Karen Langfjæran

From Two Grades to One in Lower Secondary English

An exploration of teachers' experiences from navigating curriculum change

Master's thesis in English and Foreign Language Education
Supervisor: Anita Normann
November 2022

Master's thesis

NTNU
Norwegian University of
Science and Technology

Karen Langfjæran

From Two Grades to One
in Lower Secondary English

An exploration of teachers' experiences from
navigating curriculum change

Master's thesis in English and Foreign Language Education
Supervisor: Anita Normann
November 2022

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

NTNU
Norwegian University of
Science and Technology

Abstract

As part of the introduction of the LK20 curriculum, Lower Secondary English teachers now provide one grade rather than two, reflecting a move from grading oral and written performance to assessing *overall competence*. This newly introduced change to assessment in the English subject is preceded by no trial project, and is introduced without centrally provided, subject-specific support or guidance beyond the characteristics of goal achievement. This study examines the experiences of four teachers navigating this curriculum change in the 2021/22 school year to provide insights into how the reduction in number of grades translates to the local level of implementation.

This is a qualitative study employing a longitudinal design, and the data material consists of the transcripts of 11 semi-structured interviews with four teachers, conducted at three different points in time during the 2021/22 school year to access reflections in real-time, first to establish expectations and again soon after term grading in January and June. The focus of this study is the qualitative data, but a quantitative element has been added to provide context and gain a broader perspective on certain aspects of the change. While the aim is not to generalise, the questionnaire data from 115 Lower Secondary English teachers provide a helpful complement to understand the experiences of teachers with regards to this recent phenomenon.

The qualitative analysis led to the development of three themes covering assessment of overall competence, consequences on grading and motivation, and impact on teaching and assessment practice, collegial collaboration, and subject view. The qualitative findings indicate a need for guidelines on how to assess overall competence in the English subject and for grading more complex cases where pupils demonstrate incoherent competence. The respondents raised several validity and reliability concerns, demonstrating the complexity of transitioning with few guidelines. Overall, the teachers experienced few challenges relating to grading for the majority of pupils, but they also detailed instances where grading was not straight-forward. Furthermore, findings suggest that the change has advantages and disadvantages for different pupils. The teachers experienced that one grade is positive for pupils at risk of not receiving a grade or demonstrating generally lower competence levels, but that the change could also lead to fewer pupils receiving the grade 6. The findings further indicate that those who show significantly higher competence in one component could experience higher motivation and receive a more truthful assessment with two grades. The study elaborates on and discusses the findings with the aim of offering a thorough deliberation on the teachers' experiences with transitioning from two grades to one in the English subject.

Sammendrag

Med den nye læreplanen, LK20, er det innført én karakter i stedet for to i engelskfaget på ungdomsskolen. Dette gjenspeiler en forflytning av fokus fra muntlig og skriftlig vurdering til å vurdere *samlet kompetanse*. Denne nye endringen i engelskfaget innføres uten foregående prøveprosjekt og introduseres uten sentralgitte retningslinjer eller fagspesifikt støttemateriell utover kjennetegn på måloppnåelse. Denne studien undersøker fire læreres erfaringer mens de iverksetter endringen i praksis, skoleåret 2021/22, for å bidra til innsikt i hvordan reduksjonen i antall karakterer omsettes til handling på det lokale nivået.

Dette er en kvalitativ studie med et longitudinalt forskningsdesign, og datamaterialet består av transkripsjoner av elleve semistrukturerte intervju med fire lærere. Intervjuene er gjennomført på tre tidspunkt i løpet av skoleåret 2021/22 for å få tilgang til refleksjoner mens erfaringene er ferske. Første intervju etablerte forventninger, mens andre og tredje runde ble gjennomført kort tid etter terminvurdering i januar og juni. Denne studiens fokus er de kvalitative dataene, men disse er satt i et bredere kontekstuelt perspektiv ved å legge til et kvantitativt element. Hensikten med dette er ikke å generalisere, men spørreskjemadataene fra 115 engelsklærere i ungdomsskolen bidrar som et nyttig komplement for å forstå læreres erfaringer med dette nye fenomenet.

Den kvalitative analysen førte til utviklingen av tre tema, som dekker vurdering av samlet kompetanse, konsekvenser for karaktersetting og motivasjon og innvirkning på undervisnings- og vurderingspraksis, kollegialt samarbeid og syn på engelskfaget. De kvalitative dataene gir indikasjon på et behov for tydeligere retningslinjer for hvordan lærere skal vurdere samlet kompetanse i engelskfaget og for vurdering av mer komplekse vurderingstilfeller, hvor elever viser inkonsistent kompetansenivå. Respondentene viser til flere utfordringer knyttet til validitet og reliabilitet med ny vurderingsordning, som demonstrerer kompleksiteten i overgangen til én karakter med få retningslinjer. Helhetlig opplevde lærerne få utfordringer knyttet til vurdering av majoriteten av elevene, men de hadde også erfaringer hvor karaktersetting var mer komplisert. Videre antyder funnene at endringen har fordeler og ulemper for ulike elever. Lærerne erfarte at én karakter er positivt for elever som står i fare for å ikke få karakter eller som viser generelt lavere kompetanse, men at det også kan føre til færre seksere. Oppfatningen fra intervjuene er at elever som viser spesielt god kompetanse i et område kunne opplevd mer motivasjon og fått en mer sannferdig vurdering med to karakterer. Studien utdyper og diskuterer funnene med et mål om å gi en grundig fremstilling av lærernes erfaringer i overgangen fra to karakterer til én i engelskfaget.

Acknowledgements

Content

Figures

Tables

Abbrevations

CEFR	Common European Framework of Reference for Languages
FIVIS	Forskning på individuell vurdering (Research on Individual Assessment)
NESH	The National Committee for Research Ethics in the Social Sciences and Humanities
LK20	Norwegian National Curriculum 2020
The Directorate	Norwegian Directorate for Education and Training

1 Introduction

1.1 Research aim and purpose

The new curriculum, LK20, has introduced several major changes, including a reduction of grades in the English subject. This is in line with the curriculum's explicit emphasis on *overall competence*, a term highlighted as far back as the R94 English curriculum. Here, the overall competence was based on six themes[1] with a total of 29 key points. In comparison, the LK06 English curriculum for 8th–10th grade highlighted four themes[2] and 30 competence aims. Finally, LK20 introduced three core elements[3] and 19 competence goals.

Any change to assessment regulations will always have a consequence on teaching (Dobson & Engh, 2010, p. 21), and changes in a subject's content will lead to changes in assessment, as the subject curriculum determines what is to be assessed (Hasselgreen & Ørevik, 2020, p. 368). Going from two grades to one grade in Lower Secondary English represents a change to the framework for how teachers assess, and although English teachers in Upper Secondary School have a long experience of giving one grade only, there have been no official standards or supportive framework to facilitate the transition for Lower Secondary teachers. It is thus up to individual teachers and schools to operationalise the request. This may be explained by Norway's long tradition of upholding the local level's strong position in management of school policies (Bøhn, 2016, p. 7), which translates to high autonomy for local teacher communities.

This thesis explores four Lower Secondary English teachers' experience with the transition from two grades to one. I interviewed these teachers several times during the 2021/22 school year to see how they experienced implementing the curricular change. The purpose is to provide insights on how this curriculum change is experienced on the local level. The following research question serves to guide this thesis: How do teachers experience transitioning from two grades to one in Lower Secondary English?

1.2 Thesis structure

The aim of this thesis is to present a thorough investigation of some teachers' perspectives and experiences in the transition from two grades to one. This first chapter aims to contextualise the change in a broader assessment frame of reference. It also presents the results of my background research and literature review. It is followed by a chapter presenting relevant theoretical perspectives, with particular attention to the term 'overall competence'. Further, I present the methodology and material which forms the basis for the data analysed, and then discuss the findings in light of the theoretical perspectives, literature review and background research. Finally, a concluding chapter presents possible implications of this study and research suggestions.

---

[1] Understanding oral English, understanding written English, use of oral English, use of written English, the English-speaking world (Norwegian Ministry of Church, Education and Research, 1993)
[2] Language learning, oral competence, written competence, and culture, society & literature (Norwegian Directorate for Education and Training, 2013).
[3] Communication, language learning, and textual encounters (Norwegian Directorate for Education and Training, 2020a).

1.3 Assessment Context

In the last few decades, there has been a significant increase in the emphasis on assessment in the Norwegian school system. In 2007, the Norwegian Directorate of Education and Training (henceforward, the Directorate) stated that the Norwegian educational system had a weak assessment culture and practice, and that neither schools nor teacher training institutions had a sufficient level of assessment competence (Meld. St. 16 (2006–2007)). This, of course, led to an increased emphasis on assessment in schools, and while this was commanded by OECD in its 2011 assessment of the Norwegian educational system, the report did also emphasise the need for clearer, more visible reference points and assessment criteria, and added that "There seems to be little shared understanding regarding what constitutes adequate, good and excellent performance in different subject areas and year levels" (Nusche et. al., 2011, p. 52).

The subsequent focus on assessment, e.g. the national Assessment for Learning initiative, certainly marked a change in the research, and a report from 2018 concluded that the initiative had indeed led to a more learning oriented assessment culture, increased competence and positive change of assessment practice as well as a more active use of the curriculum and a common assessment language (Norwegian Directorate for Education and Training, 2018a). Consequently, much has changed in how we view assessment in the last decade.

1.4 Background

In the 2019 hearing about the English curriculum, schools and other stakeholders were asked to answer 22 questions, including their thoughts on introducing one grade in Lower Secondary School rather than the existing two. The hearing documents established that the reason for the suggested change is the low number of hours in the subject, plus good experience in Upper Secondary school (Norwegian Directorate for Education and Training, 2019a). The former reason refers to the fact that the English subject has a total of 222 hours, which is allocated across three years (typically 2–3 lessons weekly). By comparison, the Norwegian subject has 398 hours, and the pupils receive three grades in summative assessment, whereas foreign language subjects such as Spanish and French have 222 hours in a school year, and the pupils receive one grade. As for the second reason, good experience with giving one grade in Upper Secondary schools, I have not been able to find any research or documents to support the claim of good experience. Furthermore, I am unable to find any documents regarding the move from two grades to one grade in Upper Secondary, which was implemented with the R94 reform, if not earlier.

According to an unpublished hearing summary, 86 schools prefer two grades, whereas almost 200 schools prefer one grade. This number includes Upper Secondary schools, which are not directly affected by the change. Nine out of ten Universities and University Colleges prefer one grade. The summary highlights the following arguments for providing one grade rather than two: It opens for more multidisciplinary collaboration, corresponds well with the low number of hours, provides more flexibility, and makes it easier to have a good assessment process. The arguments highlighted against the change are that the English subject loses status, oral competence will be less prioritised, and that oral and written competences should be assessed separately. Both "sides" highlight that their favourable situation makes it easier to adapt to pupils' individual needs. Hearing responses pointing to stress related to having many grades, and the English subject consisting of more than two competences, are highlighted as contributing factors to the

conclusion of reducing the number of grades (Norwegian Directorate for Education and Training, 2019b, unpublished). Furthermore, the Directorate states that the feedback show signs of skills and competences being viewed as separately entities, that this can hinder pupils from having a holistic approach to the English subject, and that the two-grade system may have encouraged a mathematical approach to assessment (Norwegian Directorate for Education and Training, 2019, unpublished).

I have also identified two relevant articles. In one of them, two representatives from the Directorate were interviewed, and explained that "we hope the pupils' competence will now be assessed as an overall competence … Oral and written skills in English are not separate entities which should be weighed against each other. Pupils use oral and written competence interactively, and so they must be given the opportunity to show their English competence by combining several basic skills" (Norwegian National Centre for English and other Foreign Languages in Education, 2020, p. 5, my translation). In the second article, researchers Brevik, Flogenfeldt and Beiler (2020, p. 47) stated that the change would hopefully make "pupils more secure as language users, because one grade reflects that different pupils master different skills in different ways". There are few mentions of the change on the Directorate's website, but I have identified one in a text relating to quality criteria for the selection of learning materials (Norwegian Directorate for Education and Training, 2021a). The change is here briefly mentioned as potentially positive for those with reading and/or writing challenges such as dyslexia, as it increases the freedom to demonstrate competence using different modalities.

I cannot find any reference to any prior trial projects or relevant research, and to the extent of my literature review, it has only been mentioned in one master's thesis. Here, teachers were asked in a digital questionnaire, "Do you think going from two grades to one is positive?". 69 out of 139 EFL teachers (54,8%) responded yes, whereas 57 (45,2%) responded that they did not think it was a positive change. It was also discussed in interviews, where one teacher said it was positive as it would hopefully result in a more holistic mastery of the subject and more confidence. A second teacher implied it might be an advantage for pupils on a lower grade level, but negative for pupils who are strong in either oral or written English (Eriksen, 2021, p. 28–29). This questionnaire result, albeit small, is in agreement with the Directorate's conclusion that the majority is positive to the change.

During my work on this thesis, I have found myself looking for information: Some has been hard to find, and some is left unanswered. I emailed the Directorate on 30th May and again on 23rd August 2022 with a number of questions related to the curriculum change but have received no reply beyond receipt of both emails. Meanwhile, I have received timely responses to other requests. Some questions still unanswered are:

- When you say that the background for one grade in Lower Secondary English is the subject's allocation of hours and good experience from Upper Secondary (as explained in the hearing documents[4]): Which knowledge basis is this based on? Is the good experience from Upper Secondary documented?
- Is this change followed up with supportive material for teachers?
- I have looked at the assessment module in UDIR's online support material[5] but cannot find anything specifically regarding this change in the English subject. I do, however, find a subject-specific page on how final grading in Mathematics should

---

[4] https://hoering.udir.no/Hoering/v2/338
[5] https://www.udir.no/kvalitet-og-kompetanse/kompetansepakker/

be understood[6], supported by example tasks. Have you considered doing something similar to explain the considerations behind the changes in the English subject?

- In 2016, a survey showed that only half of those asked, knew about existing supportive material[7]. Has the Directorate, and if so, how, actioned this to make supportive material more accessible for teachers?

Beyond the general characteristics of goal achievement[8], there is little evidence of any supportive material to facilitate this particular curriculum change in the English subject. Moreover, the Regulations to the Education Act §3[9] (2020) says nothing about how teachers should weigh the competence, ensure that the basis of assessment is thorough enough or interpret conflicting evidence of competence (Skar and Hopfenbeck, 2021, p. 6). I have found one competence development measure in Trøndelag county[10] where teachers of English participate in educational gatherings about assessment, including summative assessment. Such measures could prove to be rewarding as the Directorate has previously noted that "experience shows that introducing new curriculums, regulations or other measures in the sector without following up with competence development leads to a low degree of practice change" (Norwegian Directorate for Education and Training, 2009, p. 26, my translation). Moreover, Robinson (2018) claims that the simultaneous implementation of many changes leads to superficial development, and not lasting change in actions or practices. Potential practice change is a focal point in this thesis.

1.5 Literature Review

Despite the lack of relevant literature on this specific theme, a similar change has been implemented in the Norwegian subject in a more planned manner. The Directorate introduced a trial project in 2013 where teachers gave pupils either one or two grades at term grading until the Spring term in 10th grade, when they received three grades[11]. The Directorate's conclusion then was that it would ensure teachers a better work situation and enable them to focus less on documentation for the summative assessment basis (Seland et. al., 2018, p. 28). Prior to implementation, this was followed up with a trial project and three subsequent reports. The final report concluded that most school leaders reported that it had made term grading easier to organise and complete instruction and assessment situations in the Norwegian subject (p. 107), and that it has had a motivating effect on teachers (p. 150). Grade analysis of 197 000 pupils' overall and summative results showed no difference in learning results when comparing pupils who participated in the project with pupils who did not (p. 150).

With regards to relevant literature relating to the English subject, Sandvik et. al. (2012, p. 158) has pointed out that the developments to promoting better assessment practices

---

[6] https://www.udir.no/laring-og-trivsel/lareplanverket/fagspesifikk-stotte/standpunktvurdering-i-matematikk/

[7] https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/Hvordan-bruker-larerne-veiledninger/

[8] There is no official translation, but the direct translation is 'characteristics of goal achievement'. Engh (2012) operates with the term 'indicators of success', and Bøhn (2019b) uses 'rating scale guidelines'.

[9] The Norwegian regulations concerning assessment.

[10] https://www.matematikksenteret.no/nyheter/satsing-p%C3%A5-vurderingskompetanse-i-tr%C3%B8ndelag

[11] Nynorsk, Bokmål and Oral.

up until 2012 had primarily had a pedagogical focus and that the majority of the didactic focus of individual subjects was still to be 'translated'. In the 2019 collection of doctoral research on English didactics between 1988 and 2017, Rindal and Brevik (2019) argue that although English didactics has traditionally been a practically oriented domain developed from research conducted in more established fields, it has in recent decades firmly established itself as a research field in Norway. Fields include English writing, Digital English competence, Reading in English, Culture and Literature, and Oral Proficiency, corresponding well with the content of the English curriculum. Interestingly, only two PhDs were related to the field of assessment, which adds substance to the claim that assessment research primarily has had a pedagogical rather than a didactical focus.

These two PhD's are Tony Burner's "Formative assessment of writing in English. A school-based study of perceptions, practices and transformations" (2016) and Henrik Bøhn's "What is to be assessed? Teachers' understanding of constructs in an oral English examination in Norway" (2016). While the former is concerned with formative assessment, the latter looks at the rating process of oral English examinations. Although the study concerns Upper Secondary, Bøhn's PhD is very informative on somewhat more general ideas connected to all assessment of oral activity in the English classroom. For example, Bøhn (2016) points out that we can find traces of both the measurement paradigm and the assessment paradigm in the Norwegian educational system, and while summative assessment tends to share more features with the measurement paradigm, there are no common exam tasks or rating scales for oral English exams. He also argues that assessment contexts with no common scales "are special, and they beg for closer scrutiny" (Bøhn, 2016, p. 8). Furthermore, Bøhn (2016) concludes that Norwegian educational authorities should consider introducing common rating scale guidelines and more coherent rater training for high-stakes oral L2 testing situations. In his "PhD Revisited", Bøhn (2019a, p. 392) emphasises the need for guidelines, rather than fixed rating scales, as "the fixed scales may make the teachers focus too strongly on the criteria for assessment rather than on competence aims of the subject curriculum when teaching English". As the summative assessment in the English subject, in my view, belongs to the category of an *assessment context with no common scale*, his statements lend themselves to more general views on assessment in the English classroom. One understanding could be that such rating scales correspond to the characteristics of goal achievement, but these are very general.

Generally, the final grade is emphasised as of greater importance to the individual pupil's future prospects than exams, as the latter represent only a very small number of grades on a diploma. In 2009, little existed to document assessment practices besides the fact that the grounds were often based on more than the subject, and that the curriculum is operationalised on a local level (Norwegian Directorate for Education and Training, 2009). On those grounds, the Directorate recommended to have a research-based review conducted of the summative assessment practices held by Norwegian teachers focusing on what, as well as on what grounds, teachers emphasise when giving summative assessments (Norwegian Directorate for Education of Education, 2009, p. 42–43). A later report on summative assessment stated that written documentation of pupils' performance and teachers' notes was the main basis of summative assessment in several subjects (Prøitz & Borgen, 2010, p. 11). Moreover, it emphasised that teachers perceive a subject's degree of interpretive work as a substantial factor in providing fair assessments. It did not look at assessment in the English subject, but did conclude that overall, using assessment tools such as rubrics, criteria or point systems seemed to strengthen the experience of just assessments regardless of the subject (p. 109–110). In

2021, the Norwegian Centre for Writing Education and Research published a review on summative assessment[12]. Authors Skar and Hopfenbeck (2021) concluded that there is little research-based knowledge on teachers and school leaders' practice and the quality of teachers' summative assessment practices (p. 3), but that several studies document a lack in coherence between the final grades and exam grades, with a consistently lower result in written exams (p. 23). Furthermore, several studies show that pupils have difficulties understanding what components constitute their final grade, and similarly, this insecurity is shared by teachers in Norwegian studies. The report also highlighted that most of the reviewed research had looked at *experience* of practice; not *practices*. Also, the largest study was over 10 years old and based on limited material (Skar & Hopfenbeck, 2021).

Among the studies highlighted in the review is one which looked at Upper Secondary schools who had eliminated grading throughout the term in the Norwegian subject and found that the decreased focus on grades resulted in an increase in attention to learning progress among the pupils (Eriksen & Elstad, 2019). In a recent article, Sandvik (2022) establish that reduced focus on grades must be followed with a change in assessment practices, including a goal-oriented practice with clear goals and criteria, and established interpretation communities. A related project, Going Gradeless (2020–2024), is currently investigating consequences of reduced grading in subjects including English in Upper Secondary schools. It will be interesting to see its conclusions, compared to Eriksen and Elstad's (2019). Nevertheless, these projects look at formative assessment throughout the school year as opposed to term grading and summative assessment, but are relevant due to the current move towards reduced focus on grading.

Finally, several research projects evaluate the implementation of LK20, such as EDUCATE (2021–2025) and EVA20 (2019–2025). The former is a comparative study of classroom practices in seven subjects, including English, whereas the latter has no subject-specific aim. The EVA20 project's first report emphasised that the authorities had invited to a shared, or joint, project, and that it would inevitably lead to tension, dilemmas and disagreement (Karseth, Kvamme & Ottesen, 2020, p. 16). It is these tensions and dilemmas occurring as a result of the transition from two grades to one in the English subject that I draw attention to in this thesis.

---

[12] The basis for the review was 305 documents, 3 357 800 assessment results, and 42 publications.

2 Theoretical Background

This chapter seeks to inform about relevant theoretical perspectives regarding assessment and grading. The chapter is structured in two parts, where the first section informs about more general topics within assessment, while the second focuses on assessment in the English subject, with a particular emphasis on what encompasses competence and constructs in the English subject. The decision to include a section on assessment relating to general terms such as validity and reliability, and a presentation of functions and tensions related to assessment, is made because these are found to be important points in relation to the thesis' findings. To discuss these matters of assessment in the English subject without relating it to validity and reliability, and to discuss the curriculum change without relating it to the tension between formative and summative assessment, what assessment is and should be, and other pedagogical perspectives, is impossible. In light of the new LK20 curriculum, where teachers are instructed to view competence in a holistic manner and think interdisciplinary, a similar lens must be added to this topic.

2.1 Assessment and grading

Assessment can be viewed as a process consisting of two steps: Eliciting evidence regarding the level of knowledge or ability of a person, and then interpreting this evidence to make conclusions about the underlying knowledge (Wiliam, 2000). This interpretation process is essential, as traditional tests typically "can assess only a small part of the learning of which they are claimed to be a synopsis" (Wiliam, 2000, p. 1). Furthermore, we need to distinguish between performance and the associated cognitive skills, i.e. competence (Chomsky, 1965). Westera (2001) emphasises that 'competence' is a theoretical term, but that competent performance presumes competence.

Assessment serves a variety of functions: to monitor national standards, to provide information with which teachers, educational administrators and politicians can be held accountable, and providing a method for selecting students for employment and subsequent providers of education and training (Wiliam, 2000, p. 1). Engen (1996, p. 57) argues that this 'sorting' function of summative assessment is of great importance for individual pupils' lives, and calls it both a 'receipt' (of acquired knowledge) and a 'ticket' (for further education or work). A given grade, expressing the sorting function, can then be understood as information about individual fates, created for quick interpretation (Engen, 1996, p. 63). Arguably, however, the most important function of assessment is in supporting learning, says Wiliam (2000), and emphasises that there is potential for integration, or alignment, between formative and summative assessment. Although there is tension between the two, Wiliam (2000, p. 15) insists "we must refuse the incompatibility of 'summative' and 'formative' assessment". According to Wiliam (2000), the conceptualised distinction between formative and summative assessment has resulted in a distinction between educational assessment and *learning*; and emphasises the importance of finding ways of mitigating this tension.

This thesis has the end of year grading at its centre of attention, and Gardner (2012) argues that final (summative) assessment will potentially not contribute to students' learning. In fact, summative assessment is often referred to as Assessment of Learning,

which Gardner (2012) argues is contra-distinct to assessment to support learning, or Assessment for Learning (AfL). On the other hand, Black & Wiliam (2018, p. 552) argue that a distinction between the two is "at best unhelpful, and may even be counterproductive". Instead, they suggest that all aspects of assessment must be viewed in an integrated way to improve classroom practices focusing on assessment (Black & Wiliam, 2018). This tension between formative and summative assessment is interesting in this context, perhaps particularly because Burner (2020, p. 54) points out that in LK20, all assessment in the English subject is supposed to be formative until Year 10, as summative assessment is only described after Year 10 and Vg1. Term grading is indeed placed below the formative assessment regulations, but should "express the competence the pupil has reached based on the expected level" (Regulations to the Education Act, 2020, §3-12). This, however, resembles the optics of summative assessment, which "takes place at the end of a course or a module, providing information about the overall level of competence that a student has attained in the subject" (Hasselgreen & Ørevik, 2020, p. 372). Using the below definition of formative assessment by Broadfoot et. al., 2002, p. 1), it *can* be applied to the term grading in 8th and 9th grade, if using grades in an ipsative manner, for example to motivate for the next term:

> Assessment for Learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.

In other words, the distinction between summative and formative assessment is unclear, and as a result, this requires a local interpretation.

A second tension of importance related to summative assessment is the one between the educational ambitions and teachers' autonomy. Grading demands qualified interpretation processes by teachers (Nordgren & Odenstad, 2017, p. 49), but simultaneously, teachers in many countries are often under implicit or explicit pressure to 'teach to the test' (Black & Wiliam, 2018). Norway is no exception, and in a report from the Norwegian FIVIS research project, Sandvik et. al. (2012, p. 140) noted that we may assume that the exam is used as a guideline for assessment and content in the English subject. On the other hand, teachers are found to use their experiences and theoretical interests when shaping their independence and create their professional space as teachers (Myhr, 2014, p. 165). This balancing act may appear to be a challenging one, as teachers are expected to provide assessments which indeed serve many various functions, and sometimes conflicting ones.

Selghed (2010) argues that the construction of grading systems is problematic when used as a selection instrument, or for comparative purposes, which is the case with final summative assessment, because it contrasts the intention of giving an equal and fair assessment across the nation. The ideological principle of fairness in education is strong in the Norwegian school system and requires an educational policy framework which opens for both national criteria and locally developed criteria, where validity is central. For some purposes the validity demands are at a national level attempting to meet a national goal, whereas others are connected to the individual class and pupil (Eggen, 2010, p. 53). National standards are, however, criticised by e.g. Gordon Stobart of the Assessment Reform Group, as "counter-productive" (in Engh, 2010, p. 64). On the other hand, Bøhn (2019a) has emphasised the need for common rating scale guidelines for high-stakes assessment contexts such as EFL oral exams, albeit in an Upper Secondary context, as the lack of a common national rating scale may affect the validity and reliability of the grading.

Assessment discourse has historically been framed as a discussion of validity, reliability, purpose and assessment expressions (Eggen, 2010, p. 40). Black et. al. (2011) comprises validity and reliability in the term 'quality', and Bøhn (2020, p. 308) emphasises the importance of validity and reliability in order to secure fair, dependable and well-founded assessment. To evaluate the validity of assessment, a comparison must be made between the skills or knowledge that we want to assess, and what is actually assessed; to evaluate reliability, we consider if the assessment would produce the same results if it had been scored by different people (Cameron, 2001). Applying Black et. al.'s (2011) understanding of assessment *quality*, we must consider if we are indeed assessing what we want to assess, and whether the circumstances regarding the assessor and assessment situation can contribute to interference in a given assessment. If an assessment situation has low reliability and/or low validity, we may speak of a low quality of assessment. Bøhn (2020, p. 309) suggests that one way to improve validity and reliability is to develop a shared assessment culture among teachers, systematically discussing for example competence aims, criteria, and performance at different grade levels. Furthermore, Hasselgreen & Ørevik (2020, p. 385) sums up the main objectives of assessment in the subject of English in Norwegian schools as "providing means of assessment for both summative and formative purposes with as high levels of validity and reliability as possible" as well as providing qualified feedback which further supports the development of English competence. Neither validity nor reliability are mentioned explicitly in the subject curriculum, but both are mentioned as quality criteria in the general framework for development of exams (Norwegian Directorate for Education and Training, 2021b).

Summative assessment should be valid, reliable and fair (Norwegian Directorate for Education and Training, 2022a). This requires teachers to be *assessment literate.* Stiggins (1995 in Fjørtoft & Sandvik, 2016, p. 18–19) includes the following five aspects in assessment literacy: purpose, a deep understanding of content and connection between learning goals, assessment methods, data collection of pupils' learning outcomes, and biases in assessment. Burner (2020, p. 58) argues that assessment literacy also involves "the understanding and practice of assessing the seemingly un-assessable: competences in deep learning, democracy and citizenship, health and life skills". Assessment literacy can thus be understood as a complex term. Interestingly, teachers may not feel they have enough time to integrate changes in the classroom and may require leadership-initiated projects in order to apply new assessment practices (Fjørtoft & Sandvik (2016, p. 19). This tension between expectations and demands on one side, and time restraints on the other, is undoubtedly relevant in this context.

Finally, the purpose of summative assessment is to "determine the student's overall achievement in a specific area of learning at a particular time" (Moss, 2013). In other words, the overall achievement grade should indicate the level the pupil has reached by the end of the year and should not be based on an average of test results throughout the whole school year (Hasselgreen & Ørevik, 2020, p. 370). There can be great variations in pupils' progression, meaning that they develop at different speeds and not necessarily in defined lines, and this creates challenges for teachers when attempting to create a varied education for their pupils whilst simultaneously fulfilling the requirements in the Regulations to the Education Act §3 and securing an assessment which is as valid as possible (Langseth, 2016, p. 142). This could inevitably create challenges regarding documentation for the end of year grading, if indeed a pupil develops their competence *after* an assessment situation 'measures' it, but has no chance to show this.

21

2.2 Assessment in the English subject

According to Nordgren & Odenstad (2017, p. 10), teachers signal their subjective perspectives on subject didactics and learning through their teaching and assessment, and the latter is typically guided by several important subject didactical decisions, for example

- What is important to know in the subject? What type of learning do I want to encourage? What skills are central?
- How do I design assessment tasks so that pupils are given the opportunity to show their knowledge and ability?
- Has the question of 'for who' the assessment task is for, importance for the task's content and design? What is my view on pupils and their abilities?

These questions are all concerned with why teachers make the decisions they do, and what consequences their choices could have (Nordgren & Odenstad, 2017). Moss (2013, p. 235) argues that what teachers assess, how and why, sends "a clear message to students about what is worth learning, how it should be learned, and how well they are expected to learn it". This sub-section is guided by the first question in relation to the English subject, namely what encompasses competence and constructs in the English subject. Ideally, I would also include sub-sections regarding task design and assessment literacy among Norwegian EFL teachers, i.e. how and why assessment is conducted, to lay out a theoretical foundation regarding the two latter questions as well, but due to length restrictions, I have had to prioritise one over the others. It is important to note that this thesis attempts neither to interpret the respondents' task design nor assessment literacy, so these additions would only serve to inform relating to the research question: *How do teachers experience transitioning from two grades to one in Lower Secondary English?* Nevertheless, certain aspects are added, where relevant.

2.2.1 Overall competence

In the end of year assessment in Lower Secondary English, teachers are now to use one grade to assess *overall competence*. A Cambridge University Press & Assessment report states that "[w]ithout a comprehensive understanding of overall competence, we run the risk of creating educational programmes that are only superficially related to competence, ultimately limiting their value" (Vitello et. al., 2021, p. 7). A holistic view of competence highlights, according to Vitello et. al. (2021), an interconnectedness of different factors involved in competence, and encompasses individual internal factors such as knowledge and skills as well as attitudes and values, and contextual factors. Vitello et. al. (2021, p. 11) use the following definition:

> Competence is the ability to integrate and apply contextually-appropriate knowledge, skills and psychosocial factors (e.g. beliefs, attitudes, values and motivations) to consistently perform successfully within a specified domain.

This is relatively close to the definition used in LK20 (Norwegian Directorate for Education and Training, 2020a):

> Competence is being able to acquire and apply knowledge and skills to master challenges and solve tasks in familiar and unfamiliar contexts and situations. Competence includes understanding and the ability to reflect and think critically.

Assessing *overall competence* is further explained as evaluating a pupil's competence based on all competence aims in context of each other (not a selection) (Norwegian Directorate of Education and Training, 2022b). In a Norwegian educational context, the

term *holistic competence* was introduced in a guideline document regarding pupil assessment accompanying the L97 reform, and explained in a model showing the overall content of holistic competence the way it was understood then:

Holistic competence

Knowledge Competence   Method Competence   Social Competence
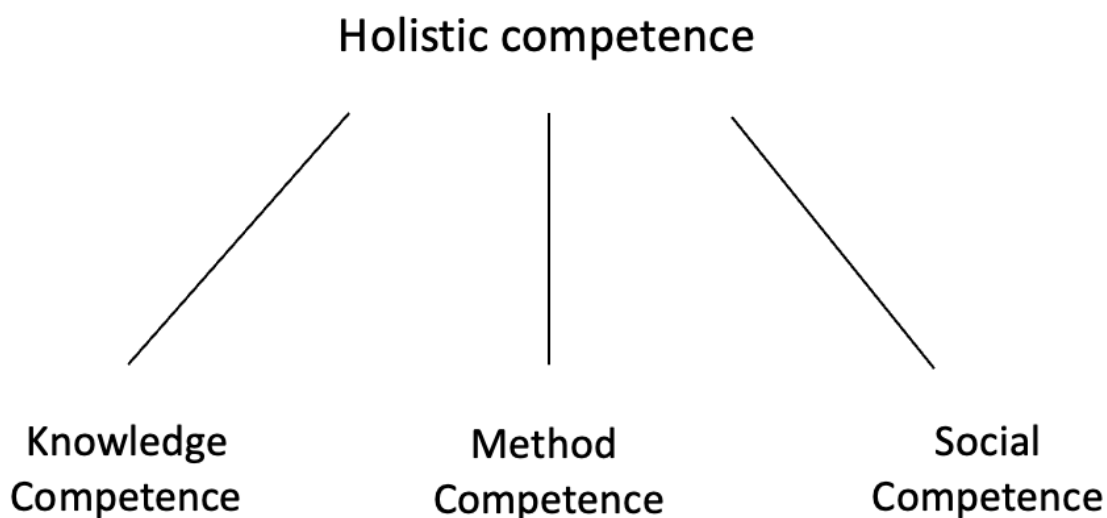
Figure 2.1: Holistic competence (Norwegian Ministry of Church, Education and Research, 1998, p. 23)

The three components[13] are interdependent, and the two latter components are necessary building blocks to develop subject competence, argues Aasland (2009, p. 200), who claims that a view on subject competence not relying on this interdependence could result in an authoritarian teaching style, using method suchs as memorising. The assessment of pupils' social and emotional competence is further emphasised in NOU 2015:8, which highlights a broad view on competence and more systematic support of pupils' social and emotional learning and development in individual subjects. Social and emotional competence is portrayed as highly important and tightly connected to the acquirement of subject competence, and includes e.g. curiosity, self-regulation, and respect for other views (NOU 2015:8, p. 80). Furthermore, NOU 2015:8 presents dilemmas in assessing a broad view on competence, because of the subjective nature of social and emotional competence, and suggests not to put weight on this in final summative assessment, but rather see it as the pupil's starting point. The below table presents what aspects are encompassed in social and emotional competence, per NOU 2015:8[14] (my translation).

---

[13] Translated from *kunnskapskompetanse, metodekompetanse og sosial kompetanse.*
[14] Categories translated from *fagspesifikk kompetanse, å kunne lære, å kunne kommunisere, samhandle og delta, og å kunne utforske og skape.*

Table 1. Aspects of social and emotional competence emphasised in the competence areas (NOU 2015:8, p. 22)

Subject-specific competence	Ethical judgment skills, dedication, attitudes towards the subject and personal learning in the subject
To learn	Endurance, expectations to personal mastery, and the ability to plan, implement and evaluate personal learning processes
To communicate, interact and participate	To express oneself and participate, show communal consideration by regulating one's own thoughts, feelings and actions, acknowledge that interaction and participation is based on interdependency, and respect and see the value of others' view points
To explore and create	Curiosity, endurance, openness to see things in new ways, and the ability to take initiative

The latter component in the above table, *to explore and create* has been incorporated in the competence aims, and the verb 'explore' is one of the main verbs used in the LK20 curriculum. Meld. St. 28 (2015–2016) emphasised that pupils' social and emotional learning should not be included in the competence aims more so than it already was in LK06, as "these skills should not be part of the basis of assessment of pupils' subject competence" (p. 23). To 'explore' has received the following definition by the Directorate (2018b, my translation):

> To explore is about experiencing and experimenting, and may [ivareta] curiosity and wonder. To explore can mean to use one's senses, to search, discover and examine. In some cases it means to investigate different sides of the same issue through an open and critical discussion. To explore can also mean to test or evaluate methods, products or equipment.

For the English subject, it is used 4 times in the competence aims for the 10th grade and 22 times in total in the English subject curriculum. To compare, it was used once for the English subject curriculum in LK06, in relation to reading (Norwegian Directorate for Education and Training, 2013). The introduction of 'explore' in competence aims is used as an example by the Directorate in a report that the foundation of summative assessment has become broader, in the sense that the core curriculum is incorporated in the competence aims, and that the core curriculum should "actualise … discussions about what sort of competence should be part of the basis for summative assessment. It is important to discuss whether the assessment of the final competence should reflect pupils' broad and holistic competence to a greater degree, and be more consistent with what is expressed in the core curriculum" (Norwegian Directorate of Education and Training, 2020c, my translation). The core curriculum states that competence aims must be seen in relation to all parts of the curriculum, including the competence aims of other subjects. Moreover, competence aims should be viewed in light of the "About the subject" chapter (Norwegian Directorate of Education and Training, 2020d), which

consists of the texts "Relevance and central values", "Core elements", "Interdisciplinary topics", and "Basic skills". However, the Directorate's (2020c) report on assessment and future examinations, mentioned above, clarifies that the content of the core curriculum beyond the values and principles incorporated in the competence aims should not be part of the basis of assessment. Based on the text "Relevance and central values" in this chapter of the subject curriculum, values integrated in the English subject are thus *(inter)cultural understanding, communication, all-round education, identity development* as well as using English to *learn, communicate and connect with others. Promoting curiosity* and *helping to prevent prejudice* are also mentioned in the subject-specific values text. Furthermore, the three core elements incorporated in the English subject are Communication, Language learning, and Textual encounters (Norwegian Directorate of Education and Training, 2020a), which are emphasised as the most important content the pupils work on (Norwegian Directorate of Education and Training, 2019d). As this is important for the content of the subject, it also becomes relevant for assessment purposes.

In 2020, three books were published on the English subject in light of the LK20 curriculum (see table below). These three books target perhaps teacher students and academics more so than practicing teachers, but the varied content shows the complexity in assessing the overall competence in the English subject. In the below table, I present the themes and topics raised in these three books (not inclusive chapters on e.g. historic perspectives), with the aim of adding context to, indeed, the complexity of the developing English subject.

Table 2. Presentation of topics in three books on the English subject after the introduction of LK20

Literature	Themes and topics
Brevik & Rindal (Eds.) (2020): Teaching English in Norwegian Classrooms. From research to practice. Universitetsforlaget.	Communication: digital competence, vocational English Language learning and literature: language use, pronunciation, strategic reading Textual encounters: literature, English use outside of school, intercultural competence
Fenner & Skulstad (Eds.) (2020): Teaching English in the 21st Century. Fagbokforlaget. 2nd edition.	Basic concepts: Communicative competence, intercultural competence Skills and competences: oral skills, writing, reading and literacy, digital technology, vocabulary, grammar, approaches to literature, multimodality
Carlsen, Dypedahl & Iversen (Eds.) (2020): Teaching and Learning English. Cappelen Damm.	Plurilingual learning and teaching, intercultural competence and culture, oral skills (speaking and listening), writing, digital learning, reading, multimodality, literature, grammar, pronunciation, vocabulary, language awareness

Brevik and Rindal (2020) presents content using the core elements as overarching themes aiming to reach teachers and student teachers, while Fenner & Skulstad (eds., 2020) attempts to discuss the interlinked questions "*why* should English be taught, *what*

is to be taught and learnt, and, finally, *how* can the subject be taught and learnt?" (p. 9). They explain the choice of skills by referring to the four basic skills specified in LK20, noting that devoting a whole chapter to multimodality shows how it has become an increasingly important part of English in the 21ˢᵗ century (p. 11). Multimodality also received a chapter in Carlsen, Dypedahl & Iversen (eds., 2020), and the chapter author Rimmereide (2020, p. 206–207) argues it is challenging to assess, as the focus is typically on the linguistic aspect rather than the overall presentation, and there is not yet an established practice. A chapter is also devoted to plurilingual learning and teaching, where Haukås and Speitz (2020) emphasise the value of plurilingualism, or *an individual's repertoire of languages,* and explain how LK20 lays a foundation for teachers to build on pupils' earlier experiences with language learning and increase their metalinguistic awareness (p. 67). Nevertheless, studies claim that it is a typical practice across Europe to keep languages apart in instruction (Haukås & Speitz, 2020). A third relatively modern concept is 'vocational English', which is framed as "not only designing instruction directed at students' future profession, but also building relevance connected to their use of English in society at large, in their youth cultures and for personal purposes" (Brevik, Skarpaas & Isaksen, 2020, p. 65). In 2004, a study established that Norwegian pupils answered they have achieved 34 % (median) of their English competence through media exposure and approximately half in school. Interestingly, media exposure outside of school correlated positively with test results, whereas media exposure in educational activities did not (Ibsen, 2004). Almost 20 years later, media exposure in the classroom has increased, and perhaps with good reason. Brevik, Skarpaas and Isaksen (2020, p. 66) emphasise that building relevance by connecting to pupils' (potential) future professions, their youth cultures and for personal purposes is key to motivate them. Nevertheless, studies have found that vocationally oriented instruction can be less successful if it is perceived as artificial or irrelevant (Brevik , Skarpaas & Isaksen, 2020). Furthermore, vocational classrooms are not too relevant for Lower Secondary, but Hestetræet & Ørevik's (2020) chapter "English vocational studies" in Fenner & Skulstad (eds., 2020) also emphasise the clear link between motivation and real-life interest. Ultimately, being able to use extramural English, i.e. informally-acquired English, has great advantages (Hestetræet & Ørevik, 2020).

With the introduction of these relatively new aspects in English didactics, the English subject based on LK20 can perhaps be interpreted as a twofold subject: On one end, there is traditional content which is part of all language learning: language use, pronunciation, reading, oral and written skills, vocabulary, grammar, writing. The other reveals a more modern component including digital competence, multimodality, vocational English, intercultural and plurilingual competence. The purpose for adding this context is to show that the subject encompasses a great variety of competences, skills, and topics. Hasselgreen and Ørevik (2020, p. 370) calls the LK20 changes to the English subject "the integrated nature of the new English subject and the ensuing need to test broader aspects of students' competence". The way I understand this 'new' English subject, then, is one where competences are intertwined in one holistic understanding, rather than viewing different competences as separate entities, and viewing these in relation to each other without comparison or judgment. In the following section, I attempt to explain what *entities* this 'new' English subject should include, supported by the Common European Framework for Languages, and what this means for assessment in the 'new' English subject.

2.2.2 Constructs in the English subject

The term *construct* is commonly used to conceptualise what is to be assessed (Bøhn, 2019b). A *construct* can be understood as abstract nouns such as *fluency*, or broader categories of concepts, alongside terms *criteria, sub-criteria* and *sub-sub-criteria*, which explain more narrowly defined performance aspects. In the latter understanding, a construct can be understood as *communication*, while a criterion can be *linguistic competence*. *Grammar* is then understood as a sub-criterion with this hierarchical understanding (Bøhn, 2015). Knowing what to evaluate is an important aspect of assessment literacy, and research has established that assessors frequently evaluate what they should not and fail to assess what they should (Bøhn, 2019b, p. 101). Bøhn (2015, p. 5) found that teachers focused on two main constructs when investigating the English oral exam context with Upper Secondary teachers, namely *communication* and *content*, with the former including for example *linguistic competence* as a sub-criterion. Bøhn (2015) explains this as fitting with a communicative approach. Hasselgreen & Ørevik (2020, p. 372–373) have identified *communicative competence* as the core ability most typically assessed in language learning contexts, encompassing *linguistic competence, sociolinguistic competence, strategic competence, and discourse competence*, and argue these can be distinctly operationalised for purposes such as making tests and reporting scales. Moreover, Skulstad (2020, p. 43) calls communicative competence "*the* single most important concept in English didactics".

One way to understand the term *construct*, then, is as comprehensive terms such as *communication;* another as abstract nouns like *fluency*. A third is that "the subject curriculum of English forms the construct against which students' overall achievements in English are measured" (Hasselgreen and Ørevik (2020, p. 375). The latter constitutes a general understanding, and this sub-chapter attempts to break this down to understand what is to be assessed in the English subject. To do so, I lean on the Council of Europe's Common European Framework of Reference for Language Learning (henceforward, CEFR), introduced in 2001.

As a general framework for language learning in Europe, the CEFR laid the foundation for the curriculum for all foreign language subjects in LK06 (Langseth, 2009) and competence aims and some of the wording in LK20 (Speitz, 2020). Speitz (2020) argues that knowledge about the CEFR contributes to a better understanding of the Norwegian subject curricula for English. In the original Framework, it was described as a "common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe" and a "common basis for the explicit description of objectives, content and methods" (Council of Europe, 2001, p. 1). In a 2020 Companion volume, it is pointed out that "[i]t has never been the intention that the CEFR should be used to justify a gate-keeping function of assessment instruments (Council of Europe, 2020, p. 11), but rather to provide a metalanguage for discussing the complexity of language proficiency and reflect on transparent and coherent learning objectives and outcomes (Council of Europe, 2020). The extended CEFR descriptors in the 2020 Companion volume replaces the original descriptors (Council of Europe, 2020). For the purpose of this thesis, the main interest is in the CEFR's comprehensive descriptors of competence. One major difference is in the move from describing competence in tables to detailing constructs belonging to different competences using figures.

It was pointed out in the 2001 Reference document that the original descriptors would need to be based on theories of language competence, although "[t]his is difficult to achieve because the available theory and research is inadequate to provide a basis for

such a description" (Council of Europe, 2001, p. 21). Langseth (2009) points out that the fact that the descriptors are not based on research, but rather on classroom experience, is a weakness. Another weakness highlighted by Langseth (2009) is that linguistic quality and quantity is described one-dimensionally, which does not reflect some pupils' reality where linguistic correctness is not achieved, yet they communicate very well in many areas. A third is that the scale is too wide and that it can take a disproportionate time to advance from one level to the next, potentially resulting in a reduction in motivation. Furthermore, details are not always provided for the full A1–C2 scale and might say "As C1" for the C2 level, or similar. What the descriptors do provide teachers and policymakers with, however, is reference material which can be used to set learning goals, develop courses, modules and language curricula (Speitz, 2020). Furthermore, the Council of Europe points out that the main function of descriptors is to help align curriculum, teaching and assessment, and insists they are only "one source for the development of standards … they are not in themselves offered as standards. They are a basis for reflection, discussion and further action … thus reference tools" (Council of Europe, 2020, p. 41). Ultimately, the benefit of these descriptors for this thesis is as an overview of competences and constructs which may be included in an understanding of relevant constructs for the English subject.

The 2020 updated framework includes comprehensive descriptors for reception, production, interaction and mediation activities and strategies, as well as plurilingual and pluricultural competences, communicative language competences, and signing competences (Council of Europe, 2020). The 2001 original self-assessment grids were related to understanding (listening, reading), speaking (spoken interaction and production), and writing (Council of Europe, 2001), and is now updated to reception (oral comprehension, reading comprehension), production (spoken, written), interaction (oral, written and online), mediation (mediating a text, mediating concepts, mediating communication). In the 2001 Reference document, common reference levels for qualitative aspects of spoken language use (range, accuracy, fluency, interaction, and coherence) were also provided (Council of Europe, 2001, p. 28–29), and later updated to include phonology (Council of Europe, 2020, p. 183–185). In the below table, I present the competences and related activities and strategies included in the Companion volume (Council of Europe, 2020).

Table 3. Competences, and related activities and strategies, based on the Council of Europe's comprehensive descriptors (Council of Europe, 2020)

Competence	Activities	Strategies
Reception (from p. 47 onwards in Council of Europe, 2020)	Oral comprehension (e.g. understanding conversation between other people) Audio-visual comprehension (watching TV, film and video) Reading comprehension (e.g. reading for information and argument)	Identifying cues and inferring
Production (from p. 61 onwards)	Oral production (e.g. sustained monologue: putting a case) Written production (e.g. creative writing)	Planning Compensating Monitoring and repair
Interaction (from p. 71)	Oral interaction (e.g. goal-oriented co-operation) Written interaction (e.g. correspondence) Online interaction (e.g. online conversation and discussion)	Turn-taking Co-operating Asking for clarification
Mediation (from p. 90)	Mediating a text (e.g. analysis and criticism of creative texts) Mediating concepts ➢ Collaborating in a group (e.g. collaborating to construct meaning) ➢ Leading group work (e.g. encouraging conceptual talk) Mediating communication (e.g. facilitating communication in delicate situations and disagreements)	Strategies to explain a new concept (e.g. adapting language) Strategies to simplify a text (e.g. amplifying a dense text)
Plurilingual and pluricultural competences (from p. 123)	Building on pluricultural repertoire Pluricultural comprehension Building on plurilingual repertoire	-
Communicative language competences (from p. 129)	Linguistic competence (e.g. vocabulary range) Sociolinguistic competence (sociolinguistic appropriateness) Pragmatic competence (e.g. coherence and cohesion)	-

If we are to understand the term *construct* as an abstract noun, such as *fluency* (Fulcher & Davidson, 2007, in Bøhn, 2015, p. 2), constructs should be interpreted as the examples in parenthesis in the above table. This means that the constructs connected to for example linguistic competence (activity connected to communicative language competences) can be understood as *general linguistic range, vocabulary range, grammatical accuracy, vocabulary control, phonological control, and orthographic* control within the CEFR framework. While I will rely on this understanding for the Discussion chapter, I will refer to the direct level below competence, for example *written interaction,* as constructs in the remainder of this chapter. This adds more context than more comprehensive alternatives, like *interaction*, but less context than *correspondence*[15].

Hasselgreen & Ørevik (2020, p. 384) emphasise that the characteristics of goal achievement for the English subject are intended to be used to set grades as well as for on-going classroom assessment. Such guidelines existed prior to the introduction of LK20, although these presented indicators for oral and written competence and have now been deleted from the Directorate's website. When they were first introduced, one intention was for them to be a tool for teachers when assessing pupils' overall competence, and it was noted that too detailed characteristics of goal achievement could lead to fragmentation. Nevertheless, it was stated that they would act more as support for summative than formative assessment (Norwegian Directorate for Education and Training, 2009, p. 15). In the below table (my translation), I have analysed the current characteristics of goal achievement reflecting good competence (3–4) in 10th grade by

---

[15] The CEFR use neither the term *construct* nor *criterion*, so there is no precedent.

what constructs they reflect, based on the mentioned understanding of the term construct.

Table 4. Characteristics of goal achievement – good – for 10th grade

Good competence in the subject	Construct interpretation
The pupil understands and communicates content from different types of oral and written texts in a mainly relevant way	Oral, reading, and audio-visual comprehension Oral and written production Linguistic competence Pragmatic competence
The pupil expresses themselves with clear language and coherence and is mostly able to adapt language and structure to purpose, receiver/audience and situation in different types of oral and written texts	Oral and written production Linguistic competence Sociolinguistic competence Pragmatic competence Strategies to explain a new concept
The pupil finds relevant information from several English-speaking sources, uses this for oral and written text production and refers to the sources	Mediating a text Oral and written production Strategies to simplify a text *Digital competence / source criticism*
The pupil processes parts of their own oral and written texts following concrete feedback	Mediating a text Strategies to simplify a text
The pupil shows intercultural competence in oral and/or written texts by explaining and reflecting over ways of living, thinking and keeping traditions, and the cultural diversity in the English-speaking world	Sociolinguistic competence Oral and written production *Intercultural competence*

These characteristics of goal achievement notably have no indicators relating to interaction or plurilingual constructs and offer little detail, with the exception of its explanation of intercultural competence in practice. Hasselgreen & Ørevik (2020) point out that the subject curriculum for English is "non-specific in nature" (p. 384), and that it leaves much interpretation onto teachers and exam committees. This is also the case for these characteristics of goal achievement, which means it is up to teachers to operationalise for assessment purposes. This widely accepted principle regarding local autonomy has nevertheless led to great differences in how teachers assess (Nordahl & Hansen, 2011, p. 31). How different teachers interpret constructs, competence aims, subject-specific values, and core elements could thus cause differences in assessment. This is something this thesis aims to shed a light on.

3 Methodology, Methods & Material

This research project aims to be a helpful addition in understanding how some Lower Secondary English teachers have experienced the transition from two grades to one, and in an attempt to do so I have conducted a total of 11 interviews with 4 teachers in the 21/22 school year, complemented by an online questionnaire distributed to Lower Secondary English teachers across Norway. In this chapter, I explain how the study was conducted, including the philosophical foundation, approaches and methods used as well as a discussion of the study's validity and reliability. Finally, I present ethical considerations and limitations of the research project.

3.1 Philosophical Foundation

Because qualitative research is concerned with building "a complex, holistic picture" and "reports detailed views of informants" (Creswell, 1998, p. 15), it is vital to reflect on the relationship between objective knowledge and perception. The data material in this study covers individual experiences and ideas, which can neither be made generalised nor seen as objective truth. According to Kant, the only thing we can say for certain is how we experience a phenomenon; we can never with full certainty say that how we study an object, truly reflects reality (Postholm & Jacobsen, 2018, p. 49). Furthermore, within constructivist theory, individual *constructions* can be explained as "an understanding of the world, building on biological, social and individual factors, an ability or an inclination to select and interpret the world" (Steffensen, 2003, p. 232, my translation). For this thesis, which looks at individual views on English subject didactics in light of a structural change in assessment, the perspectives referred to in the study reflect their individual views only – these individual teachers' perspectives, which are indeed formed by how they understand and interpret the world, including their social and individual experiences.

If we imagine a scale with objective knowledge on one side and subjective construction on the other, Bo Steffensen (2003, p. 14–15) argues that there is a middle field where there is a space for value-based choices: the area of subject didactics – where one is free to reflect, explain, justify and choose a common Bildung[16] via the chosen content in public education (Steffensen, 2003). This area of the scale will also include a range of different attitudes to and views of what the subject should and does contain (subjective experience), even though it is to some extent defined in the LK20 curriculum ('objective' knowledge). It is this area of the scale that I attempt to gain insights into in this thesis, as English teachers in Lower Secondary School transition from providing pupils with two grades to one grade. It translates to questions of operationalising of the core and subject curriculum, subjective experiences, and views and attitudes to the subject and its content, and the curriculum change.

This is a phenomenological study, attempting to study and describe the world the way it is experienced, directly and immediately (Kvarv, 2014). Phenomenology requires us to break down the habits in how to view and think about phenomena, and rather view things in the first-person perspective, with as few prejudices or reservations as possible (Kvarv, 2014). This should not be confused with a need to approach research without

---

[16] Danning; all-round development

background knowledge, as thorough insight into the chosen topic can be highly beneficial (Brinkmann & Tanggaard, 2012). I elaborate further on phenomenological analysis in 3.4.

3.2 Interviews

This research project is guided by the following research question: How do teachers experience transitioning from two grades to one in Lower Secondary English? To answer this, I found it most relevant to gather data through qualitative research, specifically interviews. The interview is a regularly used technique to gather, analyse and interpret data and information, and is particularly well suited to observe attitudes. We differentiate between *informant interviews* and *respondent interviews*, where the first refers to interviews where people are asked about their knowledge about a phenomenon and not their interpretations, whereas the latter refers to interviews where the focal point is the respondent's personal feelings, perceptions, and intentions (Kvarv, 2014). This study belongs to the latter category, as the research question is concerned with the teachers' *experience* of a phenomenon, not the phenomenon in itself. The phenomenological interview serves this purpose well, as it involves paying attention to 'what' and 'how'; what is consciously experienced, and in what context is the phenomenon experienced (Postholm & Jacobsen, 2018).

This study followed four teachers throughout the 2021/22 school year, and interviews were conducted in three rounds, meaning this is a repeated cross-sectional qualitative study. Three teachers were interviewed three times, whereas the fourth ("Peter") was interviewed in January and June 2022 only[17]. The interviews in the study were semi-structured, loosely following a thorough interview guide (Appendix E). A semi-structured interview guide creates an opportunity to follow the respondents' lead (Postholm & Jacobsen, 2018), which is something I wanted to accommodate further by preparing a longer list of questions to choose from, depending on what topics the respondents showed more interest in covering. When creating the interview questions, I attempted to follow the advice of Brinkmann and Tanggaard (2012, p. 30) to formulate the questions clearly and close to their reality, which meant paying attention the language they used as well as leaning on less academic terms, but rather use words close to the classroom reality (e.g. comprehension, or listening and reading, rather than reception skills). By collecting data at three different time points, this allowed for their views to change and for me to follow their experiences in real time. This created an opportunity for the respondents to talk about thoughts and ideas in the early stages of the school year, and for me to follow up later to see what they ended up doing and how they experienced different elements of teaching and assessment over time. This also accommodated for the collection of data in cases where participants had changed their minds over time, which could potentially highlight interesting discussion points and contribute to a more extensive understanding of the respondents' individual experiences.

Because of the Covid-19 pandemic, the interviews were not conducted face to face as intended, but rather digitally via Zoom. The interviews lasted between 40 mins and 1h 20 mins, totalling approximately 11 hours and 30 minutes. Conducting interviews in person would have been preferable to meeting digitally. One concern I had was regarding turn-taking and whether we would be able to engage in natural conversation, or if it would have a detrimental effect on our communication or developing relationships. In

---

[17] The first interview with this participant included questions from the first interview round to ensure data on important questions.

retrospect, I wish I had paid more attention to the accepting longer pauses and waiting for responses, rather than opting for strategies such as elaborating on the question or asking if they wanted me to rephrase. This is because pauses can be productive and lead to richer interviews (Gleiss & Sæther, 2021, p. 90). The more methodological insight I have gained over the course of this research project, the more novice errors I am able to identify. Nevertheless, conducting interviews via Zoom did give the respondents more flexibility in finding a time that worked with their busy schedule, and for both ends to easily reschedule when necessary.

The goal of the first interview round in Autumn 2020 was to gain insight into the respondents' background, experiences, and initial thoughts about the transition from two grades to one grade. The questions were very open at this point. After the first interview, I realised I had to reflect more about how to guide the conversation in case I had more responses where there was seemingly not as much to discuss as if one has a strong opinion about a topic. How could I gather meaningful data about the change if teachers reflected that the curriculum change would not have much of an effect on teacher behaviour, and that rather, things would mostly stay the same with no change to assessment practice? Could this be interpreted as positive change, and if so, how could I get the teacher to elaborate, and how could I make sure I got valid responses without leading the participant towards a hoped-for result? Perhaps my preconceived assumptions showed, more than anything, that I believed that teachers would have a strong opinion about it, either negative or positive. The first interview gave me a chance to reflect on how to gather data in the event that I interviewed a teacher with no strong opinion about it, which was helpful prior to conducting the next interviews.

For the second interview round, in the months of January and February 2022, I focused on their experiences with grading for the preceding term, including term papers[18]. I also asked about their expectations for the Spring term. This interview guide included a large number of questions to allow for a selection as I saw best fit under individual circumstances, without necessarily asking every question to all respondents. We did cover all important aspects but also focused more in-depth on their experiences which became apparent were of more importance or relevance to them. This was also the case for the third interview round, which took place in June 2022, immediately or shortly after the term grading, and for the respondent who taught 10th grade, the pupils' final grading in Lower Secondary School. Questions regarding the Spring term as well as their overall experiences were relevant, as well as any questions I had asked some in the 2nd round which I now wanted to elaborate on with others (all were included in the interview guide and selected where appropriate).

---

[18] Tentamen

Table 5. Overview of interview rounds, with corresponding goals

Interview Round	Goal
Interview Round 1 Autumn 2021	To gain insight into background, current assessment practices and school culture, and initial experiences and thoughts on the change
Interview Round 2 January/February 2022	To uncover experiences with grading in the Autumn 2021 term, and reflections on change, collaboration, and constructs in the English subject
Interview Round 3 June 2022	To uncover experiences with grading in the Spring 2022 term, and gain insight into real-life grading examples from the school year Access final remarks

Developing the interview guides helped not only in the actual interviews, but also in the selection of background information and theoretical foundation as I had to ask myself what I wanted to know about the topic, which included questions no respondents would be able to answer. Consequently, it helped me conduct thorough background research and led me to the path of sending questions to the Directorate as well. With that in mind, this thesis could, with certain adjustments, have bordered to the qualitative evaluation methodology (bottom-up oriented evaluation), where the goal is to systematically evaluate the outcome, results and organising with a goal to impact practical action (Dahler Larsen, 2012). However, I view the chosen approach as more appropriate for an master's thesis, in addition to being particularly interesting as it highlights teachers' views on change as it happens.

Table 6. Examples of questions from the interview guides, and background for asking

Example of questions	Reason for asking
How have you discussed assessment practice following the transition to one grade? Who initiates change and development in the English subject? (Interview Round 1)	To access information about the teachers' school culture and systems put in place to facilitate interpretive communities about assessment practice, assessment culture, and changes related to the LK20 curriculum
How have you approached assessing 'overall competence' compared to previously giving an oral and a written grade? (IR2)	To gain insight into the individual teacher's process of assessing pupils' competence, with comparison to previously providing two grades
If pupils now hypothetically were to formally complain about the grade and wanted feedback on your assessment: How do you think you would have presented the weighting of constructs and 'the overall competence'? (IR2)	To understand how the teachers interpret what should be assessed, in light of the LK20 curriculum changes and the transition to one grade. The question is asked in a way which requires the teachers to think practically about something theoretical.
Can you give some examples of pupils who have been 'between grades', and what you emphasised in the final assessment? (IR3)	To gather real-life examples of pupils who might have received a different result with two grades, and to understand what the teachers emphasise in those assessment situations.
If you imagine an assessment scale from "thinking holistically" to "assessing separate skills": Where are you on this scale, and what are your reasons? (IR3)	To facilitate reflection about the individual teacher's ideas about assessment in the English subject.

34

Interviews should ideally resemble, or follow similar structures to, ordinary conversations (Brinkmann & Tanggaard, 2012; Kvarv, 2014). I found it increasingly challenging to balance this wish for the interview to resemble natural conversation with keeping my presence at a necessary distance and not enter 'collegial' conversation rather than an interview. When I piloted the first interview round questions with a fellow student, and later a teacher, I had not paid enough attention to how I would actually speak in the interview, but rather focused on finding out what interview questions worked and whether I needed to change the wording. As a consequence, perhaps, I was not sufficiently prepared for the human factor of wanting to build a relationship in the conversation. For this purpose, the thorough interview guides were a good help. During the first interview with one of the respondents, I experienced that they changed from being initially very negative to being somewhat positive. This, of course, made me concerned that I had somehow contributed to the change in opinion. However, because I had planned a series of interviews, this gave me the chance to discuss this retrospectively and ask why she thought she had changed her mind. The longitudinal aspect of the study was helpful in other ways as well, as the school year is often filled with great intentions and time restrains in equal measures, and by following the teachers over the school year, they could inform me what came into fruition and what did not. The benefit of not having a cross-sectional study, but rather opting for a longitudinal research design, has been the additional data this has provided. For example, I would not have been able to follow the above-mentioned teacher's change of opinion and gain insight into their reflections over time. Another teacher discussed how they wanted to introduce an assessment task design later in the year, but ultimately, did not have the time to implement it. If I had only conducted one interview per teacher, I would not have had such rich and relevant data.

Another benefit from choosing a longitudinal approach is that I was able to reflect on my role as a researcher in between interview rounds, including how I could ask better questions. When I read the transcripts from the first interview round, I noticed that I would occasionally formulate topic-introductory questions with a yes/no question, e.g. "Have you talked about what competences you will emphasise?". In retrospect, I have reflected that this is a strategy I often use in informal conversation, and that my subconscious goal is to safely introduce a topic before engaging in more open questions. This is also revealed by my subsequent follow-up questions, which are typically more open 'what'/'how' questions, often chosen from the interview guide. Ultimately, when comparing the transcripts from the first interview round with the subsequent interview rounds, I had reduced the use of this strategy and relied more on open-ended questions.

This inexperience with research was also revealed to me when I was attempting to understand, or perhaps simultaneously interpret and engage with their responses. I tried to avoid formulating leading questions, as I had a genuine wish to get their truthful reflections, but I was concerned that the respondents would interpret these comments or follow-up questions as indeed leading comments and questions. One example is from my first interview with one respondent: "I interpret your answers … that there has been no guidelines", upon which they answered "You can safely interpret that, yes". Of course, my comment did not come out of nowhere and reflected perhaps a wish to clarify or get additional comments on something I felt I could safely assume from their previous response. According to Easterby-Smith et. al. (1991, p. 73), "the researcher will need to be sensitive enough, and skilled enough, to ensure that she not only understands the other person's views, but also, at times, assists individuals to explore their own beliefs". I

have found it challenging to balance these two somewhat conflicting needs, especially when I sensed insecurity in their responses.

3.2.1 Selecting Respondents

The essential selection criteria for participants partaking in phenomenological interviews is to have experienced the phenomenon being researched, recruiting typically between 3 and 25 participants (Postholm & Jacobsen, 2018, p. 118). This should be a heterogenous group (Creswell, 1998), but the context should ideally remain the same for all participants (Postholm & Jacobsen, 2018). The essential selection criteria for this study were for potential respondents to be Lower Secondary English teachers currently transitioning from two grades to one. To achieve a certain heterogenous selection, I wanted respondents representing 8th, 9th and 10th grade, although one could argue that I could have recruited teachers with experience from one grade only to further limit the context and receive responses which could more easily be compared within that one context. By broadening the selection criteria, I would potentially allow for more disparity in responses, but the benefit is potentially being able to show broader nuances in the respondents' backgrounds, which gave a greater potential for increased data material as the experiences will differ. For example, a 10th grade teacher will deal with final grading, whereas an 8th grade teacher works with pupils who are completely new to receiving grades. Heterogeneity was also covered through a 50/50 gender balance, age groups ranging from the 30s to the 50s, and years of teaching experience (from <5 to >25 years), as well as size of class and school (ranging from 200 to 500 pupils). Furthermore, the teachers are recruited from three municipalities across the country.

Table 7. Presentation of interview participants

Respondents (pseudonyms)	2021/22 School Year	Thesis-specific context
«Lisbeth»	Teaches 10th grade	Second year providing one grade
«Ragnhild»	Teaches 8th grade	Second year providing one grade
«Markus»	Teaches 9th grade	First time providing one grade
«Peter»	Teaches 8th grade	First time providing one grade

I have limited the information about each respondent to the essential minimum to ensure their anonymity is maintained (see more in 3.6). Henceforward, I will refer to the participants by their pseudonyms.

This group of respondents have relatively homogenous classes in terms of ethnicity, and two teachers also express having quite a homogenous group with regards to the general level in the class, explaining this as very high. There will be other experiences to draw on with more heterogenous pupil populations and lower general skill levels or interest in the subject, which is a weakness of this study as it has not collected data to portray teachers or pupils who could potentially reflect the most challenging aspects of the change. However, the respondents reflected thoroughly on how they imagined it would affect a range of pupils and drew on their own experiences with pupils who would benefit from having two grades. Additionally, Postholm & Jacobsen (2018, p. 118) argues that large disparities in contexts cause a challenge for the researcher in establishing an essence in

phenomenological studies. Consequently, this may equally be an unplanned benefit as well as a potential weakness, depending on perspective.

3.3 Questionnaire

Following the third round of interviews in June 2022, I distributed a digital questionnaire to complement the qualitative findings with quantitative data. At this point, all teachers had finalised the grading process for the Spring term. The idea behind this was to position the interview data in a broader perspective, and I explain this in more detail in 3.5 with regards to validity and reliability. Furthermore, I detail how I engaged analytically with the data in 3.4.2. Rather than constituting a mixed-method study, the quantitative element serves as a helpful tool to add perspective to the qualitative data.

The questionnaire (Appendix F) was named "Transitioning from two grades to one in Lower Secondary English" and was distributed in two versions in Nettskjema: one English, and one Norwegian. I wrote the questions in Norwegian first before I translated it close to the source text for an English version for whoever would be more comfortable responding in English. My main reason for this was not wanting to potentially exclude teachers whose first language is not Norwegian, but there could also be other reasons for preferring to respond in English. The questionnaire was distributed twice to a closed Facebook group[19] consisting of approximately 20 000 teachers of English across Grunnskole and Upper Secondary levels, as well as a selection of teachers, principals, and school librarians across the country.

The questionnaire consisted of 24 closed questions detailing experiences and thoughts regarding the transition from two grades to one grade. Some shared resemblance to questions from the interview guides, and others were inspired by the data gathered in interviews. The questions varied from simple questions about background and personal opinions to collegial collaboration, the use of supportive material, and experiences.

In the end, I received 87 responses on the Norwegian version and 28 responses on the English questionnaire – in total 115 responses. This is of course not representative for English teachers, nor is my aim to generalise. Nevertheless, Thrane (2018, p. 121) establishes that 120 is a large number of quantitative observations, and that it is the variation of the population, not the size, which defines the sample size requirements if precision is the aim. Had I required the sample to be representative for teachers in order to generalise, I would have distributed more systematically by contacting more schools and structuring the quantitative analysis more thoroughly than I did, for example by using SPSS. Again, I will stress that the quantitative element does not serve a function to represent or generalise the opinions of Lower Secondary English teachers, but to position the qualitative data in a broader perspective for validity purposes. Nevertheless, this sample of 115 teacher responses is representative of those 115 Lower Secondary English teachers, and these teachers' responses are important in their own right. These findings are presented in 4.5 and later used in the discussion to add context to the qualitative findings. Examples of questions along with background for asking is found in the table below.

---

[19] The group "Engelsklærere": https://www.facebook.com/groups/1434343746839024

Table 8. Examples of questionnaire questions, and background for asking

Example of questions	Reason for asking
Have you changed the way you assess as a result of transitioning to one grade (reflecting overall competence) rather than two (oral skills + written skills)? On a scale from 1 (no change in assessment practice) to 5 (full restructure of assessment practice)	To gain information about whether teachers experience that they had made changes to their assessment practice following the curriculum change. A corresponding question was asked about change in teaching, as teaching and assessment are ultimately linked.
If you responded 2 or more in the two previous questions, what have you changed? Choose all that apply.

(16 options. Examples include More/less focus on written skills / oral skills / listening / understanding) | This followed the above question and was chosen to gain insight into the results of teachers' experience with changing their assessment and/or teaching practice following the subject curriculum change. I wanted to find out what teachers found to have increased or decreased in focus. One option included 'Other'. |
| On a scale from 1 to 5, how much do you agree with the following statement? "It is easier to assess with one grade versus two grades" (1 = I do not agree. 5 = I fully agree.) | Simple statements such as this one were included to see if it was possible to get "simple answers" to what is perceived as complex. If such questions received mixed feedback, I wondered if I could conclude that it is indeed complex, or if it is possible to simplify the curriculum change. |
| What is your general experience with assessing people with a disproportionate skill set? For example, a pupil you would give a 6 in oral English and a 3 in written English with the old system. | Gain general insight into whether teachers experience that the change in grading is beneficial, similar or negative for pupils. The participants were able to respond that "My experience is that the change is entirely/mainly positive/negative" or "Similar, but I have to work more/less on adapting to their individual needs". Options also included No change and Not sure. |

3.4 Data Analysis

The data material consists of transcriptions of 11 interviews conducted with four teachers over the course of the 2021/22 school year. The initial interviews tended to be shorter, as these covered their initial thoughts whereas the subsequent interviews in the second and third interview round consisted of more in-depth conversation. This chapter has already established the research project as a phenomenological study. This section goes further in elaborating on the consequences this framework has for my chosen approach to data analysis. I begin by describing the chosen approach, interpretative phenomenological analysis (henceforward, IPA) , before I detail the procedure of analysis.

3.4.1 Interpretative Phenomenological Analysis

The goal of phenomenological analysis is to describe a phenomenon thoroughly "in all its subtlety and rich layers" (Finlay, 2014, p. 122) and to explicate the lived experience holistically (p. 136). According to van Manen (1997 in Normann, 2017, p. 615), phenomenological research can be compared to "borrowing other people's experiences and their reflections on their experiences". By choosing a phenomenological approach to analysing qualitative data material, the focus is thus on identifying *what* individuals experience and *how* they experience the phenomenon (Normann, 2017).

At its core, Finlay (2014, p. 122) distinguishes phenomenology from other research approaches focused on exploring subjective experience in its *radical phenomenological attitude*, which van Manen and van Manen (2021) establish is characterised by application or transformation of Husserl's method of reduction and epoché. This attitude requires a move away from personal experience, established knowledge, prejudices and

reservations (Kvarv, 2014; Finlay, 2014). This refers to one's own pre-understandings, frameworks and biases in search for genuine openness, and requires overcoming subjective feelings and expectations (van Manen, 2011, in Finlay, 2014, p. 123). We do so through what is known as *bracketing*, or *epoché*, which means setting aside any taken-for-granted or intuition-based knowledge in order to see the essence of the experience (Finlay, 2014; Creswell, 1998). This is perhaps connected to 'critical subjectivity', i.e. recognising one's own views, but not allowing oneself to be overwhelmed and swept along by them (Reason, 1988, in Easterby-Smith et. al., 1991, p. 39). Although this bracketing method is essential in Husserl's descriptive phenomenological approach, it is not typically advocated for in IPA (Normann, 2017). Nevertheless, I chose to engage with my personal experiences receiving one grade as a pupil in Upper Secondary and any pre-conceived notions and assumptions I may have had, and it proved to be a very helpful exercise to develop a more phenomenological attitude when engaging with the respondents and the subsequent data material. This bracketing exercise is detailed in 3.5.1. My experience is that by making my personal experiences and pre-suppositions explicit, I was distancing myself from my perspective and simultaneously building internal curiosity about potential alternative experiences.

A second distinction between the two approaches is in its use of theory. While descriptive phenomenology rejects theoretical frameworks as a help to focus the inquiry, researchers following an IPA approach do in fact often lean on a theoretical framework when interpreting the material (Normann, 2017). While I did engage substantially with background research and gaining a theoretical foundation early on in the process, I did not make the selection for the thesis until the analysis process had ended. In hindsight, I have reflected that it would have been beneficial to lean on this particular theoretical framework prior to conducting interviews and indeed in the analysis process, but ultimately, the selection is made in dialogue with findings as the narrative develops. Furthermore, while Husserl and descriptive phenomenologists speak of *essence,* researchers following an interpretative phenomenological approach are less concerned with essence, and rather look for patterns to build a thematic structure, abstracting multiple themes and identifying a few superordinate and/or sub-themes (Finlay, 2014, p. 131). This path is described in 3.4.2.

The interpretative phenomenological framework is often adopted when a researcher's prior knowledge and insights can be applied to interpret hidden meaning (Normann, 2017). By building on a personal account of receiving one grade, resulting in certain pre-suppositions, while gradually developing an understanding of the background and theoretical foundation, it became natural to steer towards IPA. According to Smith & Osborn (2003, p. 66), meaning is not transparently available, but "must be obtained through a sustained engagement with the text and a process of interpretation". This is in line with the hermeneutical tradition, where analysis is viewed as an interpretation process where the researcher investigates the relationship between individual parts and the big picture: seeing parts of the whole in relation to each other and gradually developing a nuanced understanding of the whole (Gleiss & Sæther, 2021, p. 170). In relation to using prior insights to interpret (hidden) meaning, my experience has indeed been that my previous knowledge has been helpful in both the interview and the interpretation process. In the interviews, I would occasionally use background knowledge as a way to encourage further reflection, and in my opinion, this opened for critical reflection on the reasons for their opinions, or more detailed engagement with their own experiences.

Finally, a phenomenological analysis requires what Finlay cites as "fidelity to the phenomenon as distinct from mere conceptual thematizing" (2014, p. 122). In my understanding, this means committing to finding out as much as possible about as many aspects of the phenomenon as possible, which requires a genuine curiosity and openness. I believe I have maintained this throughout the interview and interpretation process, and have attempted to present an analysis in line with the aim of IPA, which is to "say something in detail about the perceptions and understandings of this particular group rather than prematurely make more general claims" (Smith & Osborn, 2007, p. 55).

3.4.2 Procedure of Analysis

Phenomenological analysis typically follows three steps: 1) Read to get a sense of the big picture, 2) Develop meaning units, and 3) transform participants' responses to psychological phenomenological expressions (Giorgi, 1985, in Postholm & Jacobsen, 2018, p. 160). There are ultimately many ways to transform this into practice (Finlay, 2014; Smith & Osborn, 2003; Normann, 2017), and Normann (2017) argues that it is important that novice researchers do not fall into the trap of depending too much on established structures. Smith & Osborn (2003) present a stepwise procedure for interpretative phenomenological analysis (IPA), but insists it is not a prescriptive methodology, but rather "a way of doing IPA … to be adapted by researchers" (p. 67). In the analysis process, I used this procedure as my main inspiration. Although I would also lend aspects from other analysis approaches, I would remind myself to always keep the focus on experiences related to the phenomenon, even when the respondents had discussed interesting topics unrelated to the research question.

The first step of Smith & Osborn's (2003) procedure for IPA relates to reading the transcript multiple times, before starting the written interpretation process which involves initial noting of exploratory comments. Smith & Osborn (2003) recommends using the left-hand margin to annotate significant information from the responses, then the right-hand margin to document emerging themes, i.e. concise phrases which transforms the data material to capture the essential quality of findings (p. 68). The former can include for example summarising, paraphrasing, associations, connections, comments on language, and the noting is likened to a free textual analysis. For this thesis, I found myself noting down a variety of words, from single words such as 'LK20', referring to the curriculum, and short explanations, like 'benefit of the doubt', to longer summaries. Examples include 'Example: Pupil feeling anxious about oral production – solution: Audio files. Result: Can assess 'how they speak, vocabulary, grammar, everything'' and 'Class dynamic: Mainly energetic. Responsive. Wide range which includes pupils who do not participate". At this point, I also tried to establish a narrative about the individual teachers' background and classroom experience, to differentiate the human account from simple words on paper.

The third step is to develop emerging themes. It is recommended to start with an initial chronological list, while later processing this to connect themes to find similarities or develop superordinate themes. This is referred to as searching for connections across emergent themes (Smith & Osborn, 2003). In Smith & Osborn's (2003) account of IPA, they recommend producing a table of themes, with identifiers to indicate page numbers and key words to be able to identify the source for the theme, but I did not add such indicators in this process. I had to develop trust in my ability to extract meaning from the transcripts while simultaneously simplifying the process in order not to overwhelm myself with the substantial material. Rather, I collected the relevant quotes in Word documents to use for the later narrative account.

The above steps were then repeated for each individual transcript. Smith & Osborn (2003) establishes that researchers can use either emerging themes from previous transcripts or start over each time. I chose to analyse each individual transcript from scratch, as is recommended by Smith & Osborn (2003, p. 74) for a student's first IPA project. However, I kept the list(s) available so that I could remind myself of the wording if I felt it was needed. This was another way of simplifying the process due to the substantial material I had to analyse and meant I could attempt to use similar wording.

Once I had developed themes for each individual transcript, I looked for connections between the transcripts. To do so, I was inspired by what Gleiss & Sæther (2021, p. 176) refers to as code structures, where codes are connected in a structure at different levels, with some being more superordinate. This step helped in clarifying themes and more importantly, identifying similarities and differences in the identified themes and the associated data. This step was done by hand in several rounds and gradually developed from looking like mind maps to more structured 'code trees' with branches going down from what emerged as superordinate themes.

The final step in Smith & Osborn's (2003) account of the IPA approach is the write-up, which is explained as the "final statement outlining the meanings inherent in the participants' experience" and "translating the themes into a narrative account" (Smith & Osborn, 2003, p. 76). This last component meant going back and forth in the transcripts, making sure I reflected the respondents' views appropriately and correctly. In retrospect, it would have been helpful to follow Smith & Osborn's (2003) advice to create a table of themes with reference to page numbers and quotes which created the basis of the themes, but what I had rather done was create a collection of quotes which were essential to build the narrative, and supplement over time when I re-engaged with the transcripts. The write-up process included several rounds of looking over transcripts and other notes, including mind maps/code structures, while looking for connections. Because the initial noting process, development of emerging themes and mind maps were done by hand, this created a challenge while shaping the narrative account on the computer. This could of course also be related to the fact that I had 11 transcripts relating to three time points over a year, and multiple corresponding mind maps and notes. In other words, as a novice researcher I have made some experiences about how this process can be structured differently when dealing with a large data set.

The task of extracting meaning from the data and reducing these to a handful themes and sub-themes has been a time-consuming, but very interesting one, and the reduction process shaped rather naturally as I began looking for similarities and differences both when comparing the individual teacher's answers across interview rounds, and when comparing with the other teachers' responses. At the very end, I engaged with the quantitative data findings to see what additional perspectives this could bring to the thesis. I did so by printing the final responses from both questionnaire versions, identifying what questions were considered relevant in light of the qualitative data, then summarising the responses to the relevant questions. I appreciated having the option of making this element a larger part of the thesis, potentially looking at correlation between responses and age groups, years of experience, etc., but concluded the qualitative data were so rich, there was neither space nor need for a substantial analysis of the quantitative data.

3.5 Validity and Reliability

According to Postholm & Jacobsen (2018, p. 222), research quality depends on a dialogue about the substance of the produced findings, as well as two important factors: the study's limitations, and how the researcher may have impacted the final results. The first factor refers to validity, and the second to reliability, two central terms I will discuss in this section. Limitations are further mentioned in 3.6.

Validity refers to what conclusions a researcher may in reality draw from the collected data (Postholm & Jacobsen, 2018), or the quality of the data material and the researcher's interpretations and conclusions (Gleiss & Sæther, 2021). We distinguish between 'internal validity', which refers to whether the researcher's conclusion are valid for the study in question, and 'external validity', which relates to whether the results can be transferred, or generalised, to other contexts than what is studied (Postholm & Jacobsen, 2018). One relevant question related to the internal validity of this study, is *How certain can we be that something is a 'cause' and something else is 'effect'?* (Postholm & Jacobsen, 2018, p. 223). In this thesis, I have tried to pay attention to causality when collecting and analysing data through interpreting responses and asking what is clearly connected to the curriculum change, and what are more general challenges, or unrelated. The background research has also undoubtedly helped in developing this understanding. A second relevant question is whether the collected data has indeed measured what we say or think we are examining (Postholm & Jacobsen, 2018, p. 223). By constantly reminding myself of the research question, and by attempting to develop a phenomenological attitude as explained in 3.4.1, I have focused my attention to the best of my abilities on creating good opportunities for the research participants to answer detailed about the phenomenon, and later in extracting meaning from the data material which reveals the truthful experiences of these teachers in the transition from two grades to one. In terms of external validity, this is not particularly relevant to this study. If a similar change were to be implemented in other subjects, they would have other contexts. The only other comparable context would be the Norwegian subject, where pupils receive three grades in the final assessment. However, a reduction in the number of term grades has been recently implemented following a long trial period, and the subject context is different in many ways. And ultimately, the chosen research design makes generalisability impossible as it is concerned with individual experiences.

Reliability can be understood as whether results can be replicated by other researchers in other contexts, but a qualitative study will be challenging to replicate for several reasons, including that the participants will be different, and other researchers will bring their subjectivity in their research. For this reason, reliability can also be understood as how the researcher may impact the result, and this requires the researcher to reflect on their personal impact and making the entire research process transparent (Postholm & Jacobsen, 2018). Postholm & Jacobsen (2018) insist that findings in a qualitative study represent contextual knowledge, and that this includes researcher subjectivity. For this reason, I engage reflexively with my role as a researcher and the subjectivity I bring to this study in the next section, 3.5.1. Furthermore, from my understanding of Postholm & Jacobsen (2018), a study's reliability is best evaluated by the readers[20]. This whole

---

[20] See for example Postholm & Jacobsen (2018, p. 224) where it is pointed out that reliability is a result from a dialogue between the researcher and others who are interested in the research.

chapter thus becomes an attempt to make the research process transparent for the reader to judge the study's reliability.

One option to strengthen the validity and reliability of the study is triangulation, which refers to combining different researchers, research designs, data sources, and data collection methods. The intention is to describe reality from different angles to gain a more holistic image of something complex (Postholm & Jacobsen, 2018). This study is not a mixed-methods study, but the decision to include a quantitative element was done indeed to strengthen the validity and gain access to broader data which could complement the qualitative findings. By adding these quantitative data as context, I attempt not to generalise but to understand the qualitative data in light of a greater number of responses indicating answers to similar questions. Additionally, I use a significant number of quotes from the participants to build a true account of the respondents' responses, without interference from me as the researcher by paraphrasing. However, having to translate the material from Norwegian to English does mean there could always be instances where a response is translated in an ambiguous way. In this process, I was able to lean on my extensive experience as a translator and a copywriter since 2015, which has been helpful in reflecting about how to stay close to the source material yet use the English language in a meaningful way.

3.5.1 Positioning myself as a researcher

The decision to investigate the recent change in grading in Lower Secondary English came from curiosity. My own experiences from Upper Secondary contributed to this interest, as I received one grade heavily relying on written work, and where the interpretation of my oral competence was mainly based on listening tasks, answering questions by raising our hands and what was called oral questions in written tests. I was thus aware at an early point that I would not be able to position myself as fully neutral. Nevertheless, I was also very curious to see how it could impact teachers and pupils in a positive way, as it has been portrayed in a mainly positive light in the limited attention it has received.

Wishing to be transparent, I wrote down initial reflections at an early point and shared these with my supervisor. In my chosen analysis approach, IPA, such a process is not typically adopted (Normann, 2017, p. 615–616). However, I found this to be both helpful and necessary in order to get a fresh perspective on my own experiences and see if I could use this as a resource in order to develop a path towards finding out what to read, and how to prepare for and develop guides for the interviews to follow. Finlay (2014) recommends beginning the 'bracketing' process by making a list of assumptions, expectations and hopes for the findings, and discussing these with a supervisor to "consider reflexively how such presuppositions might impact on the research" (Finlay, 2014, p. 136). This writing exercise was an attempt to objectively reflect on possible positive and negative outcomes, and this early record of assumptions has been helpful at later stages of the study, as it has given me the chance to go back and see my initial ideas in light of the collected data. This writing exercise could have been the starting point for using a hypothetical deductive method, i.e. establishing a theory about a phenomenon and then testing hypotheses when collecting data (Thurén, 2012). However, it did not play a big part in the preliminary phase and acted primarily as an entry to background research. Rather, this study had an *inductive* approach, which is characterised by collecting data, then looking for patterns in the data material. The significant contrast is whether the researcher starts with data or theory: Inductive research goes from the specific (in my case, qualitative interview data) to general

information (attempting to conclude based on the data). Although I did have a written account for ideas and assumptions, I did not shape them as hypotheses to be tested or expected outcomes. Instead, it served the purpose of brainstorming.

My starting point, including my experiences and preconceived beliefs and opinions, began to take a backseat as I wrote down these assumptions and later started collecting background information. By reading hearing responses I was exposed to a range of reflections on the curriculum change. I found the teacher perspective in the hearings showed a great difference of opinion, and several concerns and optimistic views that reflected other perspectives than mine. Once I had a good overview of different perspectives, I found it surprisingly easy to move away from my personal starting point, which I was very aware was simply one experience, defined by one teacher's learning beliefs and a range of other factors. I believe all this work collecting background data, initiated by a personal experience but also a deep wish to be objective, has helped me in creating a healthy distance to my own experience and preconceived beliefs, towards a goal to focus on collecting broad and interesting data based on respondents' experiences and perspectives.

The subsequent data analysis process inevitably involves an element of subjective interpretation (Gleiss & Sæther, 2021), and a researcher will never become entirely objective about their own subjectivity (Heshusius, 1994, in Postholm & Jacobsen, 2018, p. 220). If a researcher becomes aware of their subjectivity, it can be presented as a part of the context, for which the findings are understood within (Merriam, 2002, in Postholm & Jacobsen, 2018, p. 220). In my attempt to be transparent about my own experience and use it as a resource rather than a hindrance, I found that it provided me with a retrospective pupil perspective of receiving one grade and finding it unfair, and I experienced this empathy for pupils in a similar position as a helpful factor. When moving on from this personal experience via thorough background research onto the teacher perspective, I experienced that the respondents quickly shared similar concerns without me being required to introduce it first.

According to Brinkmann & Tanggard (2012), the best interviews are conducted by researchers with an extensive knowledge about the topic investigated. Perhaps the fact that my own experiences were so impactful (and even headed me towards an English teaching degree) was helpful in how I continuously attempted to stay alert on researcher bias. Ultimately, it made me more curious of what other experiences and views there could be, and hopeful to be able to portray nuances in the study. I have been driven by wanting to portray, indeed, nuances, to show that the change from two grades to one grade is a rather complex exercise.

This extensive knowledge about the topic creates a challenge in defining my researcher positioning as an *insider* or an *outsider*. Although Postholm & Jacobsen (2018, p. 245) emphasises that a student will always be an *outsider*, an *insider* is explained by Hellaway (2007, p. 484) as "an individual who possesses a priori intimate knowledge of the community and its members", emphasising that "possessing intimate knowledge of it doesn't necessarily mean being a member of it yourself". An *outsider* is defined as the opposite, i.e. a researcher not familiar with the setting or people which are being researched (Hellaway, 2007). Although I am indeed a student, I have obtained knowledge about the teacher community through this thesis as well as previous studies, teaching practice, and conversations with teachers. As I followed the teachers in this study over time and developed a somewhat friendly, professional relationship with them,

I did find it challenging to balance what I identify as at least in part an *insider* view with a more objective researcher perspective. There is also an increased chance that *insider* researchers agree with respondents (Kvernbekk, 2005). Nevertheless, there are also benefits to having extensive knowledge and insider insights into a topic. For example, thorough literature reviews give the researcher an overview that an objective outsider could more easily have overlooked (Brinkmann & Tanggaard, 2012, p. 27).

3.6 Ethical Considerations and Limitations

In the process of working on this research, I have followed the guidelines set by the National Committee for Research Ethics in the Social Sciences and the Humanities (NESH), and the project was approved for data collection by the Norwegian Centre for Research Data (NSD) (See Appendix B). Research ethics consists of a core set of scientific norms: the truth norm, methodological norms, and institutional norms, as well as common norms, i.e. society's expectations to research (NESH, 2022, p. 5). Methodological norms, including factuality, accuracy, transparency, and accountability, is of essence, as it is a researcher's responsibility, regardless of experience, to conduct a study using a methodology and a research design which is appropriate within the research community and to the general society's expectations to research.

In this study, I did not obtain any personal information beyond the four respondents' email addresses and names, ages, years of experience plus limited employer information (name of school and municipality only). The interviews detailed no information about third parties, and the questionnaire was completed in Nettskjema, which collected no identifying data (including no collection of IP addresses) and used only closed questions to leave no chance for participants to include any identifying data.

Participation in research is based on information and consent (NESH, 2022). The participants recruited for the interviews all received an information letter (Appendix C) detailing relevant information about the study, including who to contact, and their rights as participants, e.g. the right to withdraw at any time. Prior to our first interview, they all received and digitally signed a consent form, and upon the first interview, they were asked if they had any questions to the study. Informed consent means that researchers provide sufficient and clear information, including "why they are asked to participate, what type of data is being collected, how it will be used, who will make use of the data, and for which purposes" and "a plan for the processing, storage, use and sharing of data" (NESH, 2022, p. 19). This was all included in the information letter, and the interview data was appropriately saved on an securely stored external device not used for any other purposes, and deleted after the transcription process had finished. Furthermore, the digital interviews were conducted via Zoom (encrypted, using my NTNU credentials) and I used a recording device borrowed from NTNU to record the interviews.

Anonymisation is crucial if the research is to protect the participants' identity and integrity, and means the removal of any connection between individuals and information to prevent information from being traced back to particular individuals. This includes pseudonymisation or the use of another type of key (NESH, 2022, p. 23). With this in mind, there are a few possible weaknesses with the recruitment and communication in this study. Recruitment was made using my own network, known as convenience sampling (Johnson & Christensen, 2012, p. 230). Two people contacted relevant teachers on my behalf in their network. This meant the identity of one person is known to one other person, and the identity of two others are known to one other person, although no other information about their participation has been shared. Nevertheless, for this

reason, I have anonymised the participants to full capacity to avoid any ethical concerns. This includes the anonymisation of gender, age, years of experience, as well as other categories regarded as non-essential. Postholm & Jacobsen (2018, p. 251) explain we can apply a low degree of detailing when we operate with few participants in research to further reduce the possibility to identify the participants. The fourth teacher was recruited by me. Ideally, I would have had a pool of respondents unknown to any other person, however, the respondents this study did recruit, contributed with such diverse data that I am very grateful I ended up with this particular selection of teachers. I am also grateful to the two people helping me recruit such an interesting group of teachers, and trust that the respondents, by accepting to partake, also trust their integrity, as well as mine.

There are further limitations to the study simply from the chosen research design. Firstly, a limitation relating to phenomenological studies is that the impact of research on public policy is typically limited when methods are essentially qualitative (Easterby-Smith et. al., 1991). Because this study aims to gather information that I believe to be helpful for the continuation of improving Lower Secondary English assessment, I have considered this fact carefully and added a quantitative element, which I use to show the qualitative data in a broader perspective. Nevertheless, it is important to note that this is not a mixed-method study, and the focus remains on the qualitative data as it is presented as responses from the four teachers who gracefully gave access to their reflections and experiences.

Secondly, Postholm & Jacobsen (2018, p. 227) point out that we will never detect better data than we are able to register. Although this point relates to the use of e.g. recording devices, it also shares relevance to how the quality of data and findings depends on a researcher's skills and experience level. My inexperience with qualitative research resulted in me being highly self-aware in the first interviews, and I found the responsibility to be vast. Nevertheless, with time, I became more secure in my role as I conducted more and more interviews. The fact that the data material consists of transcriptions of 11 interviews conducted with 4 respondents could be a strength for the study because it reflects data at different points in time, and also revealing an increase in comfort level both for me as a researcher and in the researcher-respondent relationship. As the data was indeed so substantial, it also required an extensive analytical engagement with the material, which could have resulted in a different selection and interpretation process by more experienced researchers. Nevertheless, as explained in 3.5, I have attempted to the best of my abilities to extract meaning from the data in a truthful and meaningful way.

4 Data Analysis & Findings

This thesis has been guided by the goal of identifying themes that could describe some teachers' perceptions of curriculum change in the English subject and their experiences from the transition from two grades to one. As the data material revealed quantities of interesting findings, I had to remind myself to be guided by the thesis' research question in order to separate relevant findings from findings not correlating to the phenomenon in question, and to establish certain focal points. The research question is: *How do teachers experience transitioning from two grades to one in Lower Secondary English?* This chapter presents three themes based on the qualitative data, then the questionnaire findings. I have chosen an approach which emphasises detailed descriptions of experiences rather than more general presentations. Additionally, I include a relatively large number of quotes to emphasise the teachers' perspective, both for transparency and to reflect their experiences using their own words. I have found it beneficial to explain certain matters using a representative quote, particularly examples of relevant assessment cases, while on some matters, I have used less detailed descriptions to depict multiple teachers' experiences.

4.1 Assessing Overall Competence

This theme encompasses responses from all four teachers, revealing reflections and experiences concerning the assessment and grading in the transition from two grades to one grade. None of the four teachers showed any concern about the use of the term *overall competence*[21], with Lisbeth saying it was "not at all problematic" and Ragnhild reflecting that the use of the term in a way validates the move from two grades to one. Markus was less convinced, and stated that, although it was "fair enough" to transition to one grade, "oral and written [competence] are two very different things when they are tested … They are connected, but they are two different ways to express yourself in English". Markus said it was clearer to assess one grade, and displayed a preference to providing two grades, which he shared with Peter, and, to begin with, Ragnhild. Peter explained how it was sometimes challenging to assess pupil's competence as it consists of varied components which could be hard to assess 'as a whole'. He put it this way:

> It was easier to assess the two competences separately … [In the Autumn term,] I have had instances of not knowing which overall grade to give, because how do I decide what's more crucial, what trumps the rest?

In different ways, all four teachers reflected about how to weigh the different components of competence against each other: How do you build an image of, or a model, of individuals' overall competence? Both Peter and Ragnhild noted that, generally, grading was not as challenging as they first had expected. Ragnhild said she was very negative to start with, but displayed an optimistic outlook and curious attitude over time, concluding in June that "it was not so bad". Lisbeth noted in the first interview round that assessing overall competence is "not that hard"; that it requires you to "look at what they have done well … then that's not as good, that's dragging you down a bit … And then we find a middle way". When we met again after Christmas, she had experienced

---

[21] "To assess overall competence is explained by the Directorate (2022b) as seeing competence aims in connection with each other.

challenges with the assessment of two pupils and ultimately contacted the Directorate and the County Governor[22] for guidance (see 4.2.3). Nevertheless, she noted in June that "it is in many ways just as easy for teachers to have one grade" and pointed out that experience could be a factor. Peter, who has the least amount of experience, reflected that a teacher's experience, competence, and interpretation are important factors in relation to assessing overall competence:

> This is new for everyone, providing one overall grade. … Whether it is habit, experience or the fact that it is split in two, I don't know, but I experienced it as simpler [giving two grades], no doubt about that.

> While previously, you could split [written and oral performance], and a pupil could get one very good grade and perhaps a weaker one in the other, they are now assessed as one. And how you weigh those up against each other … It will likely be worse for them [those with weaker competence and challenges such as dyslexia]. … And then you have to factor in a teacher's competence, experience and interpretation.

This reflection is from the first interview with Peter, and when I met him again in June, he called the weighing of the different components of competence "the core of the matter" and reflected on the lack of collegial unity and guidelines as problematic (covered in 4.3.3). Peter reflected in June that the characteristics of goal achievement help to "see what they [the authorities] are really looking for", but that teachers need more guidelines to fully comprehend and operationalise the term *overall competence* and how the components should be weighed in more complex assessment cases. One of Ragnhild's main concerns in the first interview round was how to ensure fair assessment, and how different weighing of components impact the competence levels attained by the pupils, for example if there are any consequences from reducing time spent on grammar instruction. A third concern was how to combine their overall reduced focus on grades with documentation requirements. Several of the teachers pointed out that if pupils complain on their final grades, the teacher is required to explain the assessment basis. On this topic, Lisbeth said that she expected few changes due to having one grade, but speculated that a broader assessment basis would make it easier to defend a grade in the case of receiving a complaint. On the other hand, Markus noted that teachers now might have to have the pupils' right to complain in the back of their heads when grading, and that it was more complicated now that the grade reflects, or unites, several competences. Markus also explained that there is a risk that the oral aspect of the English subject could receive less focus when assessing overall competence due to it being easier to document written competence:

> The written grade is weighted more; has always been weighted more, I think. And might possibly be easier to assess … It is more concrete [than oral assessment]. … I fear that there will be fewer oral assessments.

This concern is shared by Ragnhild and Peter, with the latter drawing on discussions with his colleagues:

> What do we do with pupils whose competence reflect a 4 in written English and 5 in oral English? Is the written component easier to add weight to, because it is easier to measure, easier to 'see', not as much room for interpretation and subjectivity? …

> We have discussed what is weighed more, what is important. And that lead us to whether oral skills are maintained. Do the oral competence aims demand more of us than we have been able to achieve? Overall competence…what is it? Can we see what it is? These were some questions we had floating in the air, but we weren't able to land any answers.

---

[22] Statsforvalteren

Lisbeth had the opposite view: "There is no doubt that written assessment is more demanding [than oral assessment], so … If something is cut, I would think that would be written assessment". In terms of other components of competence besides what is assessed through their oral and written production, all four teachers participating in the interviews, talked about how they assess listening, reading, and understanding as well as undefined activities in the classroom. Ragnhild explained in the second interview round that she had become more focused on integrating the different components of competence in her English class. For example, she would combine written and oral components in a podcast, and reading, comprehension and speaking in activities where the pupils would search online for a topic, then talk to a fellow pupil about it:

> They watched [a film] and were assigned to search a bit on Google individually, then tell one another what they found out. … This way I get … these small drips, where everyone is orally active. And for those who find it hard to put things into words, they are allowed to show their understanding in Norwegian.

This is one example of how these teachers adapt their teaching and assessment practice to pupils' individual needs. Others are presented in other sections. One assessment situation which was mentioned by everyone was having one-on-one conversations[23] where the pupils would be able to demonstrate skills and competences without an audience. This was explained as a helpful tool both for pupils and teachers; a good way to gain insight into their competence level, and explain how the pupils had demonstrated their competence, how this was interpreted and graded. Lisbeth explained this as a good way to get an impression of pupils' comprehension, and that it was "not that difficult" to assess their level of comprehension. In terms of other constructs in the English subject, Markus explained that it would be difficult to find a schematic representation as there are so many variables, but that that was not an argument against guidelines: Rather, pupils would benefit as it would result in more similar assessment basis across schools.

Another challenge, voiced by Markus, is the decoding of the term *competence* and the subject curriculum, particularly the interaction between the core curriculum and the subject curriculum. He said:

> It is claimed that the core curriculum should 'sprinkle' on everything, which is great, there's a lot of good content there – at least if you can decode it. But the problem is that it is decoded differently, based on which school you are on. So it becomes a bit random some places. Perhaps we are a bit old-fashioned, but we use the competence aims as our starting point. Even though we of course also take into account [the interdisciplinary themes] as well. But in the end, we use the competence aims.

The competence aims were problematised by several teachers, with Markus saying that while they are "rounder around the edges", they present other problems as they are not necessarily easier to operationalise. Furthermore, they give more room for interpretation. Peter reflected in similar terms:

> The competence aims have become clearer, but there is still a way to go. That balance, deciding how it is weighed, it is very difficult. And time demanding. … The curriculum has many competence aims relating to written competence, several which can be interpreted to be used for both oral and written components, and some specifically oral … What is the balance between them, and how do we use them in the subject? … But then you are looking at a mathematical formula again …

In the last interview round, I asked the teachers about their thoughts on the use of the verb *explore* in the competence aims. Peter said they had worked briefly on it but that it

---

[23] Fagsamtale

had not impacted assessment practice. Lisbeth questioned how teachers should and would assess pupils' ability to explore:

> We have worked with exploration for years. … We have peer guidance where we focus on exploration. … To put a grade on a pupil's ability to explore is odd. In many ways. … It is limited what it can tell you about your oral or written competence. … It is like grading the ability to think. It is so vague … It is entirely up to how the teacher facilitates exploration.

Ragnhild agreed it might be challenging to assess, and that it is connected to their subject interest. A second connection was to creativity, and Ragnhild mentioned introducing creative ways to work with language, such as role play and tools like story cubes and finger puppets. Markus and his colleagues had worked specifically on how to understand *exploring* with regards to the competence aims:

> It might be easier in written work. They get a research question, a dilemma or a situation which they write about, which they reflect about. Drawing on different sides of the same issue without concluding that this is how it is. That's what we have landed on as a definition. … To find information, other viewpoints, that can be part of it. But grading it … The assessment criteria becomes incredibly important; that we in advance set criteria, preferably with the pupils, to reflect that they have explored something to the extent they have.

Furthermore, Markus emphasised that it is important to break down the curriculum to something concrete which is meaningful for the classroom experience, and that they would usually set criteria with pupils ahead of assessment situations. The other teachers also emphasised pupil involvement in creating learning goals.

4.2 Consequences for Grading and Motivation

This theme takes its basis from responses from all four respondents about how to decide on what grades to give their pupils when moving away from one oral grade and one written grade. The teachers all provided examples of how they decided on grades for the Autumn and Spring term (and for Lisbeth, the final grading). It is when we were able to talk about specific cases that the greater complexity in the change truly became evident. Furthermore, this theme establishes the respondents' reflections and experiences regarding potential and experienced consequences for pupils as a result of the change from two grades to one, and pays particular attention to outcomes on motivation in relation to the grades given.

The teachers established that there was often coherence in pupils' oral and written skills, but everyone raised concerns about more complex assessment cases, where pupils displayed disproportionate skill sets. Ragnhild noted: "There is not always coherence between the oral and written [performance], which means their full competence is not measured. Some can say directly that they do not want to speak". Indeed, Ragnhild, and others, have had experience with pupils who for various reasons are reluctant to speak, and this would often result in one high and one low grade with two grades. On this topic, Markus pointed out that pupils avoiding oral activity might feel relieved by the reduction in number of grades. Nevertheless, all teachers reflected about whether the transition to one grade might be a disadvantage to pupils who are proving to be particularly academically skilled in one area, but show weaker development in other areas or in cases where for example a pupil refuses to be orally active or hand in written assignments. This was mentioned multiple times across all interview rounds. One point raised is that the grade will not necessarily reflect the pupil's true competence, but rather be an average of the overall impression. In this context, all teachers mentioned the potential impact on motivation for certain pupils. For example, Lisbeth said that motivation will suffer among

pupils who, with the two-grade system, would receive a high grade in oral English, but a lower grade in written English:

> They will not experience the motivation and the sense of accomplishment by getting that 5 or 6 … because their grade will [now] be lower because of the other proportion of the subject.

This concern was shared by Markus, who explained that pupils demonstrating an overall lower competence would also benefit from receiving two grades because it would "lift what they do best". Markus pointed out that his primary concern was that these pupils would not experience the joy of achieving a grade reflecting the strong skill set in one area, even though a higher competence in one field would now of course impact the grade positively and be reflected in conversations with the pupil. Markus further reflected:

> It would have been an advantage to have what they need to work more on reflected in black and white. They do, regardless [through feedback]. But regardless of what you comment on, whether it is written or oral skills, a grade is more of a tangible evidence on how this and that has gone. It might be easier for the pupil to relate to two [separate] grades.

On the other hand, Ragnhild pointed out that pupils who are not motivated by grades, or who continuously receive assessment signalling lower competence, fewer grades could be beneficial. Furthermore, she argued that, with one grade, teachers may be able to gather more of a basis for assessment to support grading those who would otherwise be at risk of not receiving a grade in either oral or written English:

> With one grade, you might have more material to take from … We try to stretch ourselves far to provide grades, but if you have nothing, you have nothing. You can't guess a grade. … But I feel like we try to … offer an extra chance. Try to make sure that everyone gets a grade.

> It might ease a little, because you can use the written [component]. You get an overall competence. The oral [component] could pull you down of course, if there is not a lot of assessment [basis] there, but at least you have some basis for assessment in other [components], if you don't have anything in every area.

The teachers do not apply this principle in cases of inconsistency at higher demonstrated competence levels. For example, Lisbeth reflected that a high inconsistency between oral and written production makes it difficult for pupils to achieve a higher grade based on one of those components, and that there is a need for some proximity in competence. As a result, such pupils would often "land in the middle". Lisbeth highlighted that there is an individual aspect to this, and that pupils would always be assessed on a case-to-case basis. Furthermore, she explained that the result of high inconsistency can be that one grade does not reflect a valid result, a concern shared with Markus, who explained that two grades gave a more accurate reflection of pupils' competence. Lisbeth called one pupil's overall grade "a wrong summary" because the grade did not reflect the pupil's true competence: With two grades, they would have received a written grade 4 and an oral grade 1. With one grade, the pupil received a "thin 3" and could have received a 2 as their final summative assessment.

Similarly, Ragnhild explained how a pupil was given the grade 5, while with two grades, they would have received one 5 and one 2 or 3, lowering the average substantially. The reason was linked to anxiety and the pupil not speaking in class. However, the pupil had demonstrated a very strong written performance and would deliver audio files displaying a good vocabulary and grammar, and was given a 5 overall. Pupils' classroom activity was raised as an important factor in assessment basis as a complement to more formal

assessment such as term papers, written submissions, oral presentations, and audio-files. On this topic, both Markus and Peter problematised the fact that effort is not to be a factor in assessment, yet it is an important factor in showing what they know. Peter exemplified:

> A pupil who was given a 5 in the Autumn term, landed on the grade 4 this Spring. She is a highly skilled pupil, but I experience varying efforts or motivation. This has affected her classroom behaviour, but … I see that for formal assessments, she mobilises. We had more formal assessments in the Autumn term than in the Spring term, which made it easier [to assess her competence]. … They have to work to show results. … It is not a competence aim to show effort, but if they do not show effort, they do not show what they know – which affects grading.

None of the four teachers addressed any specific *difficulty* with grading with regards to pupils with challenges which could impact their oral or written production, such as dyslexia, hearing problems, or anxiety, but several had experience with it. However, providing one grade is described as "unfortunate" and "problematic" with relation to pupils who show disproportionate performance in oral and written skills. Furthermore, psychological aspects can hinder a pupil from producing in class, limiting a teacher's insight into a pupil's competence. One of the teachers talked about their experience with having a pupil with selective mutism in their class, who rejected the offer to demonstrate oral competence using audio files. They reflected on how they graded this pupil back then, compared with one grade:

> I think we ended up not giving an oral grade on this pupil. No, wait, because they can show oral comprehension through written production. So they received the grade 2. But we did not evaluate the basis of assessment, that's how it was. The pupil was aware there would be competence aims they would not achieve. Their written grade was 4, it was a pretty strong pupil. But with one grade, I suppose they would end up with 3. … That way, it is unfortunate if there is a great difference.

Lisbeth pointed out that once dyslexia and similar challenges have been uncovered, they will typically get aids, which can be very helpful if they are used well. They can also be exempt from assessment. Furthermore, dyslexia present itself in different ways, which affects their performance and ultimately, the assessment basis and the grade received:

> If they use their aids well, a lot of the problems with writing are corrected. I have had two pupils with dyslexia [in the past year], one with issues related only to writing. So he has a vocabulary someone with reading difficulty would not have. When he learned using the aids, … a lot of it was solved. The other pupil was not as willing to use the aids, which affects the language as a whole, vocabulary and… both written and oral [performance]. I find myself sitting as an interviewer while he waits for the next question. … One of them has a strategy where he uses body language a lot. The other is more vocal and easier to have a conversation with. [In terms of assessment], I have assessed what they produce orally and in writing. [The basis of assessment] becomes scarce when you have limited production.

One point made by Lisbeth is that, more so than before, pupils might be pulled up to the higher grade if their competence level is between grade levels. She explained that she thought pupils would now be given higher grades in cases were the oral production is at a competence level reflecting the grade 4 and the written production reflects the grade 5. On the other hand, Ragnhild explained she had a pupil who she said would have received a 4 in written English and a 5 in oral English, but was given a 4 with one grade. She emphasised activity in the classroom, homework and learning goal achievement. Similarly, Markus experienced providing the higher grade in some cases where oral performance was comparatively higher.

Lisbeth explained that her above assumption would likely not apply to cases where one component reflected the grade 5 and the other 6, as "there is a threshold to pull you up to a 6 compared to pulling a pupil up to a 5 if in doubt". She thought fewer pupils might receive the highest grade, 6. Similarly, Ragnhild described how she had *"dared"* to give one pupil a 6, due to a strong demonstration of competence after receiving a 5 in the Autumn term. Furthermore, when describing how they decide on grades when in doubt, both Lisbeth and Peter used the phrase "giving the benefit of the doubt", with Lisbeth explaining further that the aim would be to give the highest possible grade that *realistically* reