2016). Three in-course assessment reforms to improve higher education learning outcomes. *Assessment and Evaluation in Higher Education*, 41:7, 1081-1099.

Sadler, P. M., & Good, E. (2006). The Impact of Self- and Peer-Grading on Student Learning. *Educational Assessment,* 1, 1-31.

Säljö, R. (2001). *Læring i Praksis: Et Sosiokulturelt Perspektiv*. Cappelen akademisk.

Schneider, C. M., & Randel, B. (2010). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. Andrade, & G.J. Cizek (Eds.). *Handbook of Formative Assessment* (pp. 251-276). Routledge.

Schunk, D.H., Pintrich, P. R., & Meece, J.L. (2010). *Motivation in education: Theory, Research, and Applications* (3th ed.). Upper Saddle River, NJ: Prentice Hall.

Schutz, A. (1972). *The Phenomenology of the Social World*. Northwestern University Press.

Senko, C., & Harackiewicz, J. M. (2005). Regulation of achievement goals: The role of competence feedback. *Journal of Educational Psychology*, 97, 320-336.

Silverman, D. (2011). *Interpreting Qualitative Data. A guide to the principles of qualitative Research.* (4th ed.). SAGE Publications Ltd.

Stiggins, R. (1999). Teams. *Journal of Staff Development*, 20:3, 17-21.

Stiggins, R. (2010). Essential Formative Assessment Competencies for Teachers and School leaders. In H. Andrade & G.J. Cizek (Eds.). *Handbook of Formative Assessment* (pp. 233-250). Routledge.

Strauss, A. & Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* Thousand Oaks, CA: Sage Publications.

Svensson, L., Ellström, P. E., & Brulin, G. (2007). Introduction – on interactive research. *International Journal of Action Research*, 3:3, 233-249.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics*. (5th ed.). Boston: Pearson Education Inc.

Thagaard, T. (2016). *Systematikk og Innlevelse. En Innføring i Kvalitativ Metode.* (4th ed.). Vigmostad & Bjørke AS.

Thompson, M., & Wiliam, D. (2008). Tight but loose: a conceptual framework for scaling up school reforms. In. E.C. Wylie (Eds.). *Tight but Loose: Scaling Up Teacher Professional Development in Diverse Contexts*. Research report RR-08-29. (pp. 1-44). Princeton, NJ: Educational Testing Service.

Timperley, H., & Alton-Lee, A. (2008). Reframing teacher professional learning: an alternative policy approach to strengthening valued outcomes for diverse learners. *Review of Research in Education*, 32, 328-369.

Topping, K. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research*, 68:3, 249-276.

Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes.* Cambridge, MA: Harvard University Press.

Wiggen, K. S., Øygarden, K. F., Bakken, P., & Pedersen, L. F. (2020). *Student Survey 2019: Main Trends* (Studiebarometeret rapport 1/2019). https://www.nokut.no/globalassets/studiebarometeret/2019/studiebarometeret-2018_hovedtendenser_1-2019.pdf

Wiliam, D. (2007). Keeping learning on track: classroom assessment and the regulation of learning. In F. K. Lester Jr (Ed.). *Second Handbook of Mathematics Teaching and Learning* (pp. 1053-1098). Greenwich, CT: Information Age Publishing.

Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: what will it take to make it work? In C. A. Dwyer (Eds.). *The Future of Assessment: Shaping Teaching and Learning* (pp. 53-82). Mahwah, NJ: Lawrence Erlbaum Associates.

Winstone, N. & Carless, D. (2020). *Designing Effective Feedback Processes in Higher Education. A Learning-Focused Approach*. London: Society for Research into Higher Education.

Winstone, N. E., & Boud, D. (2020). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education,* DOI: 10.1080/03075079.2020.1779687

Yan, Z. (2016). Student Self-assessment Practices: The Role of Gender, School Level and Goal Orientation. *Assessment in Education: Principles, Policy & Practice*, 25:2, 183-199.

Yan, Z., & Brown, G. T.L. (2017). A cyclical self-assessment process: towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education*, 42:8, 1247-1262.

Yerevan Communiqué (2015). *[No title].* http://www.ehea.info/media.ehea.info/file/2015_Yerevan/70/7/YerevanCommuniqueFinal_613707.pdf

Yorke, M., & Knight, P. (2004) Self-theories: some implications for teaching and learning in higher education. *Studies in Higher Education*, 29:1, 25-37.

Zimmermann, B.J., & Schunk, D.H. (2012). An Essential Dimension of Self-Regulated learning. In D. H. Schunk & B. J. Zimmermann (Eds.). *Motivation and Self-Regulated Learning. Theory, Research and Application* (pp. 1-30). Routledge.

APPENDICES

Appendix 1. Research permit from the Norwegian Social Science Data Service (NSD), concerning the achievement-goal questionnaire.

Norsk samfunnsvitenskapelig datatjeneste AS
NORWEGIAN SOCIAL SCIENCE DATA SERVICES

Harald Hårfagres gate 29
N-5007 Bergen
Norway
Tel +47 55 58 21 17
Fax +47 55 58 96 50
nsd@nsd.uib.no
www.nsd.uib.no
Org.nr 985 321 884

Gabrielle Hansen-Nygård
Avdeling for teknologi
Høgskolen i Sør-Trøndelag
7004 TRONDHEIM

Vår dato: 12.09.2012 Vår ref:31399 / 3 / LT Deres dato: Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 08.09.2012. Meldingen gjelder prosjektet:

31399	*Vurdering for motivasjon med mobiltelefonen i klasserommet*
Behandlingsansvarlig	*NTNU, ved institusjonens øverste leder*
Daglig ansvarlig	*Gabrielle Hansen-Nygård*

Personvernombudet har vurdert prosjektet og finner at behandlingen av personopplysninger er meldepliktig i henhold til personopplysningsloven § 31. Behandlingen tilfredsstiller kravene i personopplysningsloven.

Personvernombudets vurdering forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, eventuelle kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, http://www.nsd.uib.no/personvern/forsk_stud/skjema.html. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, http://pvo.nsd.no/prosjekt.

Personvernombudet vil ved prosjektets avslutning, 15.06.2013, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

Vigdis Namtvedt Kvalheim

Lis Tenold

Lis Tenold tlf: 55 58 33 77
Vedlegg: Prosjektvurdering

Personvernombudet for forskning

Prosjektvurdering - Kommentar

Prosjektnr: 31399

Det gis skriftlig informasjon og innhentes skriftlig samtykke. Personvernombudet finner skrivet godt utformet, men forutsetter at det angis dato for prosjektslutt, her 15.06.2013.
Personvernombudet legger til grunn for sin godkjenning at revidert skriv ettersendes før det tas kontakt med utvalget.

Innsamlede opplysninger anonymiseres ved prosjektslutt, senest 15.06.2013. Med anonymisering innebærer at navnelister slettes/makuleres, og ev. kategorisere eller slette indirekte personidentifiserbare opplysninger. Ved publisering vil ingen enkeltpersoner kunne gjenkjennes.

Appendix 2. Letter and signature for the students informed consent regarding achievement-goal questionnaire.

HØGSKOLEN I SØR-TRØNDELAG
AVDELING FOR TEKNOLOGI
PROGRAM FOR ALLMENNFAG

Informert samtykke

Forespørsel om å delta i undersøkelse:

Jeg er stipendiat ved Høgskolen i Sør-Trøndelag (HiST), avd. for teknologi (AFT). Tema for doktorgradsprosjektet mitt er studentlæring og motivasjon i matematikkfaget. Målet mitt er å forbedre undervisnings- og vurderingspraksisen i dette faget. For å kunne foreslå tiltak til forbedringer trenger jeg å vite mere om studentgruppen ved faget. Dette for å få bedre innsikt i hvordan dere som studenter opplever matematikkfaget. Spørsmålene vil bli rettet mot hvilke mål dere setter dere i dette faget, hvordan dere opplever vurderingssituasjoner som tester, og hvordan dere opplever matematikklæreren og selve klasseromsmiljøet.

HiST/AFT er ansvarlig for denne undersøkelsen. Undersøkelsen ledes av stipendiat Gabrielle Hansen.

Det er frivillig for deg å delta i undersøkelsen, og all informasjon vil bli behandlet konfidensielt. Datamaterialet vil bli anonymisert ved prosjektslutt, senest ved utgangen av 2014. Resultatene vil bli presentert slik at ingen enkeltpersoner kan gjenkjennes.

Hvis det er noe du lurer på kan du ta kontakt med meg via e-post: gabrielle.hansen@hist.no.

Undersøkelsen er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelige datatjeneste (NSD).

Samtykkeerklæring:

Jeg har mottatt skriftlig informasjon og er villig til å delta i undersøkelsen.

Navn: _____ Dato: _____

BLOKKBOKSTAVER:

Postadresse
Høgskolen i Sør-Trøndelag
Avdeling for teknologi
Program for allmennfag
7004 Trondheim

Kontoradresse
Gunnerusgt. 1
Trondheim

E-postadresse

Telefon
73559480

Telefax
73559481

Appendix 3. Interview guide teacher interviews.

Intervjuguide - Refleksjonssamtale

Flervalgsprøve/Regneprøve

1. Generelle inntrykk: Hvilke opplevelser sitter du igjen med etter arbeidskravet i går?

- Hvordan synes du det gikk i går?
    - Egen innsats?
        - Introduksjon
        - Del 1: Den individuelle prøven
        - Del 2: Gjennomgangen
    - Tanker omkring oppmøte
    - Annet?

2. Hvordan synes du det fungerer å arbeide ut ifra bestemte prinsipper?

Pedagogisk rammeverk: A good feedback practice;

1. helps clarify what good performance is;
2. facilitates the development of self-assessment in learning;
3. delivers high-quality information to students about their learning;
4. encourages teacher and peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape teaching.

<u>Denne gangen hadde vi spesielt fokus på:</u>

1. Mål og tydeliggjøring
2. Selvvurdering og refleksjon
3. Oppmuntre til dialog
    - Lærer – student
    - Student – student
4. Oppmuntre til positiv motivasjon (positive motivational self-beliefs)
5. levere kvalitetsinformasjon
6. Lukke gapet

Gå igjennom og diskuter tanker og opplevelser knyttet til implementering av de ulike prinsippene:

- Hvilken/e målsetning/er hadde du med de ulike prinsippene?
- Hvordan implementerte du prinsippene?
    - Hvilke valg tok du?
- Opplevelser knyttet til disse?
- Hvordan opplevde du at dette fungerte i praksis?
- Er det noe du føler du kunne gjort annerledes?
- Hva er du fornøyd med?
- Er det noe du er mindre fornøyd med?

3. Hvordan påvirker det deg som underviser å jobbe på en slik måte?

- Hvordan ***forberedte*** du deg til dette arbeidskravet?
- Hadde du satt deg noen mål med arbeidskravet?
    - ***Hva ønsket du å oppnå?***

4. Hvordan opplevde du studentene?

5. Hva lærte du av dette arbeidskravet?

6. Hvordan ser du for deg neste arbeidskrav?

- Målsetninger
- Forberedelser

7. Er det noe mere du vil fortelle meg?

Appendix 4. Interview guide student interviews.

Intervjuguide – arbeidskrav matematikk på forkurs vår 2013

1. Introduksjon / bakgrunnskunnskap

1.1 Presentasjon av meg selv og tema for intervju.

Det jeg har lyst å prate med dere litt om i dag er, som nevnt tidligere, vurderingsformen på forkurs i matematikk, altså arbeidskravene deres. Dere har det siste skoleåret gjennomført til sammen 12 arbeidskrav i matematikk, fordelt på seks regneprøver og seks flervalgsprøver. Jeg har i denne perioden fulgt flere av matematikklærere, og begynner å danne meg et bilde av hvordan de tenker omkring vurderingene. Jeg synes imidlertid det er minst like viktig, om ikke enda viktigere, å få et innblikk i hvordan dere som studenter har opplevd disse arbeidskravene. Kravene er jo til for dere, så hva dere synes om dem er en meget viktig del av doktorgradsprosjektet mitt, som omhandler nettopp vurdering i høyere utdanning.
Hovedtema for samtalen i dag er derfor deres opplevelser i forhold til arbeidskravene i matematikk.

1.2 Tid

- Ikke fastsatt tid på forhånd, heller ikke lagt inn pause, men det er bare å si i fra

1.3 Etiske forhåndsregler

Deltagelsen deres er 100 % frivillig. Dere svarer på det dere har lyst til å svare på, og dere har lov å avbryte, eller avslutte intervjuet når det måtte passe dere. Intervjuene vil selvsagt være anonyme; navnene dere vil bli byttet ut under transkriberingen, så det vil ikke bli mulig å spore deres identitet i det ferdige materialet. Opptakene vil også bli slettet etter transkribering.

1.4 Introduksjon av informantene

Jeg tenkte vi kunne starte med en liten runde rundt bordet, slik at alle kan gi en kort introduksjon av seg selv, da tenker jeg henholdsvis:

- Navn, alder og studiebakgrunn
    - Hvorfor begynte du på forkurs; Framtidsutsikter; Trives dere?

2. Generell oppfatning/opplevelser i forhold til arbeidskrav i matematikk:

Vi starter litt åpen og generelt, så fokuserer vi inn på ulike temaer etter hvert. Som nevnt i sted, er det gjennomført til sammen 12 arbeidskrav i matematikk på forkurs, fordelt på 6 regneprøver og 6 flervalgsprøver.

- ***2.1 Hvilke tanker og opplevelser sitter dere igjen med etter arbeidskravene?***

➔ ***Når jeg sier arbeidskrav i matematikk – hva tenker dere da?***

  - Har dere møtt opp til arbeidskravene?
    - Hvorfor/hvorfor ikke?
  - Hva synes dere om antall arbeidskrav?
  - Nytt arbeidskrav ca. hver 14 dag
  - Er det en annen måte dere ser for dere at dere kunne blitt vurdert på?

Del 1. Regneprøver

Flervalgsprøver og regneprøver er jo litt ulike måter å gjennomføre arbeidskrav på, så jeg kunne tenke meg å separere disse fra hverandre, og prate om de hver for seg. Vi kan jo begynne med regneprøvene.

- ***3.1 For å starte på begynnelsen her; hvordan vil dere beskrive en regneprøve?***
  - Kan dere fortelle meg hvordan en regneprøve foregår i klassen deres?
    - Hva skjer under en regneprøve?
  - Hva synes dere om regneprøvene?

- ***3.2 Hva opplever dere som formålet med regneprøvene?***
  - *Hva oppfatter dere som **faglærers mål** med denne formen for arbeidskrav?*
    - Hva trur dere faglærer vil med regneprøvene?
  - *Hva er **deres mål** med regneprøvene?*
    - Å få godkjent/bestått; gjøre det så bra som mulig; lære; innsats; lokalisere misforståelser; annet?

Regneprøvene er på mange måter todelt: Del 1. individuell prøve. Pause. Del 2. gjennomgang av prøven.

- **3.3 *Hvis vi retter fokus mot selve gjennomgangen, hvordan vil dere beskrive den?***
  - Hva skjer under gjennomgangen av regneprøvene?

- ***3.6. Lærer gjennomgang: hva synes dere om at lærer går gjennom og forklarer oppgavene?***
    - Føler dere at dere får en god nok forklaring rundt oppgavene?
    - Hvordan opplever dere dette i forhold til egen *læring*?
        - Lærer dere av læreres gjennomgang?
            - På hvilken måte da?
        - Er det noe som kunne vært gjort annerledes her?

- ***3.7. Hva opplever dere som deres oppgave under gjennomgangen?***
    - Hva gjør dere under gjennomgangen av regneprøven?
        - Lytter til lærer; skriver av tavlen; vurderer seg selv; annet?

- ***3.8. Hvordan vil dere vurdere eget engasjement under selve gjennomgangen?***
    - Følger dere med; interessert; passiv; deltagende?

- ***3.9. Etter at lærer har gjennomgått en oppgave på tavlen, blir dere ofte bedt om å reflektere i et par minutt, hvordan oppfatter dere denne beskjeden?***
    - *Altså, hva legger dere i det å reflektere?*
        - Hva gjorde dere under refleksjonene?
        - Hvordan tenkte dere?
        - Hva betyr det å reflektere?
    - Opplever dere at dette hadde noen form for nytteverdi for dere?
        - Hvis ja, på hvilken måte da?
        - Hvis nei; hva kunne vært gjort annerledes?
    - Fikk dere nok tid til å reflektere?

- ***3.10. Hvordan opplever dere regneprøvene i forhold til egen læring?***
- ***3.11. Hvordan opplever dere regneprøvene i forhold til egen motivasjon for å lære?***
- ***3.12. Til slutt: er det noe som kunne vært gjort annerledes i forhold til regneprøvene?***

Del 2. Flervalgsprøver

Da kan vi jo gå over til den andre formen for arbeidskrav, flervalgsprøvene. Vi starter bare på begynnelsen her og:

- ***4.1 Vi starter på begynnelsen her og; hvordan vil dere beskrive en flervalgsprøve?***
    - Kan dere fortelle meg hvordan en flervalgsprøve foregår i klassen deres?
        - Hva skjer under en flervalgsprøve?
    - Hva synes dere om flervalgsprøvene i matematikk?

- ***4.2 Hva opplever dere som formålet med flervalgsprøvene?***
    - Hva er målet med denne formen for arbeidskrav?
        - Hvorfor flervalgsprøver?
    - Er målene her annerledes i forhold til regneprøvene?
        - Faglærers mål?
        - Deres mål?

    - ***Hvordan vil dere beskrive kontinuiteten på målene deres,***
        - Altså, har dere samme mål for hver prøve?
            - Har de ulike mål i forhold til ulike arbeidskrav?

    - ***Tenkte dere annerledes om målet med arbeidskravene tidligere i høst?***
        - Tenkte deres annerledes om hensikten med arbeidskravene i høst enn nå på slutten av skoleåret?

Flervalgsprøvene er i likhet med regneprøvene todelt: Del 1. individuell test. Pause. Del 2. gjennomgang av testen.

- ***4.4 Hvordan vil dere beskrive gjennomgangen flervalgsprøvene?***
    - Hva skjer under gjennomgangen på denne formen for arbeidskrav?
    - Følte dere at dere fikk en god nok forklaring på hvorfor alternativene viste seg å være riktig eller feil?
    - Hvordan oppleves gjennomgangen i forhold til egen *læring*?

- ***4.5. Hva opplever dere som deres oppgave under gjennomgangen av flervalgsprøvene?***

Under flervalgsprøvene sitter dere i grupper og i gjennomgangen ble det lagt opp til gruppediskusjoner, hva synes dere om det?

- ***4.5 Hvilke opplevelser har dere gjort dere i forhold til gruppediskusjonene?***
    - Beskriv hva som skjer under en gruppediskusjon?
        - Diskuterte dere?
            - På hvilken måte?
            - Hvorfor ikke?
    - Opplever dere at dette hadde noen form for nytteverdi for dere?
        - Hvordan vil dere vurdere slike former for gruppediskusjoner i forhold til egen forståelse og læring?
        - Hvor viktig opplever dere andre studenter er i forhold til egen læring?
    - Burde det ha vært flere/færre slike diskusjoner?

- ***4.6 Under flervalgsprøvene brukte dere også SRS, hvilke tanker har dere gjort dere i forhold til dette?***
    - Hva synes dere om å svare på et bestemt testspørsmål en gang til ved bruk av SRS?
    - Hva er meningen med en SRS-runde etter en diskusjon?
    - Hvordan vil dere vurdere egen deltagelse i forhold til bruk av SRS?
        - Deltok dere: svarte dere på spørsmålene?
    - Har dere noen ganger unnlat å svare på et spørsmål?
        - Hvorfor

- ***4.9. Hva med klassediskusjon, hvordan har dere opplevd dem?***
    - Faglærer stiller ofte en del spørsmål ut i klassen – svarer dere?
        - Hvorfor/hvorfor ikke

- ***4.10. Hvordan opplever dere flervalgsprøvene i forhold til egen læring?***
- ***4.11. Hvordan opplever dere flervalgsprøvene i forhold til egen motivasjon for å lære?***
- ***4.12. Til slutt: er det noe som kunne vært gjort annerledes i forhold til flervalgsprøvene?***

Del 3: Avrunding og avslutning

5.1 Hvordan synes dere faglærer har fremstått under arbeidskravene?

- Virker faglærer forberedt til arbeidskravene?
    - På hvilken måte?
    - Hva kunne faglærer gjort annerledes?
- Hvilken holdning føler dere faglærer har til arbeidskravene?
- Hvordan fremstår faglærer for dere?
    - Positiv; Engasjert; Passiv; Stresset; Negativ?

6. Sist, men ikke minst; har dere ellers noen kommentarer i forhold til arbeidskravene i matematikk?

- For å summere; *Hva vil dere beskrive som deres fremste opplevelse i forhold til arbeidskravene?*

Appendix 5. Achievement-goal questionnaire.

DINE TANKER OM DEG SELV I MATEMATIKKFAGET - 1

Ikke skriv her!

Tema for doktorgradsprosjektet mitt er studentlæring og studentmotivasjon. Målet mitt er å forbedre undervisnings- og vurderingspraksisen i dette faget. For å kunne foreslå tiltak til forbedringer, trenger jeg å vite mer om hvordan du som student opplever faget. Jeg håper derfor du kan svare på noen spørsmål.

Det er frivillig å delta i undersøkelsen, og all informasjon vil bli behandlet konfidensielt. Datamaterialet vil bli anonymisert ved prosjektslutt, senest ved utgangen av 2013. Resultatene vil bli presentert slik at ingen enkeltpersoner kan gjenkjennes.

Dette er ikke en prøve – det er ingen «riktige» eller «gale» svar. Nedenfor finner du flere påstander, og du krysser av for det svaret som viser hva du føler.

For å få så sikre resultater som mulig, stiller jeg enkelte ganger flere spørsmål om samme sak. Noen av spørsmål kan derfor se ganske like ut. Les hvert spørsmål nøye, og besvar det uten å tenke på de andre spørsmålene.

Takk for at du er villig til å delta!

Gabrielle Hansen
stipendiat

HiST TRONDHEIM

NTNU
Det skapende universitet
Program for lærerutdanning

LES DETTE FØR DU STARTER!	Skjemaet skal leses maskinelt. Vennligst følg disse reglene: • *Bruk svart/blå kulepenn. Skriv tydelig, og ikke utenfor feltene.* ***Kryss av slik:*** ☒. • *Feilkryssinger kan annulleres ved å fylle hele feltet med farge. Kryss så i rett felt.* • *Sett bare ett kryss på hvert spørsmål om ikke annet er oppgitt.*

1. Kjønn: Kvinne .. ☐ 1 Mann .... ☐ 2
2. Alder: ☐☐
3. Klasse: 1FA.... ☐ 1 1FB .... ☐ 2 1FC .... ☐ 3 1FD ... ☐ 4 1FE.... ☐ 5

Sett ett kryss for hvert utsagn:

	Helt usant 1	*Litt usant* 2	*Litt sant* 3	*Helt sant* 4
4. Jeg liker å jobbe med matematikkoppgaver jeg lærer av, selv om jeg gjør mange feil	☐	☐	☐	☐
5. Grunnen til at jeg jobber med matematikkfaget er at læreren min ikke skal tro at jeg kan mindre enn de andre	☐	☐	☐	☐
6. Jeg synes jeg gjør det veldig bra i matematikkfaget	☐	☐	☐	☐
7. Jeg synes det vi lærer i dette faget er interessant	☐	☐	☐	☐
8. Jeg prøver for enhver pris å unngå å mislykkes i matematikkfaget	☐	☐	☐	☐
9. Jeg føler meg veldig bra hvis jeg er den eneste som kan svare på lærerens spørsmål i matematikktimen	☐	☐	☐	☐
10. Jeg føler meg ofte usikker før en matematikkprøve	☐	☐	☐	☐
11. En viktig grunn til at jeg jobber med matematikkfaget er at jeg liker å lære noe nytt	☐	☐	☐	☐
12. Det er veldig viktig for meg at jeg ikke fremstår som dum i matematikktimene	☐	☐	☐	☐
13. Jeg har problemer med å forstå mye av det som gjennomgås i matematikktimene	☐	☐	☐	☐
14. Jeg tror jeg vil være i stand til å bruke det jeg lærer i matematikkfaget i andre fag	☐	☐	☐	☐

Undersøkelsen gjennomføres med bistand fra SVT-IT, NTNU

Snu arket!

Husk: Bare ett kryss på hvert spørsmål.

	Helt usant 1	*Litt usant* 2	*Litt sant* 3	*Helt sant* 4
15. Jeg unngår ofte enkelte matematikkoppgaver fordi jeg er redd for å gjøre feil	☐	☐	☐	☐
16. Jeg ønsker å gjøre det bedre i matematikk enn de andre studentene i klassen min	☐	☐	☐	☐
17. Jeg blir ofte nervøs før jeg skal ha en matematikkprøve	☐	☐	☐	☐
18. Jeg liker best å jobbe med matematikkoppgaver som virkelig får meg til å tenke	☐	☐	☐	☐
19. En viktig grunn til at jeg jobber med matematikkfaget er at jeg ikke vil dumme meg ut	☐	☐	☐	☐
20. Jeg er bekymret for karakteren i min i matematikkfaget	☐	☐	☐	☐
21. For min del er dette faget bortkastet tid	☐	☐	☐	☐
22. En grunn til at jeg ikke deltar aktivt i matematikktimene er at jeg vil unngå å fremstå som mindre smart	☐	☐	☐	☐
23. Jeg føler meg vellykket hvis jeg gjør det bedre enn de fleste andre studentene	☐	☐	☐	☐
24. Jeg gruer meg ofte til prøvene i matematikk	☐	☐	☐	☐
25. En viktig grunn til at jeg jobber i matematikktimene er at jeg ønsker å bli bedre i faget	☐	☐	☐	☐
26. Et av mine viktigste mål i dette faget er å unngå å se ut som om jeg har vanskeligheter med å løse matematikkoppgavene	☐	☐	☐	☐
27. Jeg er fornøyd med resultatene mine så langt i faget	☐	☐	☐	☐
28. Jeg synes pensumet vi har i matematikkfaget er nyttig å lære	☐	☐	☐	☐
29. Hvis jeg gjør det dårlig på en matematikkprøve, foretrekker jeg at ingen får høre om det, og jeg prøver derfor å skjule det	☐	☐	☐	☐
30. Det er viktig for meg å gjøre det bedre i matematikk enn de andre studentene i klassen min	☐	☐	☐	☐
31. Jeg føler meg ofte bekymret før jeg skal ha en prøve i matematikk	☐	☐	☐	☐
32. En viktig grunn til at jeg jobber med matematikkfaget er at jeg synes det er artig	☐	☐	☐	☐
33. Jeg gjør det dårlig i dette faget	☐	☐	☐	☐
34. Jeg ønsker å studere mer matematikk etter dette faget	☐	☐	☐	☐
35. Jeg ønsker å vise læreren min at jeg er bedre i matematikk enn de andre studentene i klassen min	☐	☐	☐	☐
36. Generelt foretrekker jeg å jobbe med oppgaver som jeg vet at jeg klarer, i stedet for å prøve å løse oppgaver som kan være for vanskelige for meg	☐	☐	☐	☐
37. Jeg jobber med dette faget for å unngå at de andre i klassen skal tro jeg er mindre smart	☐	☐	☐	☐
38. Jeg synes det vi lærer i matematikkfaget er kjedelig	☐	☐	☐	☐
39. Jeg trenger ofte hjelp i matematikktimene	☐	☐	☐	☐
40. Jeg jobber med dette faget fordi jeg er interessert i matematikk	☐	☐	☐	☐

Undersøkelsen gjennomføres med bistand fra SVT-IT, NTNU

Takk for at du ville svare på spørsmålene!

PART II: THE ARTICLES

ARTICLE 1

This article is awaiting publication and is not included in NTNU Open

ARTICLE 2

Studies in Educational Evaluation 66 (2020) 100904

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/locate/stueduc

Formative assessment as a collaborative act. Teachers` intention and students` experience: Two sides of the same coin, or?☆

Gabrielle Hansen

Department of Mathematical Science, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology. NTNU, Faculty of Information Technology and Electrical Engineering, 7491 Trondheim, Norway

ARTICLE INFO

Keywords:
Formative assessment
Higher education
Teacher evaluation and student evaluation

ABSTRACT

The study presented in this article, conducted in a European higher education institution, explores the value of dialogue as a means of facilitating alignment between the teacher's and student's understanding of a formative assessment practice. The overriding objective is to examine the degree of alignment between the teacher's intentions in using the "seven principles of good feedback practice" and the student's experiences of them in practice. The findings from this study differ from previous research findings on experience of assessment and feedback; while the latter broadly acknowledge a significant level of mismatch between student and teacher viewpoints, the findings from this study reveal a clear majority of common features between the parties' perceptions. The findings argue for the importance of the teachers' efforts to develop a mutual learning dialogue and the active effort and participation by both parties in such formative activities as self-assessment, reflection as feedback and dialogue.

1. Introduction

Formative assessment is fundamentally a *collaborative* act that takes place between the teaching staff and students where the primary purpose is to enhance the capability of the latter to the fullest extent possible (Yorke, 2003). The quality of this interaction is the very core of pedagogy (Black & William, 1998), and the key determinant for the successful outcomes of any educational changes (Sadler, 1998). This means that the effects of successful interventions exceed the addition of a few new routines to existing practice; the changes amount to a complete re-negotiation of the classroom contract that advances learning (Thompson & William, 2008).

In recent years, scholarly writing on formative assessment and feedback in higher education has flourished (O`Donovan, Rust, & Price, 2016; Orsmond, Maw, Park, Gomez, & Crook, 2013; Pitt & Norton, 2017). This has been encouraged in part by the desire to respond to the consistently low student satisfaction rating in several national student surveys looking into the overarching area of "assessment and feedback" in higher education (Boud & Molloy, 2013; Wiggen, Øygarden, Bakken, & Pedersen, 2019). Importantly, however, not only the students are dissatisfied with this area, so indeed are the teachers (Duncan, 2007; Nicol, 2010). A well-known scenario found in the research literature is that teachers find their feedback useful and blame the students for not using it, whereas the students complain about how useful the feedback they receive really is (Carless, 2006; Chanock, 2000; Duncan, 2007; Havnes, Smith, Dysthe, & Ludvigsen, 2012; Walker, 2009; Weaver, 2006). Even though validity can be claimed from both perspectives (Yorke, 2003), the outlines of formative assessment as a collaborative act appear to be weak or even more so, distinctly deficient.

In an attempt to bridge the gap between teacher and student understanding, a number of researchers have pointed out the importance of increasing student engagement and understanding through interactive *dialogues* and relational feedback, empowering students to evolve from being dependent on teacher-led feedback to being able to generate their own feedback on learning and progression, guiding students to understand the nature of quality and effectively evaluate and negotiate information, assuring that feedback is timely, and finally, considering the future role of grades (Carless et al., 2013; Lopez-Pastor & Sicilia-Camacho, 2017; Nicol, 2010; Nicol & Macfarlane-Dick, 2006; Orsmond et al., 2013; Orsmond & Merry, 2011; O`Donovan et al., 2016; Sadler, 1998). All in all, this has initiated a shift in focus on assessment and feedback, from providing detailed information to a student in which the focus is on information for fostering self-regulation (Boud & Molloy, 2013), to placing the development of student self-regulation at the core

☆ The author has a doctoral student position at the Norwegian University of Science and Technology (NTNU). This work was supported by a grant from NTNU. The founding source had no active involvement in the research process, such as the development of the study design; collection, analysis and interpretation of data; writing of the manuscript; and in the decision to submit the article for publication.

E-mail address: Gabrielle.Hansen@ntnu.no.

https://doi.org/10.1016/j.stueduc.2020.100904
Received 13 September 2019; Received in revised form 22 June 2020; Accepted 30 June 2020

of formative feedback practices (Carless, Salter, Yang, & Lam, 2011).

This shift in focus is also consistent with the assumptions underpinning the second Bologna decade up to 2020 (Pereira, Flores, & Niklasson, 2016), where student-centered learning and lifelong learning have been identified as a higher education priority area for the countries participating in the Bologna Process (Bucharest Communiqué, 2012; Leuven/Louvain-la-Neuve Communiqué, 2009; Yerevan Communiqué, 2015). This approach conveys the notion of students as constructivist learners and active participants with shared responsibilities for outcomes (European University Association (EUA, 2019), and encourages the use of learner-centred teaching and assessment practices that enable students to participate in their own learning and to develop critical thinking, autonomy, self-confidence and reflection (Hoidn, 2016).

Nevertheless, higher education, unfortunately, more often than not is facing a dilemma when it comes to feedback and assessment – namely that there is a huge gap between the empirical potential and the actual practice (O`Donovan et al., 2016), similar to the fact that the progression towards implementing the pedagogical concept of student-centred learning in European higher education has been rather slow (Hoidn, 2016).

Aiming to expand on the existing research literature on formative assessment and feedback, the study presented here, which has been conducted in a European higher education institution, uses one of the most influential accounts of feedback in higher education, namely Nicol and Macfarlane-Dick's model (2006) positing the *"seven principles of good feedback practice"* as a pedagogical framework to facilitate and create a dialogic assessment context in mathematics that is characterized by active student participation, self-assessment, reflection, peer dialogue and student-teacher dialogue. The overriding *objective* is to examine the degree of alignment between the teachers' *intentions* in using these principles and the students' *experiences* of them in practice. More specifically, this article seeks to answer the question: *to what extent are the teachers' beliefs about what they are doing in a dialogic formative assessment concordant with how the students experience it?*

2. Literature review and theoretical grounding

2.1. Perceptions of assessment and feedback in higher education

While there is increasing interest and research in the literature when it comes to formative assessment and feedback (Carless et al., 2017), there is less evidence of empirical research that directly examines experiences and perceptions of assessment (Mulliner & Tucker, 2017). The majority of these studies, however, has explored the attitudes to, beliefs about and understanding of *written* feedback (Carless, 2006; Duncan, 2007; Hounsell, McCune, Hounsell, & Litjens, 2008).

In general, research highlighting teachers` perspective suggest that teachers often believe that they are providing feedback to their students, and that their feedback is both detailed, fair, understandable, constructive and encouraging (Carless, 2006; MacLellan, 2001; Lid, Pedersen, & Damen, 2018; Mulliner & Tucker, 2017). Moreover, teachers often have an implicit assumption that the students know how to apply their feedback, and thus expect them to read and act on the feedback in some way and in doing so achieve better understanding (Orsmond & Merry, 2011; Johnson & Molloy, 2018). Research highlighting students' perspectives, on the other hand, suggests that they do not always understand the comments they receive (Weaver, 2006), and they do not necessarily interpret the teachers' comments in the way their teachers intended them to (Chanock, 2000; Orsmond & Merry, 2011; Walker, 2009). Moreover, students seldom have a homogenous view of what effective feedback is and how it could be used (Poulos & Mahony, 2008). They use feedback in different ways, for example to enhance motivation, enhance learning, encourage reflection, and clarify their progress (Orsmond, Merry, & Reiling, 2005). According to Orsmond et al. (2005), this implies that students are using feedback actively to make sense of their work. This is quite contrary to the perception of some teachers that students do not care about feedback information and are only interested in their grades (Duncan, 2007; Mulliner & Tucker, 2017). Recent research also shows that students who are assessed by methods which require their active involvement (e.g. learner-centered methods) view assessment as a fairer and more effective process than students who are assessed according to more traditional methods, such as examinations and written tests (Flores, Simao, Barros, & Pereira, 2015). It has also been suggested that students' emotional reactions play a significant role in determining how they perceive and act on the feedback they receive (Pitt & Norton, 2017).

A more limited number of studies have involved both parties in formative assessment and examined the level of alignment between student and teacher viewpoints. Although few in number, they broadly acknowledge the parties' lack of a shared understanding (Carless, 2006; Dunworth & Sánchez, 2016; Havnes et al., 2012; MacLellan, 2001; Mulliner & Tucker, 2017; Orsmond & Merry, 2011). The research literature has clear characteristics in that it focuses on the current formative assessment culture and highlights four important deficiencies: lack of alignment between the feedback provided by teachers and its utilization by students, lack of a shared focus on learning, lack of shared meaning between teacher(s) and students, and finally, lack of dialogue between teacher(s) and students.

2.2. Theoretical foundation: making sense is a social process

Similar to the work of Carless et al. (2011), this article uses a broad definition of feedback as "all dialogue to support learning in both formal and informal situations" (Askew & Lodge, 2000, p. 1). In essence, dialogic feedback can be either discussions/communication between teacher(s) and an individual or group of students, or between students (Orsmond et al., 2013). These types of dialogue are important because they allow students to make sense of new knowledge they encounter and help them to develop new conceptual understandings. More concrete dialogic feedback promotes interactive exchanges where interpretations are shared, meaning is negotiated and expectations are clarified, and it aims to provide opportunities for students to interact on notions of quality and standards in the discipline (Carless et al., 2011).

According to Yorke (2003), the essence of formative assessment and feedback has been described well by Wood (1987), as cited in Yorke, 2003), who discussed students' "maximum performance" in the light of Vygotsky (1978) "zone of proximal development" and posited the idea that teachers and students must collaborate actively to produce a best performance. Making sense is, thus, a social process (Bruner & Haste, 1987).

2.2.1. Dialogue as critically important for all sense-making

From a sociocultural and interactive perspective on learning, dialogue as a theoretical concept is described as critically important for achieving understanding and knowledge development (Bakhtin, 1981; Rommetveit, 1974). In other words, meaning and understanding, and thus learning, are created through interaction (Dysthe, 1996, 2008).

2.2.1.1. Internal and external. From a theoretical perspective, researchers operate with an expanded concept of dialogue, including both an external and internal dialogue where the former involves interlocutors who are present then and there, whilst the latter takes place within the individual and can be described as the interactions between the individual student and the subject content (Dysthe, 1996). According to Dysthe (1996), the internal dialogue is an essential part of the students' learning process, often referred to as the last phase of a three-part learning frequency, where the first is input (the teacher presents information), the second is processing (group work), whilst the third is "consolidating" knowledge (through reflection). The first phase is generally organized by the teacher and is often incorporated as an

active part of a learning activity. However, the third phase, the internal dialogue through reflection, is not as widespread in higher education (Boud, Ajjawi, Dawson, & Tai, 2018, Dysthe, 1996), which limits the students' ability to take an active role and involve higher order processes of self-regulation in their own learning (Carless, 2019).

2.2.1.2. Symmetry and asymmetry. Another important element in relation to dialogue, especially bearing the external dialogue in mind, is the degree of symmetry and asymmetry in dialogue situations (Dysthe, 1996). Basically, the teacher-student relationship is always asymmetrical, which is not necessarily an impediment to dialogue, rather quite the opposite. At the same time, the degree of asymmetry is often connected to different types of status, as between student and teacher. A key influencing factor in the dialogue between the teacher and students is therefore what Rommetveit (1991) calls "distribution of epistemic responsibility". One of the partners in a dialogue can thus have complete control over it by being the one to decide what is worth talking about and by mastering the "correct" language.

2.3. Seven principles of good feedback practice

Of key relevance to the notion of dialogue is the influential model created by Nicol and Macfarlane-Dick (2006) where self-regulated learning underpins seven principles of feedback. The aim of the model is to create a shift in focus so students are seen as having a proactive rather than passive role in generating and using feedback. Based on their analysis of extensive research material on formative feedback, they identified the following seven principles of good feedback practice:

1 Helps to clarify what good performance is;
2 Facilitates the development of self-assessment in learning;
3 Delivers high-quality information to students about their learning;
4 Encourages teacher and peer dialogue around learning;
5 Encourages positive motivational beliefs and self-esteem;
6 Provides opportunities to close the gap between current and desired performance;
7 Provides information to teachers that can be used to help shape teaching.

The seven feedback principles are not new: their value is that each principle is supported by substantial body of research, and that they are all defined in relation to their contribution to the development of learner self-regulation (see Nicol & Macfarlane-Dick, 2006, for further details and descriptions). The construct of self-regulation refers to the degree to which students can regulate aspects of their thinking, motivation and behavior during learning using a repertoire of mastery-oriented strategies (Pintrich & Zusho, 2002), such as the setting of and orientation towards leaning goals, assessing goal progress, seeking assistance, expending effort and persistence, and adjusting learning strategies (Zimmermann & Schunk, 2012).

Unfortunately, however, not every student displays these mastery-oriented qualities, and worse, many students abandon them just when they are most needed (Dweck, 1999, Dweck & Master, 2012). More concretely, the path to productive student decision making passes through their emotional reactions to assessment and what those emotions cause learners to do in response (Forsythe & Johnson, 2017; Stiggins, 2010).

2.3.1. Assessment and emotion: the notion of self-theories

A theory that can help shed a keener light on this issue is found in the extensive research of Carol Dweck and colleagues on students' theories of intelligence. Evolving within a social-cognitive framework, it is based on the idea that people develop beliefs that organize their world and give meaning to their experiences. Social cognitive theory emphasizes the importance of social influence on human behavior and assumes a triadic reciprocality between personal factors, behaviors, and environmental influences as they interact with and affect one another (Bandura, 1986; 2001, 1997).

In short, mastery-oriented qualities develop from the way students understand intelligence, and there are two very different ways through which students understand intelligence (Dweck & Legget, 1988). Some students believe that their intelligence is a fixed trait, they have a certain amount of it and that is pretty much it (entity view). On the other hand, some students believe that intelligence is malleable and can be changed (incremental view). These views are important because they affect how students respond to learning difficulties, external feedback, and their commitment to the self-regulation of learning (Dweck, 1999). Those with an entity view (fixed) interpret failure as a reflection of their low ability and are likely to give up, whereas those with an incremental view (malleable) interpret failure as a challenge to be overcome and increase their efforts (Dweck & Master, 2012; Dweck & Molden, 2005; Forsythe & Johnson, 2017). The validity of this theory has also been confirmed in higher education (Forsythe & Johnson, 2017; Yorke & Knight, 2004). Therefore, when designing an assessment practice, care should be taken to differentiate between what is "mastery" and "performance" oriented (Bryan & Clegg, 2019; Crooks, 1988). Unfortunately, research shows that much of the feedback students receive in higher education is performance oriented (Dweck & Master, 2012).

3. The context of the study

The research presented here was initiated by a cooperative relationship between two mathematics teachers and the author of this article, referred to in the following as "author". Together they formed a small team with the shared intention of establishing and applying a formative assessment practice in mathematics. "Author" had the role of guiding the teachers, both individually and as a team, and introduced the teachers to the "seven principles of good feedback practice". Through dialogue within the team, training in practice and joint reflection, they agreed on how the principles should be applied in practice (see section 3.3.1 and 3.3.2 for more detailed descriptions). "Author" also observed class assessments and had subsequent conversations with each of the teachers. The team also met regularly (at 14-day intervals) to discuss and evaluate each assessment session and learn from each other's experiences.

3.1. Preparatory engineering

The study in this article is based on data collected from two preparatory engineering courses held by "author" in 2019 (details removed for peer review) in Norway. The preparatory engineering course is for students who want to pursue an engineering degree but who do not have all the necessary requirements (e.g. mathematics, physics, social studies and languages) from upper-secondary school. There are approximately 50 students in each preparatory course, each lasting for a year, where graduates acquire the necessary curriculum requirements corresponding to the two last years of upper-secondary education.

Each preparatory course has its own mathematics teacher. The curriculum, instructions, number of teaching hours (ten hours per week) and assessment practice are the same in both courses. A special feature of the preparatory courses, bearing in mind the student group in mathematics, is that there is great variation both in background (they can have different vocational backgrounds and it also varies when it comes to when they were last in school; many come directly from upper secondary education, while others have been working for some time and are more distanced from their schooldays) and the academic level of the students.

3.2. Assessment requirements

The assessment practice in mathematics involves 14 formative

assessments and a summative assessment (final exam) over one school year. The 14 formative assessments consist of six mathematics tests (where students individually solve various mathematics tasks, either with or without a calculator), six multiple-choice tests and two mock exams. The mathematics and multiple-choice tests are given at approximately 14-day intervals, whilst the mock exams are held at the end of the fall and spring semesters, with the final exam at the end of the spring semester. This study includes the mathematics and multiple-choice tests, thus a total of 12 formative assessments.

The team wanted to promote and establish a strong focus on learning, and thereby reduce the focus on performance, errors, and correction, as recommended by research on self-theories. Therefore, grounded in the "seven principles of good feedback practice" (Nicol & Macfarlane-Dick, 2006), various learning objectives, such as self-assessment, making an effort, reflection, and dialogue, replaced the more traditional performance objectives. To gain approval in the various mathematics assessments, the students were measured against specific learning rather than performance goals. This meant that it was not their academic performance (number of correct answers on a test) that determined whether their assessment performance was approved, it was rather whether they had focused on and put an effort into reaching the *learning goals* of the assessments.

Thus, what distinguishes this assessment practice from the various assessment practices used in previous studies on assessment and feedback is simply what is being assessed: students' academic performance or students' attempts to learn, and their efforts to continue their learning. To achieve an overriding focus on mastery and learning, teacher-led grading was replaced by *self-assessment* through *reflection.* However, it is important to clarify that this did not involve the students writing "correct" or "incorrect" on various tasks, but rather *reflecting* on their mathematical solutions, choice of strategies, level of effort, learning outcomes and further progress. Their reflections thus served as a *criterion* for approval of the assessments and was submitted to the teacher after the assessments. Of the 14 formative assessments, ten had to be approved, according to the criterion described above, before the students can sit for the final summative exam

3.3. Formative assessment inspired by the "seven principles of good feedback practice"

All the formative assessments included in this study were given in classrooms and lasted about two teaching periods (45 min each). During the first 45 min, the students worked individually (on about 6–10 questions), followed by a 10-minute break. After the break, the main part of the "formative aftermath" began, namely the *interactive dialogic review*. Below is a more detailed description of how the teachers implemented "the seven principles of good feedback practice" when working with six mathematics tests and six multiple-choice tests over one school year. The procedures in these two types of assessment have clear differences and will therefore be presented separately.

3.3.1. Assessment procedures: mathematics test

A mathematics test was initiated with a short teacher introduction (*principle 1*) where the following were presented to the students: the academic goals of the test (e.g. which part of the curriculum the students would be questioned on), followed by an emphasis on the main purpose of the test, namely to create an opportunity for students to learn from their misconceptions, to make an effort and experience mastery (*principle 5*). The teacher pointed out two important "learning tools" that the students were to use: self-assessment and reflection (*principle 2*). As mentioned above, the students' academic performances were "moderated" in the sense that it was not the number of correct answers on a test that the teacher endorsed as important, but rather how the students dealt with their performance during the interactive dialogic review (*principle 3*).

After the introduction, the students started working on the test questions individually. When the test was finished, they were given a short break before they returned to the classroom and a joint plenary interactive dialogic review began.

During the interactive dialogic review all the questions and their solutions were presented by the teacher and discussed with the entire class. The students were repeatedly encouraged to participate and thus be in dialogue with the teacher (*principle 4*). The students' main task in this review was to evaluate and assess their own performance (*principle 2*). As part of this self-assessment, they were regularly given time to individually *reflect* on their own achievements (*principle 2*), and were encouraged to consider the following questions during their reflection: how did you solve this task; which method did you use; are there other methods or strategies you could have used; how can you improve; where should you put further effort, and; how did you cope?

The students wrote down their reflections and submitted them to the teacher after the interactive dialogic review.

3.3.2. Assessment procedure: multiple-choice tests

As with the mathematics tests, the multiple-choice tests were also introduced by the teacher (*principle 1*). The introduction was fairly similar, typically giving a presentation of the academic goals and the main purpose of the test to create an opportunity for the students to learn from their misconceptions, make an effort and experience mastery (*principle 5*). The difference compared to the former test was the presence of learning tools. In addition to self-assessment and reflection, the multiple-choice tests also explored the potential for learning through peer dialogue (*principle 4*). Prior to the test, the students were therefore placed in smaller groups of approximately four members.

After the introduction, the students started working on the test questions individually. Towards the end of the test they received the various alternatives for the multiple-choice questions and responded by using mobile technology: the student response system (SRS). Using this technology, teachers gain immediate access to student performances (Nielsen, Hansen, & Stav, 2016), in other words they had an overview of the students' understanding (what questions they have answered more or less correctly) and used this as a compass during the interactive dialogic review (*principle 3*).

During this review all the questions and their subsequent solutions were presented by the teacher in class and discussed in a plenary session. As with the mathematics tests, the students evaluated and assessed their own performance (*principle 2*) and were repeatedly encouraged to participate in this process (*principle 4*). Moreover, the students also discussed a number of questions with each other in small groups (*principle 4*). They were advised to use the following tools and methodologies during the discussions: share with the rest of the group how you responded to the question and the reasons and arguments behind your response; try to get everyone in the group to participate; listen to your peers, evaluate all responses, and; try to reveal misunderstandings and learn from them together.

The discussions were often completed with a second SRS round, where the students anonymously answered a test question once more. In doing this, they had what we can call a "second chance", an opportunity to learn from their misconceptions in-group, and respond again (*principle 6*). At the end of the interactive dialogic review, each student wrote down a final individual reflection (*principle 2*) on their performance (e.g. learning outcome, misconceptions, further effort, and peer dialogue) and submitted it to the teacher.

4. Research methodology

This study, with its origins in a larger intervention study using a qualitative approach within an interactive action research framework, aims to enrich our understanding of experiences of formative assessment in higher education, and importantly, provides insight into the degree of alignment between teacher and student viewpoints.

4.1. Participants

To protect the identity of the teachers participating in the research, they will not be described in more detail here, except to mention that they are both Norwegian nationals and have many years' experience of teaching mathematics. The students, in turn, were part of a larger student group and can thus be described in more detail.

First, between 20 and 25 students volunteered as participants from each class, meaning that the sample comprised students from the preparatory courses of both teachers involved. The only selection criterion established was that the students had attended the formative assessments in mathematics. The starting point for the selection of student participants was thus convenience, also called strategic selection (Johannessen, Tufte, & Kristoffersen, 2004).

Furthermore, the mean age of the students was 24 years within a range spanning from 19 to 33. They were evenly represented in terms of gender, and all were Norwegian nationals. Moreover, they had a great variety in backgrounds (including for example theatre, photography, literature and sociology; some of them had previous education, such as electricians and carpenters, while others came directly from high school) and a common goal of becoming an engineer (within various disciplines).

4.2. Research instruments

To gain insight into the teachers' and students' viewpoints of the formative assessment practice in mathematics, a qualitative methodology, with a phenomenological approach to research, was used. The aim of this research is thus to understand the research participants' perspectives (Kvale & Brinkmann, 2009; Postholm, 2005). To obtain knowledge about the participants' own experiences, intentions and opinions, in-depth interviews were found to be a suitable method. This is also the most commonly used data collection strategy within phenomenological research (Postholm, 2005).

4.2.1. Teacher viewpoints: semi-structured interviews

To elicit the teacher's intentions behind their various choices and their experiences of them in practice, semi-structured interviews were carried out regularly throughout the school year. As this study is aimed at examining the degree of alignment between the teachers' intentions in using "the seven principles of good feedback practice" and the students' experiences of them in practice, the teachers' experiences of their own change process will be reported elsewhere. The interviews thus afforded "author" the opportunity to seek clarification in relation to the teachers' intentions and experiences, and to discuss how these are interrelated.

4.2.2. Student viewpoints: focus-group interviews

To illuminate the students' perceptions of the formative assessment practice, focus-group interviews were conducted. This approach was chosen primarily because this is an independent method for acquiring the views of participants on various topics (Johannessen et al., 2004). Furthermore, it is a type of group interview where conversation and discussion processes are key elements. One of the advantages of focus-group interviews is that when handled properly, they can become extremely dynamic (Berg, 2007).

4.3. Data collection procedures

The data collection process can be summarized as follows:

- Semi-structured interviews with the teachers were carried out regularly throughout the school year. All in all, the data collection process for the teacher viewpoints amounts to 12 semi-structured interviews with two mathematics teachers (twenty-four interviews in total), and twelve classroom observations (twenty-four observations in total).
- Four semi-structured focus-group interviews, with four students in each group, were conducted with the students at the end of the spring semester.

All qualitative data material was digitally recorded, subsequently transcribed verbatim, and both the teachers and the students were given pseudonyms to protect their identity. More detailed descriptions follow below.

The semi-structured interviews with the teachers were completed shortly after the various formative assessments and served as interactive reflective conversations in which the teachers described and reflected on their assessment practice and responded to various probing questions. Each interview lasted about one, to one and a half hours. "Author" used a semi-structured question guide related to the teachers' implementation and application of the "seven principles". For each interview, the teachers brought their reflection notes which they wrote immediately after the formative assessments, and "author" brought the classroom observation notes. These were used as a tool to facilitate reflection on the various interview questions. More concretely, the semi-structured interviews focused on the teachers' preparations ahead of the formative assessments, their main objectives, their various choices relating to the implementation of particular principles, and their experiences relating to their practice, their students, and lastly, their plans for the next formative assessment. In sum, the interviews served as an in-depth summary and evaluation of the completed formative assessment, and furthermore, as a planning session for the next formative assessment.

Regarding the focus-group interviews with the students, there were, as stated above, between 20 and 25 students who volunteered as participants from each class. These were randomly distributed into four groups, with four students from each class. Both the classes were thus represented with two groups each. The students were informed prior to the interview that their participation was entirely voluntary, and they could withdraw at any time, or if they wished, at a later date they could have their data withdrawn from the study.

The facilitator of the focus-group interviews, "author", used a semi-structured question guide. It is important to point out that the students were not aware that the teachers were using the "seven principles of good feedback practice" as pedagogical inspiration for their own practice. The guide was therefore not related to these seven principles and the students were instead asked open questions about how the assessments in mathematics were undertaken and how they experienced them. The aim was that what the students highlighted as interesting and important would be discussed in the focus groups.

4.4. Data analysis procedures

All the qualitative data material in this study was analyzed by means of the constant comparative method, developed by Strauss and Corbin (1998), that involves structuring and seeing patterns in the data material to collect information that belongs together. A key characteristic is that the categories emerge from data itself. The data material from the different samples involved was analyzed separately to better explore and highlight the parties' foremost intentions and experiences in relation to the formative assessment practice in mathematics. This is in accordance with an approach suggested by Kathy Charmaz (2001, 2003), and included coding line by line, focused coding and categorization. After both analyses, the main categories, and the experiences that underpinned them, were compared to explore the understanding about the correspondence between the teachers' and students' experiences.

It is important to point out that the analytical steps were used as a *method of analysis* to be thorough and develop a rich description of the experience of the participants. More precisely: grounded theory was used as a method, as opposed to grounded theory as a product of the

method. The analysis has thus *not* aimed for concurrent involvement, theoretical sampling, or theoretical saturation.

5. Findings

The data relating to the teacher viewpoint revealed one main category "a learning arena", with three sub-categories, "reflection", "dialogue" and "introduction", whereas the data relating to the student viewpoint revealed the categories "reflection" and "social relations" (between students and student and teacher).

Initially, it is important to emphasize the point of departure for the parties involved in this study; namely a strong sense of meaningless previous assessment experiences. Moreover, both parties point to the focus on learning as the main distinction between past and present experiences. For the teachers, creating a *learning arena* (Teacher category 1) was the very essence and the foremost intention of their assessment practice in mathematics. The decision to implement self-assessment through *reflection* (teacher sub-category 1) was a key part of this intention. Using reflection, the teachers wanted the students to increase their awareness of their own learning by putting their own understanding, challenges and further solutions and efforts into words. In addition to considering reflection as an important learning tool for the students, the teachers also felt that it was an important tool for developing a relation to their students as their reflections gave them insight into the thoughts and points of view of the students that they would otherwise not have had no access to. One of the teachers, Michael, explains the value of student reflections as follows:

> "You know, I get so much more quality feedback as a teacher. I feel that the reflections help me see everybody better, I know them better. Their reflections have been successful beyond all expectations. It's highly meaningful for me as a teacher."

For the students, and in accordance with the teachers' intentions, self-assessment, and *reflection* (student category 1) represented a genuine opportunity to think actively and become more aware within their own learning process. According to the students, the reflection process functioned as important feedback on their own learning. By finding words for their own thoughts, solutions, and possible misunderstandings, they found that they were more aware of their own choices of methods and the underlying causes. The students experienced strong links between words and learning and felt that the reflections gave them insight into their own understanding and further effort. This is how one of the student groups explains it:

> Ola: "You know, you were giving yourself feedback."
> Megan: "Yes, and because of that, I feel that I become more aware of my own situation in this subject. I learn what I am good and not good at, and what I have to do before the next time. Actually, raising the awareness of my own skills, really."

The students point out that they were somewhat uncertain and skeptical at the start of the semester when reflection was presented as the main criterion for the assessments in mathematics. They were used to grades and the need to score a certain percentage on a test set correctly to achieve a particular grade. Thus, having the focus on effort through reflection represented something entirely new, but they were convinced when they found the teachers to be both engaged and genuinely involved, which created a *relation* of trust (student category 2). Another important requirement for the students' reflections may be summed up by the keyword "immediacy"; only ten minutes passed from when the test was finished before the review with reflections started. For the students this was simply the alpha and omega of the entire experience of reflection. The reason is very simple. The students find it very difficult, if not impossible, to reflect back in time in a good and constructive way. The experiences the students brought with them were of a relatively traditional assessment, where the teacher collects, corrects and grades, and return the test after some weeks. The challenge with this practice is that the attention and engagement of the students in terms of their own performance is lost while the teacher spends time correcting.

Beyond encouraging reflection, the teachers also aimed to create a common learning arena where students, and students and teachers could be in *dialogue* with each other and work together (teacher sub-category 2). Both the teachers and students felt that the group discussions represented a very clear switching of roles in the classroom, where the students assumed the role as the most active party. The teachers introduced several measures to initiate dialogues relating to the subject and shared learning in the groups, including changing the positions of the desks to create group settings, presenting specific discussions tools and hints, and using response technology to round off and summarize the discussions. In accordance with the teachers' intentions, the students highlight all these ideas as important for both starting up and rounding off the discussions, describing *peer dialogue* (student category 2) as a golden opportunity to learn from each other.

But the parties were not quite so harmonious about the dialogues in plenary sessions. While the teachers pointed to the reduction in physical distance, regular involvement and attempts to create a cooperative "we" environment as important for initiating dialogues with the students in plenary sessions, the students found it more natural and simpler to explain what they were thinking and experiencing to a co-student in the group discussions, finding it rather more difficult to express this to a teacher in a plenary session. In the groups they are all students, they have all experienced not understanding something, they may often be struggling with the same problems and they use concepts which everyone in the student group will understand, making it more natural to discuss more actively with other students than with the teacher in a plenary session. However, they pointed out that everybody participated in the plenary sessions when using response technology, where the student responses immediately appeared on the board, and together with the teacher they assessed whether there was a need for further review and explanation of the task. For the students, the use of the SRS became a form of quality assurance, where they received confirmation or rejection of the level of understanding in the group. If the student responses showed that the groups had misunderstood, the teacher would review the tasks with them again, which they felt strengthened the learning experience.

There was more harmony on the issue of a calm and stress-free assessment context. The teachers were aware that the focus on learning rather than achievement represented something new and different in an assessment context, for themselves as well as for the students. For this reason, they considered it their duty to convince and persuade the students about the value of a learning focus, as well as the value of the tools to be used when bringing this focus into practice. Time and effort, previously expended on the task of correcting student performance, were therefore now used to develop constructive *introductions* (teacher sub-category 3), where the expectations about their own role as teachers and the active role of their students were made clear, which they found had a calming effect on the students. Howard, one of the teachers, describes his intentions with the introduction in this way:

> "I want them (the students) to see me as … as sincere. I want them to believe that they can learn more, that they can learn from mistakes they make in a test. That this is where real learning takes place. That a test is only a beginning, not the end. I tell them every time, 'this is a learning arena'. If they see it like that, I have succeeded."

The students confirm this, highlighting the introductions as a particularly important part of the new learning focus that calmed them down. The students also point out the importance of actions and that the teachers' introductions were not only fancy words used to embellish the opening of an assessment session. On the contrary, they stated that the teachers' focus on learning permeated the entire interactive review. The teachers' introductions and efforts during the review were also seen as a measure of their credibility and engagement (student category 2).

When the teachers took the time to talk with the students at the start of each assessment and highlight the focus on learning, and also gave them tools to work with in the learning process, this was, according to the students, a clear sign that they had a genuine intention with the assessments, and that they were committed to their own intentions. According to the students, the teachers were there to help them, not to check them, as previous experience would have them believe. This motivated and bolstered their self-confidence. One of the students, Emma, explains:

> "I don't get this sense that if I do really bad on a test that I would be looked down on. It's more the learning which is part of the test. In the review and explanation I learn more about what we're working on and about myself. This applies to both what I master and don't master, and how I still have a chance to work with it and understand it. It's not like it's too late to learn. And this is an important message, you can learn even if you have misunderstood."

The students thus learned that a test could actually be a golden opportunity for learning, which can be said to satisfy the teachers' intention to create an arena for learning. The students were grateful to the teachers, and often felt that they should reward their teachers and make them proud. For the students, the best way of rewarding the teachers was to increase their effort and try to learn.

The students were, on the other hand, far less satisfied with what occurred after the assessments, and criticized their teachers for not using the results of the assessments more actively in the further teaching in mathematics. According to the students, the teachers had acquired much valuable information after each assessment. After a mathematics test the teacher would be able to read their reflections, and after the multiple-choice test, where the students answered the tasks using the SRS, the teachers had access to histograms that could give them an indication of what the students had understood and what they found difficult. The students did not feel that the information available to the teachers was applied to the subsequent teaching in mathematics. They criticised this, pointing out that they understand that the teachers have to work to a timetable, but also find that the information is too important to not be used more actively.

6. Discussion

This article has aimed to expand on existing literature on experiences of formative assessment and feedback. The overriding objective has been to examine the level of alignment between teachers' *intentions* when using Nicol and Macfarlane-Dick's model (2006), where self-regulated learning underpins seven principles of feedback practice, and the students' *experiences* in practice. Below, the findings from the study examined in this article and the level of alignment will be reviewed and discussed in relation to the potential value of dialogue using current empirical results and theory within the field.

6.1. Empirical implications: dialogue as a mediating factor

The findings from this study generally differ from previous research findings on the experience of assessment and feedback (Carless, 2006; Dunworth & Sánchez, 2016; Havnes et al., 2012; MacLellan, 2001; Mulliner & Tucker, 2017; Orsmond & Merry, 2011); while the latter broadly acknowledge a significant level of mismatch between student and teacher viewpoints, the findings from the study here reveal a clear majority of common features between the students' and teachers' experiences. Furthermore, the intentions and experiences presented in this article differ quite substantially from the usual norm within higher education. Through the implementation and use of a formative framework, this article emphasizes the students' role in making sense of their own learning and understanding, and in developing their self-evaluative abilities through reflection and dialogue, which may, to some degree, explain why the tendencies presented in the previous literature on experience of assessment and feedback are challenged.

For starters, this article chose a broad definition of feedback as "all dialogue that supports learning in both formal and informal situations" (Askew & Lodge, 2000, p. 1), highlighting that learning from feedback does not mean transferring knowledge from the teacher to the student, but rather constructing it in a process of social interaction (Dunworth & Sánchez, 2016). In other words, it is dialogic (Boud & Molloy, 2013; Nicol, 2010; Orsmond et al., 2013). This means that the immediate post-test period, which the teachers called the interactive dialogic review, embraced all forms of dialogue as feedback, including the dialogue between the teacher and the students, the peer dialogues and, last but not least, the students' reflections, their internal dialogue (Dysthe, 1996).

For the participants in this study, self-assessment through reflection was especially important. The teachers became better acquainted with their students by reading their reflections and developed a stronger relation with them. The students became better acquainted with themselves, their academic reasoning and resources, and their level of understanding, and also increased their level of awareness in their own learning process. In general, the students describe a process where they became more independent in their own learning. This confirms both theory on self-regulated learning, which sees self-assessment and reflection as effective methods students can use to acquaint themselves with self-regulating aspects of their own learning (Nicol & Macfarlane-Dick, 2006), thus becoming more aware of and active in constructing their own learning process, and not least, the theoretical value of reflection as an internal dialogue (Dysthe, 1996). According to Dysthe (1996), an internal dialogue between the student and the subject content is an essential part of each student's learning frequency. In order to sort out different perspectives, reflect over options and implications and connect a new understanding to what they already know, students can use their inner dialogue as a tool in their own development of knowledge. Without such abilities and awareness, planning for future learning simply becomes difficult (Boud et al., 2018; Orsmond & Merry, 2011).

In addition to seeing reflection and the inner dialogue as effective methods, the participants in this study were also very interested in, and not least influenced by, the external dialogue between the parties and the relation that was created between the teacher and student in the formative assessments. An interesting way to describe this relation is to call it a social classroom contract, a conditional contract between two parties, where both have a responsibility in relation to each other.

One way of understanding the teacher's intentions and actions towards creating a new assessment practice, with its new and underlying contract, is to relate them to one of the key assumptions in Bandura (1986,1997, 2001) social cognitive theory. That is the concept of reciprocal interactions, the view that (a) personal factors in the form of cognition, affect and biological events, (b) behavior, and (c) environmental influences create interactions that result in triadic reciprocality (Bandura, 1986). This behavioral-environmental link can be exemplified by the efforts of the teachers involved in this study. The teachers' main intention was to facilitate a holistic focus on learning that strengthens the students' desire to learn. To do this, they tried to foster learning experiences by improving the students' emotional states, self-belief and previous habits of thinking in relation to assessments in mathematics (personal) by enhancing their self-regulatory practices through, for example, self-assessment and reflection (behavior) and finally, by facilitating a social classroom structure that did not undermine students' learning experiences (environment), as more traditional and performance-oriented assessment structures that accentuate the importance of ability and social comparison might do.

As effective models and social persuaders, two key contributors within social cognitive learning theory (Pajares, 2012), the teachers used the introductions to point out that missteps are a natural and important part of a learning process that students can overcome through their own effort and persistence, and together with their peers and teacher. Competence and ability are thus both changeable and

controllable aspects of the environment. Finally, as a persuader, it was important for the teachers that their students perceived the message regarding a learning focus as genuine. The focus, therefore, could not just represent a few nice words at the beginning of each assessment. On the contrary, it had to permeate the entire practice with its roles and underlying social contract.

A sign that the teachers, at least to some extent, succeeded, is that the students emphasize the teachers' genuine interest and engagement as essential factors for their own dedicated participation. The students trusted the teachers' intentions and actions and explained their own level of trust in terms of the teachers' continuous efforts, for example through the introductions and use of various tools. This confirms earlier studies which point to trust as a particularly important dimension in an assessment practice (Boud & Molloy, 2013; Carless, 2009). The students were grateful to their teachers because they felt that they put them and their learning first, in contrast to what they describe as spending time on meaningless traditional assessment methods. Therefore, they wanted to reward their teachers and make them proud, they had a desire to master the work requirements and put more effort into the subject.

In sum, a new classroom contract, a complex network of shared understandings and agreed ways of working, was established in the assessment practice in mathematics. This was a contract where the teacher's role had changed from a focus on delivering the right answers to a focus on facilitating learning experiences, and where the student's role changed from passive to active participant, and finally, where the student-teacher relationship changed from being adversaries to being collaborators. This reinforces the main findings from previous and vital intervention studies on formative assessment (Thompson & William, 2008).

Importantly, however, a social contract does not simply exist as is; it must be continuously maintained through mutual understandings and expectations regarding its content. Without this, a breach of contract may occur. To try to exemplify a possible breach in this study, we can look at the students' "criticism of the teachers" for not using the results of the assessments more actively in their further teaching. Simply put, the students expected more, but the teachers did not deliver. This could indicate that the dialogue was not clear when it came to what the students could expect from further teaching.

The students' criticism is both interesting and indeed fair. For practical reasons, such as the time aspect and available resources, this study, similar to other research carried out in schools (Black, Harrison, Lee, Marshall, & William, 2003), uses the immediate post-test period as an opportunity for formative work. This means that the seventh and last principle, "provides information to teachers that can be used to help shape teaching", was not implemented to the same degree as the first six principles. Their critique is also quite thought-provoking given that the students had no advance knowledge on the education framework the teachers were using; they were not aware that one of the seven principles had not been fully implemented. It is thus very interesting that precisely this principle, number seven, is the basis for their criticism of the teachers. To put it another way, several of the intentions of the teachers, all with rationales and inspiration from six of the principles for good feedback practice represent the foremost experiences of the students in terms of the assessment practice in mathematics. The only principle the teachers did not implement fully turned out to be the focus of the strongest criticism from the students. This can be interpreted as showing that the relation between teacher and students was strong and mutual.

In addition to the criticisms described above, there was also less agreement about the dialogues in the plenary sessions. Whilst the teachers experienced an inclusive "we environment", where the students took an active part in smaller group and plenary discussions, the students felt somewhat differently about this, saying that participating in the group discussions was much more natural for them. In the groups, they were on the same level, talked a common language, in the sense of using terms and concepts between them that everyone understood, they had often struggled with similar problems and they all shared the experience of not understanding. Together these elements helped to facilitate active participation in the group discussions and, not least, make them constructive in terms of their own learning. The plenary sessions, on the other hand, also included an expert, namely the teacher.

One way of understanding this difference is to see the dialogue between the teacher and students as somewhat asymmetrical. Even though it is not difficult to argue that asymmetry is an advantageous dialogue situation (Dysthe, 1996), it is important to point out that this same asymmetry can very quickly become an effective impediment to a constructive dialogue because one of the parties, the students, experiences that the other party, the teacher, is the one who masters the right language (Rommetveit, 1991). This means that the teacher and students have different status and are on different levels in the dialogue, which creates an uneven "distribution of epistemic responsibility". In other words, the fact that we know different things in different fields can make a dialogue both necessary and useful but at the same time it can create a skewed distribution between the parties that challenges the participation of those with the "less correct" language.

The students' preference for dialogue in smaller groups does not mean that they rejected all participation in the plenary interactive dialogic review. They participated freely when they had the opportunity to use response technology. As opposed to the verbal participation in the plenary session, this form of participation was anonymous and might have been easier for the students to participate in, and not least, safer in terms of their self-esteem.

The emotional aspect of assessment has often been underrated in education, nevertheless, it is essential to the student's motivation (Clegg & Bryan, 2006). In short, being judged for the quality of your work is a potentially humiliating experience (Stiggins, 2010). Studies shows that students who ascribe to an entity view believe that assessment is an all-encompassing activity that defines them (Dweck, 1999). If they fail the task, they are failures (Forsythe & Johnson, 2017). Motivation and self-esteem are thus likely to be enhanced when a course fosters a notion of ability as incremental rather than fixed, and has many low-stakes assessment tasks with feedback geared to providing information about progress and achievement, rather than high-stakes summative assessment tasks where information is only about success or failure, or about how students compare to their peers (Bryan & Clegg, 2019; Dweck & Master, 2012; Nicol & Macfarlane-Dick, 2006).

In stark contrast to these last-mentioned practices, the teachers in this study focused on specific learning goals with the aim of improving and learning. This means that the traditional focus on achievement, which allows for social comparison, was toned down considerably, however, perhaps, not completely eliminated. For the students, the plenary dialogues could still have been characterized as an opportunity for social comparison. This means that some of the students may have been afraid of appearing less intelligent in comparison to the rest of the class. This may also explain why they state that they more than willingly participated in the plenary session when response technology was being used; anonymity removed the chance of being compared and thus "exposed".

Bearing this and future research in mind, it could be that by trying to promote a common interactive dialogue in the plenary sessions the teachers in this study may have facilitated more for a performance context than a mastering context. Considering the students' experiences and self-theories, perhaps the group discussions and response technology should have been the main focus, and student participation in the plenary dialogue should have been toned down. At the same time, it is important to point out that the students found the repeated invitations by the teachers to join a common plenary dialogue to be a characteristic of an engaged and interested teacher, which they felt bolstered their own involvement and not least, their interaction with the teachers.

7. Limitations

An important challenge when using the constant comparative method is that researchers can influence the findings from the analysis, especially in relation to the categories that have been developed. For this reason, it is often recommended that several researchers code the data material together (Charmaz, 2001). This was not done in this analysis. To increase the validity of the findings here, one or more researchers should have been included.

In addition, further research could benefit from including a more diverse student group, in the form of, for example, different nationals and cultures. Students seldom come to their learning environment with the same prior experiences, thus diversity, beyond the variation in the student group included in this study, is important to take in account when examining how various student groups respond to formative activities in practice.

8. Conclusion

This article shows that alignment between the views of teachers and students on formative assessment is relevant and feasible. It finds that students and teachers think in the same way about several important aspects of a formative framework. However, this alignment does not come out of the blue. It is the result of a focused dialogue and active efforts and participation by both parties. In the study here the dialogue was initiated by the teachers and maintained through continuous and deliberate effort, which in turn convinced and engaged the students in such formative activities as self-assessment, reflection and dialogue, whilst also being challenged by plenary exposure and asymmetry in the dialogue situation. In a more metaphorical language, it takes two to tango, confirming the fundamentals of formative assessment as a collaborative act. Nevertheless, in order to truly embrace assessment as a collaborative act, the emotional aspect of assessment, from the students' point of view, must be involved in every aspect of the educational change and the new contract development.

Acknowledgement

The author would like to acknowledge the teachers and students who volunteered to participate in this study.

References

Askew, S., & Lodge, C. (2000). Gifts, ping-pong and loops – Linking feedback and learning. In S. Askew (Ed.). *Feedback for learning* (pp. 1–17). London: Routledge.
Bakhtin, M. M. (1981). In C. Holquist (Ed.). *The dialogic imagination*. Austin TX: University of Texas Press.
Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1–26.
Berg, B. L. (2007). Chapter five: Focus group interviewing. In B. L. Berg (Ed.). *Qualitative research methods for the social sciences* (pp. 144–170). Pearson Education Inc.
Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education Principles Policy and Practice, 5*, 7–74.
Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for Learning. Putting it into practice*. Open University Press.
Boud, D., & Molloy, E. (2013). Changing conceptions of feedback. In D. Boud, & E. Molloy (Eds.). *Feedback in higher and professional education: understanding it and doing it well* (pp. 11–33). Oxon: Routledge.
Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (2018). *Developing evaluative judgement in higher education*. Routledge.
Bruner, J. S., & Haste, H. (1987). *Making sense: The child's construction of the world*. London: Methuen.
Bryan, C., & Clegg, K. (2019). *Innovative assessment in higher education* (2nd ed.). Routledge.
Bucharest Communiqué (2012). *Making the most of our potential: Consolidating the European higher education area*. http://www.ehea.info/media.ehea.info/file/2012_Bucharest/67/3/Bucharest_Communique_2012_610673.pdf.
Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31*(2), 219–233.
Carless, D. (2009). Trust, distrust and their impact on assessment reform. *Assessment and Evaluation in Higher Education, 34*(1), 79–89.
Carless, D. (2019). Feedback loops and the longer-term: Towards feedback spirals. *Assessment and Evaluation in Higher Education, 44*(5), 705–714.
Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education, 36*, 395–407.
Carless, D. (2013). Sustainable feedback and the development of student self-evaluative capacities. In S. Merry, M. Price, D. Carless, & M. Taras (Eds.). *Reconceptualising feedback in higher education: Developing dialogue with students* (pp. 113–122). Oxon: Routledge.
Carless, D. (2017). Scaling up assessment for learning: Progress and prospects. In D. Carless, S. Brigdes, C. K. W. Chan, & R. Glofcheski (Eds.). *Scaling up assessment for learning in higher education* (pp. 3–17). Singapore: Springer.
Chanock, K. (2000). Comments on essays: Do students understand what tutors write? *Teaching in Higher Education, 5*(1), 95–105.
Charmaz, K. (2001). Grounded theory. In J. A. Smith, R. Harre, & L. Langenhove (Eds.). *Rethinking methods in psychology* (pp. 27–49). London: Sage Publications.
Charmaz, K. (2003). Grounded theory. In J. A. Smith (Ed.). *Qualitative psychology: A practical guide to research methods* (pp. 81–110). London: Sage Publications.
Clegg, K., & Bryan, C. (2006). Reflections, rationales and realities. In C. Bryan, & K. Clegg (Eds.). *Innovative assessment in higher education* (pp. 216–227). Routledge.
Crooks, T. J. (1988). The impact of classroom assessment evaluation practices on students. *Review of Educational Research, 58*, 438–481.
Duncan, N. (2007). "Feed-forward": Improving students' use of tutors' comments. *Assessment and Evaluation in Higher Education, 32*(3), 271–283.
Dunworth, K., & Sánchez, H. S. (2016). Perceptions of quality in staff-student written feedback in higher education: A case study. *Teaching in Higher Education, 21*(5), 576–589.
Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.
Dweck, C. S., & Legget, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273.
Dweck, C. S., & Master, A. (2012). Self-theories motivate self-regulated learning. In D. H. Schunk, & B. J. Zimmermann (Eds.). *Motivation and self-regulated learning. Theory, research and application* (pp. 31–51). Routledge.
Dweck, C. S., & Molden, D. C. (2005). Self-theories. Their impact on competence motivation and acquisition. In A. Elliot, & C. S. Dweck (Eds.). *Handbook of competence and motivation* (pp. 122–140). New York/London: The Guilford Press.
Dysthe, O. (1996). *Ulike Perspektiv på Læring og Læringsforskning. [Different Perspectives on Learning and Learning Research]*. Cappelen Akademiske Forlag AS.
Dysthe, O. (2008). *Dialog, Samspel og Læring. [Dialogue, Interaction and Learning]*. Oslo: Abstrakt forlag.
European University Association (EUA) (2019). *Student-centred learning: Approaches to quality assurance (Anna gover, Tia Loukkola & Helene Peterbauer)*. https://eua.eu/resources/publications/884:the-quality-assurance-of-student-centred-learning-approaches-to-quality-assurance.html.
Flores, M. A., Simao, A. M. V., Barros, A., & Pereira, D. (2015). Perception of effectiveness, fairness and feedback of assessment methods: A study in higher education. *Studies in Higher Education, 40*(9), 1523–1534.
Forsythe, A., & Johnson, S. (2017). Thanks, but no-thanks for the feedback. *Assessment and Evaluation in Higher Education, 42*(6), 850–859.
Havnes, A., Smith, K., Dysthe, O., & Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation, 38*, 21–27.
Hoidn, S. (2016). The pedagogical concept of student-centred learning in the context of european higher education reforms. *European Scientific Journal, 12*, 1857–7881.
Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research & Development, 27*(1), 55–67.
Johannessen, A., Tufte, P. A., & Kristoffersen, L. (2004). *Introduction to social science methods*. Abstrakt forlag AS.
Kvale, S., & Brinkmann, S. (2009). *Det kvalitative forskningsintervju* (2th ed.). Oslo Gyldendal Akademiske.
Leuven/Louvain-la-Neuve Communiqué (2009). *The bologna process 2020 – The European higher education area in the new decade*. http://www.ond.vlaanderen.be/hogeronderwijs/bologna/conference/documents/leuven_louvain-la-neuve_communiqué_april_2009.pdf.
Lid, S. E., Pedersen, L. F., & Damen, M. L. (2018). *Teacher survey 2017: Main trends*. https://www.nokut.no/globalassets/studiebarometeret/underviserundersokelsen/lid_pedersen_damen_underviserundersokelsen2017_hovedtendenser_2-2018.pdf.
MacLellan, E. (2001). Assessment for learning: The differing perceptions of tutors and students. *Assessment and Evaluation in Higher Education, 26*(2), 307–318.
Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: Perceptions of students and academics. *Assessment and Evaluation in Higher Education, 42*(2), 266–288.
Nicol, D. J. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment and Evaluation in Higher Education, 35*, 501–517.
Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*, 199–218.
Nielsen, K. L., Hansen, G., & Stav, J. B. (2016). How the initial thinking period affects student argumentation during peer instruction: Students experiences versus observation. *Studies in Higher Education, 41*(1), 124–138.
O`Donovan, B., Rust, C., & Price, M. (2016). A scholarly approach to solving the feedback dilemma in practice. *Assessment and Evaluation in Higher Education, 41*(6), 938–949.
Orsmond, P., & Merry, S. (2011). Feedback alignment: Effective and ineffective links between tutors' and students' understanding of coursework feedback. *Assessment and Evaluation in Higher Education, 36*(2), 125–136.
Orsmond, P., Merry, S., & Reiling, K. (2005). Biology students' utilization of tutors'

formative feedback: A qualitative interview study. *Assessment and Evaluation in Higher Education, 30*(4), 369–386.
Orsmond, P., Maw, S. J., Park, J. R., Gomez, S., & Crook, A. C. (2013). Moving feedback forward: Theory to practice. *Assessment and Evaluation in Higher Education, 38*(2), 240–252.
Pajares, F. (2012). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk, & B. J. Zimmermann (Eds.). *Motivation and self-regulated learning. Theory, research and application* (pp. 111–139). Routledge.
Pereira, D., Flores, M. A., & Niklasson, L. (2016). Assessment revisited: A review of research in Assessment and Evaluation in higher Education. *Assessment and Evaluation in Higher Education, 41*, 1008–1032.
Pintrich, P. R., & Zusho, A. (2002). Student motivation and self-regulated learning in the college classroom. In J. C. Smart, & W. G. Tierney (Vol. Eds.), *Higher education: Handbook of theory and research: Vol. XVII*. New York: Agathon Press.
Pitt, E., & Norton, L. (2017). "Now that"s the feedback I want!' Students' reactions to feedback on graded work and what they do with it. *Assessment and Evaluation in Higher Education, 42*(4), 499–516.
Postholm, M. B. (2005). *Kvalitativ metode. En innføring med fokus på fenomenologi, etnografi og kasusstudier*. Oslo: Universitetsforlaget.
Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: The students' perspective. *Assessment and Evaluation in Higher Education, 33*(2), 143–154.
Rommetveit, R. (1974). *On message structure.* London/New York: Wiley.
Rommetveit, R. (1991). Dominance and asymmetries in a doll's house. In I. Markova, & K. Foppa (Eds.). *Asymmetries in dialogue*. Hemel Hempstead: Harvester/Wheatsheaf.
Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education, 5*, 77–84.
Stiggins, R. (2010). Essential formative assessment competencies for teachers and school leaders. In H. Andrade, & G. J. Cizek (Eds.). *Handbook of formative assessment* (pp. 233–250). Routledge.
Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Thousand Oaks, CA: Sage Publications, Inc.
Thompson, M., & William, D. (2008). `Tight but loose: A conceptual framework for scaling up school reforms. In E. C. Wylie (Ed.). *Tight but loose: scaling up teacher professional development in diverse contexts (Research report RR-08-29)* (pp. 1–44). Princeton, NJ: Educational Testing Service.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.
Walker, M. (2009). An investigation into written comments on assignments: do students find them usable? *Assessment and Evaluation in Higher Education, 31*(1), 67–78.
Weaver, M. (2006). Do students value feedback? Students' perception of tutors' written feedback. *Assessment and Evaluation in Higher Education, 31*(3), 379–394.
Wiggen, K. S., Øygarden, K. F., Bakken, P., & Pedersen, L. F. (2019). *Student survey 2019: Main trends (Studiebarometeret rapport 1/2019)*. https://www.nokut.no/globalassets/studiebarometeret/2019/studiebarometeret-2018_hovedtendenser_1-2019.pdf.
Wood, R. (1987). *Measurement and assessment in education and psychology: Collected papers 1967-1987.* London: Falmer.
Yerevan Communiqué (2015). *No title]*. http://www.ehea.info/media.ehea.info/file/2015_Yerevan/70/7/YerevanCommuniqueFinal_613707.pdf.
Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*, 477–501.
Yorke, M., & Knight, P. (2004). Self-theories: Some implications for teaching and learning in higher education. *Studies in Higher Education, 29*(1), 25–37.
Zimmermann, B. J., & Schunk, D. H. (2012). An essential dimension of self-regulated learning. In D. H. Schunk, & B. J. Zimmermann (Eds.). *Motivation and self-regulated learning. Theory, research and application* (pp. 1–30). Routledge.

ARTICLE 3

Studies in Educational Evaluation 56 (2018) 59–70

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/locate/stueduc

Formative assessment as a future step in maintaining the mastery-approach and performance-avoidance goal stability☆

Gabrielle Hansen*, Regine Ringdal[1]

Department of General Science, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, NTNU, Faculty of Information Technology and Electrical Engineering, Norway

ARTICLE INFO

Keywords:
Achievement goals
Stability
Change
Formative assessment
Feedback

ABSTRACT

The study presented in this article examines which achievement-goal pattern students pursue in a formative assessment practice that facilitates mastery and learning opportunities. An explanatory mixed-method design with three complementary data-analytic approaches (differential continuity, mean-level change and individual-level change) and four focus-group interviews were used to examine this topic. In five preparatory engineering courses "seven principles of good feedback practice" were implemented as an educational tool to facilitate and create a formative assessment practice aimed at promoting the significance of mastery and learning experiences. In contradiction to previous research, the findings in this study suggests an alternative achievement-goal pattern, and has to some extent succeeded in avoiding the proliferation of unfortunate motivational patterns found in earlier studies.

The findings of this study argues for the importance of teachers' efforts in relation to the development of students' achievement-goal patterns, and furthermore for maintaining achievement-goal stability.

1. Introduction

Performance feedback has been identified as an important variable in students' achievement-goal patterns within higher education. For instance, poor performance on an achievement task is associated with a decrease in mastery-approach goals and an increase in performance-avoidance goals (Senko & Harackiewicz, 2005). The discovery of this unfortunate pattern has given rise to an important question in the research literature: How can we maintain stability in students' pursuit of mastery-approach goals (Fryer & Elliot, 2007; Muis & Edwards, 2009)? Fryer and Elliot (2007) illustrate this:

> As educators, we would clearly like our students to endorse mastery-approach goals and steer clear of performance-avoidance goals. However, if students initially endorse mastery-approach goals, are these likely to remain stable of their own accord over time, or will substantial effort on the part of teachers and administrators be required to ensure that high levels of these goals are maintained? (p. 712)

The research literature on formative assessment describes feedback as an extremely important, if not critically important part of students' learning processes (Black & William, 1998; Hattie & Timperley, 2007). Specifically, with respect to higher education, the position generally taken has been that feedback is vital for the development of effective learning, in part because assessment procedures play a key role in shaping learning behaviour, and feedback can significantly accelerate that process (Sadler, 2010). Nonetheless, as seen in research on achievement-goal stability and change, feedback can also lead to less favourable patterns in that negative performance feedback can cause a decline in mastery-approach goals and an increase in performance-avoidance goals, a motivational pattern that undermines students' learning (Midgley, Middelton & Kaplan, 2001). As an educational tool, performance feedback should ideally promote students' wishes to resolve their misconceptions and increase their understanding. It should not inhibit students' desire to learn, which a decrease in mastery-approach goals and an increase in performance-avoidance goals may suggest. If the contribution of performance feedback is inhibition, it may be appropriate to reflect on the meaning of performance feedback, and more importantly, consider the *assessment practice* that underlies it. Feedback does not exist in a vacuum; it is part of a larger assessment

☆ The first author has a doctoral student position at the Norwegian University of Science and Technology (NTNU). This work was supported by a grant from NTNU. The funding source had no active involvement in the research process, such as the development of the study design; collection, analysis and interpretation of data; writing of the manuscript; and in the decision to submit the article for publication.

* Corresponding author at: Department of General Science, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology. NTNU, Faculty of Information Technology and Electrical Engineering, 7491 Trondheim, Norway
E-mail address: Gabrielle.Hansen@ntnu.no (G. Hansen).
[1] Present address: NTNU, Faulty of Health and Social Science, 7491 Trondheim, Norway.

https://doi.org/10.1016/j.stueduc.2017.11.005
Received 21 March 2017; Received in revised form 26 October 2017; Accepted 21 November 2017

context, a context which might be worth exploring. A relevant question to ask is whether educators are doing something wrong if feedback, which is intended to function as an educational tool, contributes to inhibition instead of strengthening the students' desire to learn.

The aim of this article is to expand on existing research literature on achievement-goal stability and change by studying students' achievement goals within a continuous formative and mastery-based assessment practice. To portray a broader picture of students' achievement-goal patterns, the study presented in this article uses theory of formative assessment as a framework to facilitate and create a mastery-based assessment practice in mathematics. The overriding *objective* is to examine which achievement-goal patterns students pursue within this context. Furthermore, the aim is to explore these observations with students' interviews and existing research literature in the fields of achievement-goal stability and change and formative assessment. The pedagogical framework for this objective is *"seven principles of good feedback practice"* by Nicol and Macfarlane-Dick (2006).

1.1. Formative assessment and the seven principles of good feedback practice

Assessment literature normally distinguishes between two types of assessment, *summative* and *formative*. The former is designed to rank, approve or control students' skills, and measures whether defined learning goals have been achieved (Sadler, 1998). Feedback information is provided after a particular type of work has been completed – normally given as a grade or some sort of achievement mark. In contrast, assessments can also generate feedback *during* a learning process, which enables students to improve their own learning and achievement. When an assessment serves these last-mentioned purposes, it is called *formative assessment* (Sadler, 1998). An increasing focus within the research literature on formative assessment is aimed at the students and how they can evolve from dependency on teacher-led feedback to being able to generate their own feedback on learning and progression and thus develop as independent learners who are able to monitor, evaluate and regulate their own learning (Cartney, 2010; Nicol & Macfarlane-Dick, 2006; Nicol, 2010). This means placing the development of student self-regulation at the core of feedback processes (Carless, Salter, Yang & Lam, 2011).

There is a substantial and growing body of evidence showing that feedback in the context of formative assessment has a strong impact on learning (Evans, 2013; Hattie & Timperley, 2007). Feedback has also become an increasingly important aspect of higher education learning and teaching strategies (Brown, 2010). However, even though there has been considerable development in research on feedback in recent years, there is surprisingly little awareness of what needs to be done to improve it and good ideas are often not translated into action (Boud & Molloy, 2013). In other words, although some principles of effective assessment feedback design have been established, the implementation of such designs has been demonstrably more problematic (Evans, 2013). This implies that current feedback practices within higher education are not fit for purpose (Carless et al., 2011; Evans, 2013) and in need of re-engineering (Carless, 2013). Furthermore, feedback is highlighted as one of the most problematic aspects of college student experiences (Blair, Wyburn-Powel, Goodwin & Shields, 2014; Carless et al., 2011).

According to Sadler (2010), the main challenge within higher education lies less with the quality of the feedback than with the assumption that telling, even detailed telling, is the most appropriate route to improvement in learning. In other words, the student role in feedback processes is in need of enhancement (Blair et al., 2014; Carless et al., 2011; Sadler, 2010). In an attempt to encourage more interaction with feedback, a number of academics have pointed to the need to engage students in interactive dialogues and thus reflect the reality of communication being a two-way process. (Black & McCormick, 2010; Blair et al., 2014; Carless et al., 2011; Carless, 2013; Donovan, Rust & Price, 2016; Hounsell et al., 2008; López-Pastor & Sicilia-Camacho, 2017; Nicol and Macfarlane-Dick, 2006; Sadler, 1998). The core of this argument is the need to step away from the "transmission" feedback model (Blair et al., 2014).

In 2006, Nicol and Macfarlane-Dick analysed extensive research material on formative assessment and feedback. Their aim was to create a shift in focus whereby students could be seen as having a proactive rather than a passive role in generating and using feedback. Based on the analysis, they identified the following seven principles of good feedback practice:

Good feedback practice:

1. helps clarify what good performance is;
2. facilitates the development of self-assessment in learning;
3. delivers high-quality information to students about their learning;
4. encourages teacher and peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape teaching.

These are familiar principles; their underlying value is supported by a substantial amount of research and they are all defined in terms of their contribution to the development of self-regulatory learning (Nicol, 2007).

1.2. Achievement goals

In recent decades, achievement-goal theory has emerged as an important theoretical perspective on students' motivation in school (Han, 2016; Kaplan & Maehr, 2007; Schunk, Pintrich & Meece, 2010). This theory is concerned with the purposes a learner adopts for achievement behaviour (Middelton, Kaplan & Midgley, 2004).

This article examines the trichotomous goal framework that includes the mastery-approach (students whose primary purpose of engaging in academic activities is to develop their competencies), the performance-approach (students who strive to appear competent and demonstrate high ability) and the performance-avoidance approach (students who strive to conceal their relative incompetence and avoid negative judgments). Similar to the distinction between the performance goals, a distinction has also been assigned to mastery goals, although the avoidance component of mastery goals still remains somewhat undefined theoretically and operationally (Tuominen-Soini, Salmela-Aro & Niemivirta, 2011).

Most research that has adopted the trichotomous goal framework has focused on various consequences of pursuing different achievement goals. A large body of research has compared the effects of these goals on important educational outcomes and each achievement goal has been associated with different patterns of cognition, affect and behaviour. These results have been summarised by others (e.g. Ames 1992; Elliot, 1999; Schunk et al., 2010) and will not be described here. Although achievement-goal research is a prominent approach to motivation, only a small number of studies have explored the issue of stability and change in students' achievement-goal endorsement over time (Fryer & Elliot, 2007; Han, 2016; Muis & Edwards. 2009; Senko & Harackiewicz, 2005; Tuominen-Soini et al., 2011).

A common underlying assumption in the research literature is that students' pursuit of achievement goals in a particular course remains relatively stable over time (Senko & Harackiewicz, 2005). One reason to anticipate this stability is that achievement goals represent concrete aims that emerge from personality characteristics, such as achievement motives and temperaments (Harackiewicz, Barron & Elliot, 1998). However, this is an assumption that recent research has begun to challenge, suggesting that although achievement goals may be stable, they can also be subject to substantial change (Fryer & Elliot, 2007;

Han, 2016; Muis & Edwards, 2009; Senko & Harackiewicz, 2005; Tuominen-Soini et al., 2011). It has also been suggested that students might regulate their pursuit of achievement goals based on instructional environments and/or perceptions of classroom goals that they confront (Fryer & Elliot, 2012). Research on this issue (Fryer & Elliot, 2007; Muis & Edwards, 2009; Senko & Harackiewicz, 2005) has suggested two possible types of goal changes, goal switching (e.g. switching from one goal type to another) and goal intensification (e.g. strengthening or reducing the pursuit of a goal). These studies support a multiple-goal perspective and the assertion that very often students do not actually pursue pure goals but rather multiple goals, and these can interact with one another (Han, 2016).

The claims that achievement goals may not be as stable as initially anticipated has led researchers to question fundamental aspects of achievement-goal patterns and processes, such as *when* these changes occur (Senko & Harackiewicz, 2005) and *what* causes them (Fryer & Elliot, 2007; Muis & Edwards, 2009).

1.3. Achievement goal stability and change

Although the issue of achievement-goal stability and change is clearly important, it has received little empirical attention in the achievement-goal literature (Fryer & Elliot, 2012). To date, few articles have been published that directly examine the issue of goal stability and change. In those that do, some researchers have focused on shifts in achievement goals across the transition from elementary to middle school (Anderman & Anderman, 1999; Anderman & Midgley, 1997; Madjar, Cohen & Shoval, 2017; Urdan & Midgley, 2003) and across the transition to upper secondary education and tertiary education (Tuominen-Soini et al., 2011), while others have examined shifts in students' achievement goals within an academic year (Bong, 2005; Warburton, 2017). Even fewer articles have examined the nature of stability and change in achievement goals across a sequence of various tasks (Han, 2016; Muis & Edwards, 2009). Senko and Harackiewicz (2005) and Fryer and Elliot (2007) examined changes over time across a sequence of similar tasks to control for the effect of varying tasks, whilst Muis and Edwards (2009) examined stability and change across different tasks. Common to these three last-mentioned studies is a self-regulation approach to describing achievement-goal stability and change: a suggestion that learners might adjust their achievement goals in ways similar to how individuals adjust goals within a self-regulation context, and an empirical exploration of factors that might affect the patterns and processes of achievement-goal endorsement within higher education. A review of these studies follows.

Senko and Harackiewicz (2005) examined the degree to which pursuit of achievement goals is regulated in response to ongoing competence feedback. Results from their research indicate that although achievement goals are generally stable throughout a semester, they are also responsive to competence feedback. Poor exam performance predicted a significant decrease in subsequent mastery-approach and performance-approach goals, and an increase in subsequent performance-avoidance goals. Senko and Harackiewicz (2005) identified performance feedback as a possible important variable in students' achievement-goal endorsement.

Fryer and Elliot (2007) examined the nature of stability and change in achievement-goal endorsement over time. Unique to their analytic approach, they examined changes at both the group level and individual level. Results from their study provide clear and consistent evidence of goal stability and change. Mastery-approach goals showed a sample-level decrease over time, whereas performance-avoidance goals showed a sample-level increase. The sample-level results provided further evidence of stability for performance-approach goals and mastery-avoidance goals. The individual-level change findings indicated that mastery-approach goals decreased for each time period, performance-avoidance goals increased, whilst there were nearly equal amounts of reliable increase and decrease in performance-approach goals and mastery-avoidance goals over time. The strongest shifts in achievement-goal change were between the first and second points in time. In addition, their results established fear of failure as an individual-difference variable.

Muis and Edwards (2009) empirically explored factors that might affect goal regulation and questioned the extent to which the type of task influences achievement-goal regulation. Findings from their study provided reliable evidence that students' level of achievement goals remains relatively stable but also changes over the course of a semester. For example, they found that performance-approach goals generally had the highest level of stability, whereas performance-avoidance goals and mastery goals had lower levels of stability. However, their study does not provide solid evidence that the type of task influences goal change. They rather suggest that there might be other factors, such as fear of failure, anxiety or interest, that are more predictive of goal change than merely the task itself.

1.4. Assessment practice and achievement-goal stability and change

Performance feedback is, as seen above, accorded great value in the research literature on achievement-goal stability and change. However, there are few elaborate descriptions of the assessment practice underpinning the use of feedback, leaving an inadequate overall understanding of the meaning of assessments in the process of achievement-goal endorsement. Furthermore, studies on achievement-goal stability and change often choose to maintain a continuous educational environment with the intention of *limiting* the influences of varying contexts and tasks (Fryer & Elliot, 2007; Senko & Harackiewicz, 2005), with the exception of Muis and Edwards (2009), who varied the types of tasks. What is not necessarily limited in these studies, however, is the possible influence this continuous educational environment might have had on students' achievement-goal patterns, including the possible influence of a continuous assessment practice, provided that one considers assessments to be a part of the environment. If such is the case, then assessments might affect students' achievement goals (Ames, 1992). An interesting question here is whether this is one of the reasons why the research literature reveals quite similar achievement-goal patterns, namely a decrease in mastery approach and an increase in performance avoidance.

2. The context of the study

The study presented in this article examines achievement-goal endorsement within the context of higher education. The majority of studies investigating longitudinal changes in students' achievement goal endorsement, focuses on younger students (e.g., elementary and middle school students), which makes it important to investigate students achievement goal pattern among older students as well (Tuominen-Soini et al., 2011). In addition, given that teaching becomes more focused on testing at higher grades (Daniel & Poth, 2017), higher education was considered an appropriate educational context for current study, which has developed a formative use of previous small summative tests.

This study has its origins in a cooperative project between one of the authors of this article, "author", and five mathematics teachers from a preparatory engineering course. Together they formed a collaborative team called "team mathematics". The purpose of this team was to establish and apply a formative assessment practice in mathematics. The pedagogical framework for this objective were the "seven principles of good feedback practice" by Nicol and Macfarlane-Dick (2006). This framework was chosen because it is directly aimed at developing independence in learning and involving students as active participants in their own learning and assessment process. It is important to clarify that this framework was used as an intervention tool for the teachers. Thus, it was not used as a tool for studying achievement-goal endorsement in that there were no hypotheses that the framework would lead to certain

changes in students' achievement-goal patterns. The team used these principles as the basis for creating a formative assessment practice that could facilitate students' learning opportunities.

"Author" had the role of educational counsellor and guided the teachers, both individually and as a team. "Author" introduced the teachers to the various principles, and through dialogue within the team, training in practice and joint reflection, they agreed on how the principles should be applied in practice (see Sections 2.3.1 and 2.3.2 for more detailed descriptions). "Author" was also present during and an observer of class assessments, and had subsequent individual conversations with each of the teachers. This was done to ensure that the practice that was conducted in the various classrooms was in line with the team's intentions and agreement about the implementation of these principles. The team also met regularly (at 14-day intervals) to discuss and evaluate each assessment session and learn from each other's experiences.

The fifth principle, "encourage positive motivation and self-esteem", served as one of the main principles for the team and resulted in the following formative assessment structure in mathematics:

- Numerous small "low-risk" assessments as opposed to a few larger performance-based assessments;
- A continuous emphasis on learning goals as opposed to performance goals;
- Self-assessment as opposed to teacher-led grading;
- Performance feedback as "approved/not approved" as opposed to a normative grading system.

2.1. Preparatory engineering course and class requirements

The study in this article is based on data collected from five preparatory engineering courses at the Norwegian University of Science and Technology (NTNU). A preparatory engineering course consists of students who want to pursue an engineering degree but who do not have all the necessary requirements (e.g. mathematics, physics, social science and languages) from upper-secondary school. There are approximately 50 students in each preparatory course, each lasting for a year, where graduates acquire the necessary curriculum requirements corresponding to the two last years of upper-secondary education.

Each preparatory course has its own mathematics teacher. The curriculum, instructions, number of teaching hours (ten hours per week) and assessment practice are the same in all five courses.

2.2. Assessment requirements

The assessment practice in mathematics involves 14 formative assessments during a school year: Six mathematics tests (where students individually solve various mathematics tasks, calculated either with or without a calculator), six multiple-choice tests and two mock exams before a final summative exam. The mathematics and multiple-choice tests are given at approximately 14-day intervals, whilst the mock exams are held at the end of the fall and spring semesters. Of the 14 formative assessments, 10 have to be approved before the students can sit for the final summative exam.

All formative assessments, with the exception of the mock exams (which are held in a location that can accommodate several hundred students, such as a sports hall), are given in classrooms and last about two teaching periods (45 min each). During the first 45 min, the students work individually (on about 6 to 10 questions), followed by a 10-min break. After the break the main part of the assessment begins, the *interactive dialogic review*. This is a joint plenary review of the test questions and their possible solutions, with an emphasis on immediate performance feedback, dialogue and active student participation.

2.3. Formative assessment following the "seven principles of good feedback practice"

To gain approval in the various mathematics assessments, the students were measured against specific learning, *not* performance goals. This means that it was not their academic performance (number of correct answers on a test) that determined whether their assessment performance was approved, it was rather whether they had focused on and put an effort into reaching the learning objectives of the assessment. Thus, what distinguishes this assessment practice from assessment practices used in other studies on achievement-goal stability and change is simply what is being assessed: students' academic performances or students' attempts to learn from their misconceptions.

2.3.1. Assessment procedure: mathematics test

A mathematics test was initiated with a short teacher introduction (*principle 1*). Here the following was presented to the students: the academic goals of the test (e.g. which part of the curriculum students would be questioned on), followed by an emphasis on the main purpose of the test, namely to create an opportunity for students to learn from their misconceptions, make an effort and experience mastery (*principle 5*). The teacher pointed out two important "learning tools" for the students to use: self-assessment and reflection (*principle 2*). Students' academic performances were "moderated" in the sense that it was not the number of correct answers on a test that the teacher endorsed as important, but rather how the students dealt with their performance during the interactive dialogic review (*principle 3*).

After the introduction, the students started working on the test questions individually. When the test was finished, they were given a short break before they returned to the classroom and a session of joint plenary interactive dialogic review began.

During the interactive dialogic review, all the questions and their solutions were presented by the teacher and discussed in plenum. The students were repeatedly encouraged to participate and thus be in dialogue with the teacher (*principle 4*). The students' main task in the interactive dialogic review was to evaluate and assess their own performance (*principle 2*). As part of this self-assessment, they were regularly given time to individually *reflect* on their own achievements (*principle 2*), and were encouraged to consider the following questions during reflection:

- How did you solve this task?
- Which method did you use?
- Are there other methods or strategies you could have used?
- How can you improve?
- Where should you put further effort?
- How did you cope?

The students wrote down their reflections and handed them in to the teacher after the interactive dialogic review. Their reflections served as a criterion for approval of the assessment. This criterion also applied to the multiple-choice tests. Below are a few examples of how the students reflected during the interactive dialogic review.

- "I see now that I'm struggling a bit with integration. I find it difficult to understand when to use e.g. substitution or integration by parts, and how to use these methods to solve the problem. I managed to determine both volume and area all right, except I forgot that area cannot be a negative quantity. I have no problems at all working with geometric and arithmetic series. That is good to know."
- "Now the time has come to connect the dots. Especially in probability theory, I tend to do things without thinking them through — I need to start identifying what I'm supposed to calculate, using a Venn diagram. I also need to develop a better understanding of what the problem is about, and what the result of various calculations should look like so I will know how to proceed. Next time I'll draw a

Venn diagram so I can see it for myself! It's very logical when you see it. Because then I know what to include and not. As for integration, which is a subject I know I'm struggling with, I didn't look through the various methods of integration, and didn't reflect on what I need to do. Consequently, my calculations were sloppy!"

2.3.2. Assessment procedure: multiple choice tests

As with the mathematics tests, the multiple-choice tests also started with a teacher introduction (*principle 1*). The content of the introduction was fairly similar, typically giving a presentation of academic goals and the test's main purpose. The difference from the former test was the learning tools. In addition to self-assessment and reflection, multiple-choice tests also explored the potential for learning through peer dialogue (*principle 4*).

After the introduction, the students started working on the test questions individually. They responded to the questions using a mobile technology called student response system (SRS), also known as electronic voting systems and classroom communication systems (Boyle & Nicol, 2003; Draper & Brown, 2004; Dufrense et al., 1996; Trees & Jackson, 2007). This technology enables teachers to gain immediate access to student performances. In other words, the teacher has an overview of the students' understanding (what questions they have answered more or less correctly), and can use this as a compass during the interactive dialogic review (*principle 3*).

During the interactive dialogic review, all the questions and their subsequent solutions were presented by the teacher in class and discussed in plenum. As with the mathematics tests, the students evaluated and assessed their own performance (*principle 2*), and were repeatedly encouraged to participate in this process (*principle 4*). Moreover, the students also discussed a number of questions with each other in small groups (*principle 4*). The students were advised to use the following tools and methodologies during the discussion:

- Share with the rest of the group how you responded to the question and the reasons and arguments behind your response
- Try to get everyone in the group to participate
- Listen to your peers
- Evaluate all responses
- Try to reveal misunderstandings and learn from them together

The discussions were often completed with a second round of SRS, where the students anonymously answered a test question once more. In doing this, the students had what we can call a "second chance", an opportunity to learn from their misconceptions in-group, and respond again (*principle 6*). At the end of the interactive dialogic review, each student wrote down a final individual reflection (*principle 2*) on their performance (e.g. learning outcome, misconceptions, further effort and peer dialogue), and submitted it to the teacher.

In sum, this study developed a formative use of small and previously summative tests. Similar to previous important research conducted in schools (Black, Harrison, Lee, Marshall & William, 2003), this study used the aftermath of tests as an opportunity for formative work.

3. Study design

The study presented in this article uses the sequential explanatory mixed-method design. This is a procedure for collecting, analysing and "mixing" or integrating both quantitative and qualitative data at some stage of the research process within a single study so a better understanding of the research problem can be gained (Creswell 2003). The sequential explanatory design also involves collecting and analysing quantitative and qualitative data in two consecutive phases within one study (Ivankova, Creswell & Stick, 2006). For this study, the research proceeded in the following manner: One questionnaire measuring students' achievement goals was given during two semesters (the fall semester of 2013 and spring semester of 2014), followed by focus-group interviews with the students. The research was guided by the following questions:

- Which achievement goal patterns do students pursue in a formative assessment practice in mathematics?
- How did the students perceive the formative assessment practice in mathematics?

The achievement-goal questionnaire was administered six times during two semesters (three times per semester) in the same student classes, but with a different number of students per semester.

The purpose of administering the questionnaire in both the fall and spring semester was to gain a more comprehensive understanding of students' achievement-goal patterns. Earlier studies promote the first achievement-goal measurements at the beginning of the fall semester as the most important ones in terms of change. The purpose of investigating the fall semester was thus to examine stability and change within a formative assessment practice during the beginning of a school year. However, although the first semester has previously been shown to exhibit the greatest achievement-goal changes, few studies investigate student's achievement-goal patterns during the spring semester. This study therefore includes both semesters in a school year. Moreover, a different type of assessment was included in the spring semester. A mock exam was included to involve an assessment situation that is closer to the final exam and the assessment structure used in previous studies on achievement goal stability and change. Therefore, in order to observe the students' achievement-goal patterns thoroughly, and through different assessment contexts, it was crucial to involve both semesters in the study.

In an effort to gain a broader impression of the assessment practice undertaken in this study, focus-group interviews were conducted with the students at the end of the school year. An important question here was whether the students perceived the assessment practice as having emphasis on a mastery-goal structure.

To measure achievement-goal stability and change, this article followed Fryer and Elliot's (2007) analytic approach and used three of the four complementary analytic procedures presented in their studies: differential continuity, mean-level change and individual-level change. These procedures were performed in SPSS, version 22.0. The intention behind these analytical procedures was to use the statistics to observe and describe students' achievement-goal patterns in a formative assessment practice. Furthermore, the aim was to explore these observations with students' interviews and existing research literature in the fields of achievement-goal stability and change and formative assessment. The analysis presented in this article is thus descriptive.

3.1. Analytical procedures

Differential continuity represents the level of rank-order consistency maintained in a construct over time within a sample and is measured by calculating Pearson product-moment correlations (Fryer & Elliot, 2007). High test-retest reliabilities for differential continuity provide evidence of goal stability (little change in intensity), whereas moderate-to-low-reliabilities signify moderate to high changes in goal intensity (Muis & Edwards, 2009).

Mean-level change represents the degree to which the average amount of a construct changes over time within a sample (Fryer & Elliot, 2007), and is measured with paired-sample *t*-tests. Significant mean changes in goals suggest changes in goal intensity; the greater the difference between scores, the more change in goal intensity (Muis & Edwards, 2009).

Individual-level change represents the magnitude of increase or decrease in a construct over time exhibited by an individual (Fryer & Elliot, 2007) and is analysed using a *reliable change index* (RCI) (Christensen & Mendoza, 1986; Jacobson & Truax, 1991). RCI allows for an assessment of whether an individual shows a significant increase,

decrease or no change in scores from one time to the next. RCI values lower than 1.96 or higher than 1.96 are unlikely to occur by chance and are thus considered indicative of reliable change. If change is random, the distribution of RCI values should be normal, with approximately 2.5% of values below −1.96, approximately 2.5% of values above 1.96, and approximately 95% of values between −1.96 and 1.96 (Fryer & Elliot, 2007).

3.2. Achievement goal measures

Mastery-approach goals, performance-approach and performance-avoidance goals were measured through the student version of PALS (Midgley et al., 1998), a 17-item achievement-goal questionnaire. In the scale, six items assessed mastery-approach goals, five items assessed performance-approach and six items assessed performance-avoidance goals. Participants responded to the items on a scale ranging from one (not at all true for me) to four (very true for me). Each sub-scale was estimated as the mean of the individual item scores.

The items were ranged from 1 to 4 following considerations based on a small pilot study conducted in a preparatory engineering mathematics class prior to the current study. In the pilot study, the scale ranged from 1 to 5, with the majority of the students placing their answer in the middle of the scale. Follow-up interviews with the students revealed that many of them did not read the questions, but automatically placed their answer in the middle to avoid taking a stand. The scale range was therefore converted to four response categories.

Furthermore, the old version of PALS (1998) was more suitable for a translation from English into Norwegian than the later version (PALS, 2000). After translating both versions, the 1998 version appeared clearer and more specific. For the purpose of this study, the wording of the items was specified for the students' thoughts about themselves in mathematics instead of more general statements. For example, the statement "I like classwork that I'll learn from it even if I make a lot of mistakes" was altered to "I like working with mathematics assignments that I learn from, even if I make a lot of mistakes". Moreover, the translation of the items from English into Norwegian also led to small adjustments.

3.3. Measurement procedure

The achievement-goal questionnaire was given three times (T1-T3) during the fall and spring semesters according to the following procedure: just before the students participated in an assessment, either a mathematics test, a multiple-choice test or a mock exam, they were asked to sign a consent form and respond to the attached questionnaire. After completing the questionnaire, the mathematics assessment began. Before the first data collection, the students were informed about giving written consent.

During the fall semester, the first achievement-goal questionnaire (T1) was given at the beginning of the semester, before the students had any assessments and received subsequent performance feedback. This was given before a mathematics test. The second achievement-goal questionnaire (T2) was given six weeks later, during midterm, before a mathematics test. The third achievement-goal questionnaire (T3) was given at the end of the semester, nine weeks after the second measurement and prior to a multiple-choice test.

During the spring semester, the first achievement-goal questionnaire (T1) was given at the beginning of the spring semester, just two weeks before a mathematics test. The second questionnaire (T2), however, was given eleven weeks later, during midterm, before a mock exam. The third questionnaire (T3) was given at the end of the semester, four weeks after the second measurement and prior to a mathematics test.

3.4. Focus-group interviews

Four semi-structured focus-group interviews, with four students in each group, were conducted at the end of the spring semester. These were further analysed using the *constant comparative method* (Strauss & Corbin, 1998), which involves structuring and searching for patterns in the data, and thus collecting information that belongs together into categories. The analysis also involves examining possible relationships between the various categories which the data material is linked to.

4. Results

4.1. Achievement-goal stability and change – fall

4.1.1. Participants

The results are based on data collected from five preparatory engineering courses at NTNU from the fall semester of 2013. The participants comprised 190 students (141 male and 49 female) from five mathematics courses. The mean age of the students was 22 years with a range from 18 to 42. Participation in the study was voluntary.

Only participants who were present during all three-time periods (T1, T2 and T3) during the fall semester were included in the statistical analysis, reducing the final sample size from 217 to 190 participants.

4.1.2. Differential continuity

Pearson product-moment correlation was used to examine differential continuity in the students' achievement-goal endorsement across three time periods. Table 1 presents correlations and Cronbach's alpha reliability coefficients for each achievement goal at three points in time. Correlations ranged from 0.69 to 0.78, which indicates a high level of stability for each of the three achievement goals.

4.1.3. Mean-level change

Principal component analysis (PCA) with an oblique rotation was used to ensure structural stability of the PALS for this sample, before the examination of mean-level change. The extraction of possible factors was conducted using *Kaiser's criterion* and *scree plot*. Three PCAs were performed to assess the goal measurements across the time periods.

Through all three points in time, three factors had an eigenvalue above Kaiser's criterion of 1, and explained a total of 60% (T1); 61% (T2) and 62% (T3) of the total variation in the variables. The first factor was assumed to represent *performance-avoidance goals*, the second factor addressed *mastery-approach goals* and the third factor addressed *performance-approach goals*. All factor loadings were above 0.4 and can thus be considered strong loadings (Tabachnick & Fidell, 2007). The analysis resulted in four items being removed from their original scale: two items assessing the mastery-approach goal, one item assessing the performance-approach goal and one item assessing the performance-avoidance goal. These were items that had high factor loadings on more than one factor, reducing the number of items measuring students' achievement goals from 17 to 13.

In sum, the principal component analysis resulted in three factor structures. The overall Cronbach's alpha value of the three factors is above the criterion of 0.7, indicating satisfactory internal consistency

Table 1
Study 1: Intercorrelations and Reliabilities.

	T1 to T2		T2 to T3		T1 to T3	
Goal type		α		α		α
MAP goals	0.72***	0.76	0.75***	0.74	0.69***	0.75
PAP goals	0.71***	0.79	0.78***	0.81	0.69***	0.83
PAV goals	0.75***	0.80	0.76***	0.81	0.71***	0.83

Note. T = time; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.
*** $p < 0.001$.

Table 2
Study 1: Descriptive statistics.

	T1		T2		T3	
Factor structures	*M*	*SD*	*M*	*SD*	*M*	*SD*
MAP	3.31	0.51	3.28	0.50	3.22	0.52
PAP	2.78	0.72	2.63	0.75	2.56	0.79
PAV	1.66	0.61	1.68	0.61	1.70	0.61

Note. T = time; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.

Table 3
Study 1: Mean-Level Change.

	T1 to T2		T2 to T3		T1 to T3	
Goal type	t	d	t	d	t	d
MAP	t(189) = 1.094	0.06	t(189) = 2.309	0.12	t(189) = 3.096***	0.18
PAP	t(189) = 3.585***	0.20	t(189) = 1.934	0.09	t(189) = 5.015***	0.29
PAV	t(185) = −0.647	0.00	t(185) = −0.623	−0.03	t(189) = −1.162	−0.07

Note. T = time; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.
*** p < 0.001.

(DeVellis, 2012). Table 2 presents a brief overview of the means and standard deviations of the three factors across the three time periods.

To examine the degree of mean-level change in the students' achievement-goal endorsement, paired *t*-tests were used. Bonferroni adjustments were made within each goal to control for Type I error rates across the three measurements for each sub-scale. Table 3 presents the t values and Cohen's d effect sizes from these analyses.

The mean-level change analysis revealed three significant changes in achievement goals: the mastery-approach goals *decreased* significantly between T1 and T3; the performance-approach goals showed a significant *decrease* between T1 and T2, and T1 and T3. The performance-avoidance goals did not change significantly between any of the points in time.

4.1.4. Individual-level change

To examine individual change, an RCI was calculated and was used to explore whether the students changed their level of achievement-goal endorsement between the three time periods. Table 4 presents the percentages of students who showed a reliable decrease, increase or no change.

The students generally displayed great stability in achievement-goal endorsement for each point in time. When aggregated at the group level, each achievement goal showed minor changes across all three time periods. The results revealed a continuous decrease in performance-approach goals, particularly between T1 and T3, confirming the mean-level change analysis. The decline in mastery-approach goals from the mean-level change analysis, however, was not confirmed. Performance-avoidance displayed the highest level of stability through all three time periods, with a stability level of 94.7% between T1 and T2, followed by 94.7% between T2 and T3 and 93.2% between T1 and T3.

4.2. Achievement-goal stability and change – spring

4.2.1. Participants

The results are based on data collected in five preparatory engineering courses at NTNU from the spring semester of 2014. The participants were 96 students (69 male and 27 female) from five mathematics courses. The mean age of the students was 23 years with a range from 19 to 43. Participation in the study was voluntary.

Only participants who were present during all three measurements (T1, T2 and T3) during the spring semester have been included in the analysis. Those excluded had only attended the first and second, second and third, or first and third measurements, reducing the final sample to from 206 to 96 participants.

4.2.2. Differential continuity

The Pearson product-moment correlation was used to examine differential continuity in the students' achievement-goal endorsement across three time periods. Table 5 presents correlations and Cronbach's alpha reliability coefficients for each achievement goal at the three points in time. Correlations ranged from 0.68 to 0.86, which consistently indicate a high level of stability for each of the three achievement goals.

4.2.3. Mean-level change

A principal component analysis was used to ensure structural stability before the examination of mean-level change. Three initial analyses were performed to reveal the eigenvalues of the potential factors.

Through all three time periods, three factors had an eigenvalue above Kaiser's criterion of 1, and explained a total of 67% (T1); 75% (T2) and 72% (T3) of the total variation in the variables. In addition to the eigenvalue criteria, a scree-plot was used to visualise the number of possible factors. Based on the cluster of variables that loads on the various factors, the first factor is assumed to represent *performance-avoidance goals*. Factor number two addresses *performance-approach goals*, while the third factor involves *mastery-approach goals*. The analysis resulted in five items being removed from the original scale: three items assessing mastery-approach goals, one item assessing performance-approach goals and one item assessing performance-avoidance goals. These were items that had high factor loadings on more than one factor, reducing the number of items measuring students' achievement goals from 17 to 12.

All factor loadings in all three analyses are above 0.4, and the overall alpha values of all four factors are far above 0.7. Table 6 provides an overview of the mean and standard deviation for the three factors across the three time periods.

Based on a total of three factor analyses, factorial invariance was confirmed. Stability was documented, Cronbach alpha values were high and changes in scores over time can therefore be interpreted as true change in the students' achievement-goal endorsement across the three time periods.

To examine the degree of mean-level change in the students'

Table 4
Study 1: Reliable Changes in Achievement Goal Endorsement.

Goal type	T1–T2			T2–T3			T1–T3		
	% inc	% same	% dec	% inc	% same	% dec	% inc	% same	% dec
MAP	2.1	97.4	0.5	5.8	90.5	3.7	1.6	96.3	2.1
PAP	1.6	93.2	5.3	2.6	91.1	6.3	1.1	91.6	7.4
PAV	2.6	94.7	2.6	2.6	94.7	2.6	3.2	93.2	3.7

Note. T = time; inc = increase; dec = decrease; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.

Table 5
Study 2: Intercorrelations and Reliabilities.

Goal type	T1 to T2	α	T2 to T3	α	T1 to T3	α
MAP goals	0.78***	0.79	0.75***	0.87	0.81***	0.86
PAP goals	0.86***	0.86	0.85***	0.88	0.85***	0.85
PAV goals	0.72***	0.82	0.68***	0.87	0.68***	0.87

Note. T = time; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.
*** $p < 0.001$.

Table 6
Study 2: Descriptive statistics.

	T1		T2		T3	
Factor structures	*M*	*SD*	*M*	*SD*	*M*	*SD*
MAP	3.09	0.66	3.09	0.72	3.17	0.68
PAP	2.55	0.82	2.59	0.82	2.56	0.81
PAV	1.66	0.60	1.79	0.68	1.73	0.65

Note. T = time; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.

Table 7
Study 2: Mean-Level Change.

Goal type	T1 to T2 t	d	T2 to T3 t	d	T1 to T3 t	d
MAP	t(93) = 0.74	0.00	t(93) = −1.656	0.11	t(93) = −1.831	0.11
PAP	t(95) = −0.980	0.05	t(95) = 0.801	0.04	t(95) = −0.168	−0.09
PAV	t(95) = −2.614**	−0.20	t(95) = 1.020	0.08	t(95) = 1.020	−0.12

Note. T = time; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.
** $p < 0.01$.

achievement-goal endorsement, paired *t*-tests were used. Bonferroni adjustments were made within each goal to control for Type I error rates across the three measurements for each sub-scale. Table 7 presents the t-values and Cohen's d effect sizes from these analyses.

Mastery-approach goals increased between T1 and T3, but the increase was not statistically significant. Performance-avoidance increased significantly between T1 and T2. Performance-approach goals increased between T1 and T2, but the increase was not statistically significant.

4.2.4. Individual-level change

To examine individual change, a reliable change index was calculated. RCIs were calculated to explore whether individual participants changed their level of achievement-goal endorsement between the three time periods. Table 8 presents the percentages of individuals who showed a reliable decrease, increase or no change. When aggregated at the group level, each achievement goal showed small changes across all three time periods.

The participants generally displayed high stability in achievement-goal endorsement for each time period. Performance avoidance showed the highest level of stability through all time periods, with a stability level of 95.8% from T1 and T2, followed by 94.8% stability from T2 and T3 and 95.8% between T1 and T3. Performance-approach goals also exhibit great stability across all time periods. The results display a minor increase in mastery-approach goals, particularly between T1 and T3, supporting the mean-level change analysis.

4.2.5. Focus-group interviews

Throughout the analysis process three codes distinguished themselves and were elevated into categories as they appeared to be the students' most important experiences in relation to the formative assessment practice in mathematics.

For the students, the assessment practice in mathematics was a new way of testing in the subject. According to them, a test usually involves the assignment of a grade, where they are accustomed to the following practice: During a test the students must perform, next the teacher collects their papers, the teacher then corrects their performances and grades them, and finally, after some weeks, their performances are handed back to them. The students are clear that they find the established practice more or less meaningless in terms of learning. Nonetheless, this is the practice they expect. The formative assessment practice therefore represented something new and different.

According to the students, three parts of the assessment process distinguished the assessment practice in mathematics from more traditional assessment practice: the teachers' *message* about their intentions (category 1), training in *reflection* (category 2) and the opportunity to *discuss* amongst themselves (category 3). One of the students (Peter) described the mathematics tests as follows:

> "It's not about testing in that sense, but in mathematics it's about the learning process, about reasoning why you did what you did."

According to the students, the teachers had a very clear message for the students; the assessments in mathematics should function as an *arena for learning*. This was the purpose of the various tests, and in the review, the students should have the opportunity to learn from their own misconceptions — and that was the intention of the whole test: Learning. Emily, one of the female students, explained it this way:

> "It creates some peace of mind, in a way. In the introduction, XXX (the teacher) focuses on the review, that this is where we should learn. We learn during this process. Not necessarily from taking the test, but rather reflecting on what we have done or by discussing it. That's the point, in other words."

The introduction the teachers gave before each assessment appeared to be a particularly important element in the assessment practice. First, the purpose of the tests was presented and highlighted; the tests were there for learning purposes. The students felt that this message had a calming effect on them, which allowed them to relax and try to focus on their own skills and understanding. Moreover, the introductions showed the students that the teachers were dedicated to the intentions. The fact that the teachers took the time to talk to the students at the

Table 8
Study 2: Reliable Changes in Achievement-Goal Endorsement.

Goal type	T1–T2			T2–T3			T1–T3		
	% inc	% same	% dec	% inc	% same	% dec	% inc	% same	% dec
MAP	4.3	91.5	4.3	2.1	96.8	1.1	6.4	91.5	2.1
PAP	2.1	94.8	3.1	2.1	94.8	3.1	3.1	92.7	4.2
PAV	3.1	95.8	1.0	2.1	94.8	3.1	2.1	95.8	2.1

Note. T = time; inc = increase; dec = decrease; MAP = mastery-approach; PAP = performance-approach; PAV = performance-avoidance.

start of each test and to emphasise the focus on learning repeatedly had to mean that they perceived the tests as meaningful and important. Lars described his perception of his teacher thus:

> "What he (the teacher) says during the introduction, it's not something he says for the fun of it. He's trying to reach us. He reaches me. He's serious. Even better, he has an intention with the test, and what he wants is that we should learn. This contrasts with all the teachers who just drone on, testing aimlessly. XXX (the teacher) dares to have a focus. And it's a meaningful focus. The other teachers have loads to learn from him."

The message about intentions given during the introduction became even more apparent for the students during the review of the different tests. The reviews highlighted the value and use of individual reflection (category 2) and discussion in small groups (category 3). The students interpreted this as "measures" being initiated so that the message from the introduction could be realised in practice, namely to create an arena for learning. In other words, category 1 (the teachers' message) served as a prerequisite for the experiences related to category 2 (training in reflection) and category 3 (peer discussion).

For the students, self-assessment through reflection (category 2) was perceived as a tool for *raising awareness*. The reflections required an active thought process on their part; they needed to think through what they had done correctly, less correctly and why what they had done proved to be more or less correct. Through such a process, the students felt that they gave *themselves* feedback. According to them, this helped them to gain a better overview of their own understanding, or just as important, lack of understanding. Sara explained it as follows:

> "Yes, that was when I understood that 'okay, this I understand' or "this I don't get". But I didn't see it before I wrote it down, before, I, well, before I was forced to really think about it."

Svein summed it up in this way:

> "I think, knowing that there's such a large proportion dedicated to it, that they value reflection as high as they do. It becomes a requirement for awareness about what one is doing."

While the reflections were emphasised as an important tool for creating awareness about their own skills, the group discussions (category 3) represented an opportunity to actually apply these skills in practice. By discussing various tasks together in small groups the students heard the points of view of others, while also promoting their own. For the students, the group discussions presented an opportunity to learn from each other. Mona perceived the discussions in this way:

> "Yes. It's really great. That's to say, if you made a mistake, then you get it explained. But if you've got it right, then you can explain it, and then you learn it … even better. It's a win-win."

5. Discussion

This article focuses on an important but largely overlooked issue in the achievement-goal literature, namely the issue of achievement-goal stability and change. More specifically, it questions which achievement-goal patterns students pursue in a formative assessment practice and how students experience an assessment practice in mathematics that facilitates learning.

5.1. Summary of the results from the fall semester

The results from the fall semester deviate from previous studies, especially when it comes to changes earlier studies have demonstrated in mastery-approach and performance-avoidance goals. Previous research has revealed a rather large decline in mastery-approach goals, particularly from T1 to T2, and an increase in performance-avoidance goals. However, this study finds a mastery-approach goal stability between T1 and T2, followed by a minor decrease between T1 and T3 and performance-avoidance stability. Performance-approach goals underwent the greatest changes in study 1, with a significant decrease from T1 to T2, followed by a subsequent decrease from T1 to T3. Thus, students became less performance-oriented throughout the semester.

The individual-level change analysis confirmed and expanded the patterns presented from the mean-level change analysis, and revealed an overall stability for all three achievement goals. Only minor increases and decreases were detected. Unlike the results from the mean-level change analysis, which showed a significant decrease in mastery-approach goals between T1 and T3, the individual-level change analysis presented a small increase in mastery-approach goals between T2 and T3. The mean-level change analysis presented no significant change during this period, which may have been caused by a similar number of increases and decreases cancelling each other out and giving the appearance of no group change in scores. Furthermore, the stability of performance-avoidance goals was confirmed and clarified through the individual-level change analysis. The majority of the students remained stable in their endorsement of performance-avoidance goals during all three time periods. Finally, results from the individual-level change analysis verified the decline in performance-approach goals, with the largest decline between T1 and T3.

5.2. Summary of the results from the spring semester

In comparison to the results from the fall semester, the results from the spring semester revealed great stability in the students' achievement-goal endorsement over time, whilst confirming minor changes. The results revealed one significant change at the group level during the spring semester: an increase in performance-avoidance goals from T1 to T2. In other words, performance-avoidance goals increased before students participated in a mock-exam assessment. According to the individual-level change analysis, performance-avoidance goals exhibited an overall high stability level.

Similar to the fall, results of the mean-level change analysis revealed stability in mastery-approach goals from T1 to T2 during the spring. The individual-level change analysis detected changes in mastery-approach goals between T1 and T2, however, the changes were similar in the number of increases and decreases and have thus cancelled each other out at the group level. The mean-level change analysis also indicates an increase in mastery-approach goals from T1 and T3, but the increase was not statistically significant. However, the increase was confirmed through the individual-level change analysis, which suggests that the students had become more mastery-oriented at the end of the school year.

Based on the mean-level change analysis, performance-approach goals appeared as the most stable achievement goal at the group level. Performance-approach goals increased between T1 and T2, but the increase was not statistically significant. At the individual level, the biggest changes in performance-approach goals occurred between T1 to T3, with a minor decrease.

5.3. Summary of the focus-group interviews

What were the students' perceptions of the formative-assessment practice conducted in this study? An interesting issue relating to the interviews was how similar the students' experiences were. They came from different preparatory courses, and many of them had different backgrounds (ranging from vocational backgrounds, military backgrounds, various work experience). These differences notwithstanding, the groups had very similar experiences of the formative assessment practice in their mathematics course, and very similar ideas about what they emphasised as important. This may be interpreted to mean that the involved teachers have been clear in their intentions, choices and actions in their interaction with the students.

The perception the students appeared to agree the most on was the

experience of the purpose of the assessments, i.e. learning. It had not gone unnoticed by the students that the teachers wanted the assessments to function as an arena for learning. This intention was clearly communicated to them at the beginning of each assessment: The assessments were to be an arena where they could address their own misconceptions, and through reflection and discussion experience mastery and learning. For the students, the purpose appeared to be very clear, and they perceived the assessment context as meaningful in terms of their own learning. They experienced increased awareness of their own skills through self-assessment and reflection, and felt that they increased their understanding of mathematics through discussions in small groups. Finally, the students perceived their mathematics teachers as critically important. The teachers presented the purpose of the assessments and facilitated and initiated measures for learning through reflection and discussion.

Bearing the students' experiences in mind, the assessment practice in mathematics may thus be described as learning oriented, not performance oriented.

5.4. Theoretical implications

The study presented in this article has to some extent challenged the initial tendencies evident in the research literature. Firstly, the findings in this study point to a more stable pursuit of mastery-approach goals than previous studies. Secondly, the performance-avoidance goals exhibited various levels of stability. An interesting focus looking ahead is whether the assessment practices that were implemented in the current study may have contributed to the results presented in this article.

Research has repeatedly postulated over the past 25 years that mastery approaches to instruction are more beneficial than performance approaches (Fryer & Elliot, 2012; Daniel & Poth, 2017), making it widely accepted, and extremely important, to encourage teachers and administrators to foster certain types of achievement goals over others (Maehr & Midgley, 1996). Recent research further suggests that the conceptualization of instruction from either a mastery or a performance perspective applies equally to assessment (Daniel & Poth, 2017). According to Ames (1992), the ways in which students are evaluated is one of the most salient classroom factors that can affect their motivation. This goes beyond the question of whether students are evaluated; it also concerns students' perceptions of the meaning of the evaluative information (Mac Iver, 1987). Hence, depending on how an assessment is structured, students may be oriented towards different goals that further stimulate different patterns of motivation. The study in this article is descriptive and therefore cannot claim that the use of a mastery-oriented assessment practice will serve as a causal reason for the disclosure of a somewhat different achievement-goal pattern than what has been presented in previous studies. There are, however, several empirical reasons for promoting a mastery-oriented assessment practice as a *possible* factor of influence.

The assessment practices provide one of the clearest distinctions between the study in this article and previous studies on achievement-goal stability and change. Most studies on this theme contain an assessment practice of some sort. However, little has been written about the actual assessment practices. Instead, most studies have focused on a common outcome of various assessments, namely the effect of performance feedback. While previous studies have been relatively vague in their descriptions of their chosen assessment practice, the study in this article set out to implement theory of formative assessment as a framework to facilitate and create a mastery-oriented assessment practice.

To ensure that the study in this article had a focus on mastery and learning, and thereby reduce the focus on performance, an entire assessment culture in mathematics had to be changed. Various learning objectives, such as self-assessment, making an effort, reflection and dialogue, all in line with the "seven principles of good feedback practice" (Nicol & Macfarlane-Dick, 2006), replaced more traditional performance objectives. Moreover, the assessments were not completed once performance feedback, such as a teacher-led grading process, had been given. To achieve an overriding focus on mastery and learning, teacher-led grading was replaced by self-assessment. The students evaluated themselves during the interactive dialogic review, which did not involve them writing correctly or incorrectly on various tasks, but rather *reflecting* on their mathematical solutions, choice of strategies, level of effort, learning outcomes and further progress. The students participated in a joint interactive dialogic review for an elaborate review of the test questions immediately after completing the test, focused on unravelling their misconceptions and creating further learning opportunities.

The teachers played an important role during the various assessments as messengers, role models and partners. Their role was clearly emphasised in the student interviews. The teachers' main task was to motivate and convince the students that the assessments were aimed at learning. This was highlighted during every assessment because students typically associate a test with performance, not learning. Through the teacher's efforts, the students had to learn that the goal of each assessment was not to grade their skills, or to socially compare or rank their achievements, but to gain insight into what they had done and how they could proceed. If the students were to be convinced, it was not enough to communicate this message verbally. Confirming previous studies on assessment and feedback (Blair et al., 2014; Carless, 2013; Sadler, 2010). From the students' perspective, the real value came when the teachers' message was exemplified through practice, such as self-assessment, reflection and peer dialogue, in other words, when the students experienced an interaction between words and actions. This interaction convinced them that the teachers had faith in the message they were imparting, which further created a practice based on mutual trust.

All in all, the teachers opened a door to an unknown assessment practice in mathematics for the students. This was a practice that did not reflect the transmission model the students knew inside and out. Instead, they were introduced to a practice aimed at enhancing their active learning and providing a meaningful dialogue. The students' experiences reflect the aim of the seven principles of good feedback practice (Nicol & Macfarlane-Dick, 2006) and support a growing tradition in the research literature on formative assessment and feedback promoting the importance of increasing student involvement in feedback practices through interactive dialogues (Blair et al., 2014; Carless, 2013; Donovan et al., 2016; López-Pastor & Sicilia-Camacho, 2017).

When using a more traditional assessment practice (following the transmission model) and teacher-led grading, students' efforts to learn from their own misconceptions, or their participation in discussions about their performances or their reflections on their efforts and learning strategies are rarely what ultimately matters or is rewarded, rather it is their performances *per se*. If performance feedback (e.g. a grade) is what students are ultimately left with after an assessment, how does a teacher persuade them to focus on learning or mastery? An achievement focus can easily overshadow a focus on mastery and learning. For this reason, the study in this article has tried to facilitate a holistic focus on learning and mastery, which may explain why to some extent it challenges the initial tendencies presented in the achievement-goal literature. The students' academic performance was never the subject of teacher-led grading. The overall focus was on the students and their effort to learn through reflection, self-assessment and peer dialogue.

The significance of contextual factors challenges Muis and Edwards' (2009) suggestion that antecedents to goal adoption (e.g. fear of failure) might be more essential to goal change than contextual factors. They endorsed the use of a mastery-oriented classroom goal structure and teacher-led grading in their studies, and their findings indicated a decrease in mastery-approach goals. This led them to suggest that classroom interventions are not as powerful over time if antecedents to goal adoption are not taken into consideration. This is a constructive suggestion and should be paid more attention to in future research.

However, it is important to emphasise that a future focus on antecedents must not overshadow, underestimate or come at the expense of a broader focus on actual actions as a source of influence on students' motivation and their desire to learn and evolve (e.g. teachers' actions, formulations and attitudes, and the responsibility of educational institutions). This article also indicates a decrease in mastery-approach goals. During the fall semester, mastery-approach goals were subject to a significant decline between T1 and T3. One way to interpret and further develop the findings presented here is to direct further attention towards other possible influences, such as antecedents like fear of failure. An equally important option, however, is to facilitate an even stronger focus on the use of mastery-oriented structures and examine contextual factors that can help to further maintain mastery-approach and performance-avoidance goal stability beyond the findings presented in this study.

Furthermore, in the fall semester, performance-avoidance goals remained low and stable throughout the entire study. A possible source of influence for this might have been the teachers and what they provided in terms of both words and actions. The assessments were transformed into learning arenas, thus giving few reasons for performance-avoidance goals to increase. As one of the girls from the interviews described the focus on learning: "It creates some peace of mind". The focus on learning, versus performance, relieved the students of stress and had a calming effect on them. They were used to a test, especially in a subject such as mathematics, being an arena for performance, followed by teacher-led grading and the awarding of a grade. According to the students, a focus on learning appeared as less threatening and more meaningful. This could also possibly explain why performance-approach goals showed a continuous decline during the fall. There were few advantages for students to pursue performance. Instead, tools such as reflection and peer dialogue were implemented as key assessment criteria. Performance-avoidance goal stability, however, was not verified throughout the spring semester. The students' pursuit of performance-avoidance goals increased prior to participation in the mock exam. An interesting question to consider here is whether this increase was due to an assessment that was more similar to the final summative exam, and to a structure that was more similar to the assessment structure used in previous studies on achievement goal stability and change.

In addition to a continuous focus on the development of mastery-oriented practices, future research within the achievement goal tradition might benefit from a greater awareness of what really *defines* and *facilitates* a mastery-oriented assessment practice. In other words, this article calls for a greater *awareness* of what is actually required in order to create a mastery-oriented practice, and puts greater emphasis on the role of teachers and their potential impact on students' achievement-goal patterns.

The role of teachers in classroom interventions is also emphasised in the achievement-goal literature. According to Ames (1992), teachers structure their classrooms and their own goals influence their beliefs and subsequent actions. Thus, changing a classroom structure may also require changing a teacher's goals, belief system and broader views about learning and motivation (Ames, 1992). Recently, Daniel and Poth (2017) examined the relationships between pre-service teachers' conceptions of assessment and their intended approaches to classroom instruction and assessment. Their findings indicate that pre-service teachers think similarly about instruction and assessment and link their intended approach to instruction to their intended approach to assessment by adopting either a mastery or a performance approach for both practices (Daniel & Poth, 2017). Furthermore, it has been clearly documented that student's perceptions of teachers, classrooms and schools can have an important influence on students' achievement-goal adoption (Fryer & Elliot, 2012). However, few studies have reported experimental intervention research that focuses on achievement goals using the trichotomous or 2 × 2 framework (Fryer & Elliot, 2012). While the study in this article has attempted to bridge this gap, it is not experimental but rather descriptive. It aims to provide a more comprehensive understanding of the subject in question through observing students' achievement-goal pattern within a certain context, and elaborate on the findings through the students' perceptions and previous research.

Knowledge about the students' different motivational patterns and the way these patterns remain stable or change over time is valuable for educational practice (Tuominen-Soini et al., 2011), and the main challenge for teachers when shaping their classroom and learning environments is to consider the differences in students' perceptions (Fryer & Elliot, 2012). Clearly, these are important areas for future research. Future research could benefit from complementing descriptive studies, such as the one presented in this article, with experimental intervention research that focuses on achievement goals using the trichtomous or 2 × 2 framework, as recommended by Fryer and Elliot (2012), and thus examine possible interactions between student's perception (of teachers, classrooms, assessment context etc.) and achievement-goal stability and change. This could be truly valuable for educational practices, highlighting the value and importance of the individual perceptions, which together will then facilitate a learning environment.

6. Limitations

The statements on mastery goals in the student version of PALS (Midgley et al., 1998) have been criticised. Some of the statements used to measure mastery goals contain affective content similar to a measurement of interest (Midgley et al., 2000).

The achievement-goal questionnaire used in this study only has four response categories. By increasing the number of alternatives from four to five, the analysis could have captured more changes in the students' achievement-goal patterns.

The achievement-goal questionnaire were collected before various types of assessment. The students may have experienced these assessments in different ways which could have in turn affected how they responded to the questionnaire. Muis and Edwards (2009) examined the extent to which the type of task influences achievement-goal stability and change. Results from their study did not provide solid evidence that the type of task influences goal regulation. Their results, however, cannot be automatically transferred to the study in this article.

The achievement-goal questionnaire used during the fall and spring semester are identical, which means that students have received and responded to the same questionnaire a total of six times, which may also have influenced the final results.

7. Conclusion

This study argues for the importance of teachers' efforts in relation to the development of students' achievement-goal patterns, and furthermore for maintaining achievement-goal stability. By using feedback as a learning tool in a formative assessment practice rather than as a measurement tool for performances, feedback can reinforce its position as a learning tool in higher education. Through incorporating assessment and feedback in a new way in the achievement-goal literature, a more comprehensive understanding of the value of feedback, formative assessment and students' achievement-goal patterns may be attained.

Acknowledgement

The authors would like to acknowledge the teachers who volunteered to participate in this study. Their effort and commitment inspired and motivated us throughout the research process. Thanks are also extended to Alex Strømme for his helpful comments that greatly the manuscript, and to Thomas Dahl who provided insight and expertise that greatly assisted the analytical procedures. Finally, we would like to thank the students who participated in the achievement-

goal measurements and focus-group interviews. Their experiences improved our understanding.

References

Ames, C. (1992). Classrooms: Goals, structures: And student motivation. *Journal of Educational Psychology, 84*, 261–271.

Anderman, L. H., & Anderman, E. M. (1999). Social predictors of changes in students' achievement goal orientations. *Contemporary Educational Psychology, 24*, 2–37.

Anderman, E. M., & Midgley, C. (1997). Changes in achievement goal orientations, perceived academic competence, and grades across the transition to middle-level schools. *Contemporary Educational Psychology, 22*, 269–298.

Black, P., & McCormick, R. (2010). Reflections and new directions. *Assessment & Evaluation in Higher Education, 35*, 493–499.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*, 7–74.

Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for Learning. Putting it into practice.* Open University Press.

Blair, A., Wyburn-Powel, A., Goodwin, M., & Shields, S. (2014). Can dialogue help to improve feedback on examinations? *Studies in Higher Education, 39*, 1039–1054.

Bong, M. (2005). Within-grade changes in Korean girls' motivation and perceptions of the learning environment across domains and achievement levels. *Journal of Educational Psychology, 97*, 656–672.

Boud, D., & Molloy, E. (2013). Changing conceptions of feedback. In D. Boud, & E. Molloy (Eds.). *Feedback in higher and professional education: Understanding it and doing it well* (pp. 11–33). Oxon: Routledge.

Boyle, J. T., & Nicol, D. J. (2003). Using classroom communication systems to support interaction and discussion in large class settings. *Association for Learning Technology Journal, 11*, 43–57.

Brown, S. (2010). Afterword. *Assessment and Evaluation in Higher Education, 35*, 347–349.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education, 36*, 395–407.

Carless, D. (2013). Sustainable feedback and the development of student self-evaluative capacities. In S. Merry, M. Price, D. Carless, & M. Taras (Eds.). *Reconceptualising feedback in higher education: Developing dialogue with students* (pp. 113–122). Oxon: Routledge.

Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment & Evaluation in Higher Education, 35*, 551–564.

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy, 17*, 305–308.

Creswell, J. (2003). *Research design: Qualitative, quantitative and mixed methods approaches.* Thousand Oaks, CA: Sage.

Daniel, L. M., & Poth, C. A. (2017). Relationships between pre-service teachers' conception of assessment, approaches to instruction, and assessment: An achievement goal theory perspective. *Educational Psychology*. http://dx.doi.org/10.1080/01443410.2017.1293800.

DeVellis, R. F. (2012). *Scale development: Theory and applications* (3th ed.). London: SAGE Publications, Inc.

Donovan, B., Rust, C., & Price, M. (2016). A scholarly approach to solving the feedback dilemma in practice. *Assessment & Evaluation in Higher Education, 41*, 938–949.

Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning, 20*, 81–94.

Dufrense, R. J., Gerace, W. J., Leonard, W. J., Mestre, J. P., & Wenk, L. (1996). Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education, 7*, 3–47.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 169–189.

Evans, C. (2013). Making sense of assessment in higher education. *Review of Educational Research, 83*, 70–120.

Fryer, J. W., & Elliot, A. J. (2007). Stability and change in achievement goals. *Journal of Educational Psychology, 99*, 700–714.

Fryer, J. W., & Elliot, A. J. (2012). Self-Regulation of achievement goal pursuit. In D. H. Schunk, & B. J. Zimmerman (Eds.). *Motivation and self-Regulated learning. theory, research and application* (pp. 52–68). Oxon: Routledge.

Han, C. (2016). Change and stability in achievement goal based on instructional tasks of a college classroom. *American Journal of Educational Research, 4*, 999–1007.

Harackiewicz, J. M., Barron, K. E., & Elliot, A. J. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist, 33*, 1–21.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81–112.

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research and Development, 27*, 55–67.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Fields Methods, 18*, 3–20.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review, 19*, 141–184.

López-Pastor, V., & Sicilia-Camacho, A. (2017). Formative and shared assessment in higher education: Lessons learned and challenges for the future. *Assessment & Evaluation in Higher Education, 42*, 77–97.

Mac Iver, D. (1987). Classroom factors and student characteristics predicting students' use of achievement standards during self-assessment. *Child Development, 58*, 1258–1271.

Madjar, N., Cohen, V., & Shoval, G. (2017). Longitudinal analysis of the trajectories of academic and social motivation across the transition from elementary to middle school. *Educational Psychology*, 1–27.

Maehr, M. L., & Midgley, C. (1996). *Transforming school cultures.* Boulder, CO: West view Press.

Middelton, M., Kaplan, A., & Midgley, C. (2004). The change in middle school students' achievement goals in mathematics over time. *Social Psychology in Education, 7*, 289–311.

Midgley, C., Kaplan, A., Middleton, M., Urdan, T., Maehr, M. L., Hicks, L., et al. (1998). Development and validation of scales assessing students' achievement goal orientation. *Contemporary Educational Psychology, 23*, 113–131.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Patterns of adaptive learning survey (PALS).* University of Michigan.

Midgley, C., Middelton, M., & Kaplan, A. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? *Journal of Educational Psychology, 93*, 77–86.

Muis, K. R., & Edwards, O. (2009). Examining the stability of achievement goal orientation. *Contemporary Educational Psychology, 34*, 265–277.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*, 199–218.

Nicol, D. J. (2007). E-assessment by design: Using multiple-choice tests to good effect. *Journal of Further and Higher Education, 31*, 53–64.

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education, 35*, 501–517.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education, 5*, 77–84.

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education, 35*, 535–550.

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2010). *Motivation in education: Theory, research, and applications* (3th ed.). Upper Saddle River, NJ: Prentice Hall.

Senko, C., & Harackiewicz, J. M. (2005). Regulation of achievement goals: The role of competence feedback. *Journal of Educational Psychology, 97*, 320–336.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Thousand Oaks, CA: Sage: Publications, Inc.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education Inc.

Trees, A. R., & Jackson, M. H. (2007). The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media & Technology, 32*, 21–40.

Tuominen-Soini, H., Salmela-Aro, K., & Niemivirta, M. (2011). Stability and change in achievement goal orientations: A person-centered approach. *Contemporary Educational Psychology, 36*, 82–100.

Urdan, T. C., & Midgley, C. (2003). Changes in the perceived classroom goal structure and pattern of adaptive learning during early adolescence. *Contemporary Educational Psychology, 28*, 524–551.

Warburton, V. E. (2017). Peer and teacher influences on the motivational climate in physical education: A longitudinal perspective on achievement goal adoption. *Contemporary Educational Psychology, 51*, 303–314.

ISBN 978-82-326-5193-1 (printed ver.)
ISBN 978-82-326-6175-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology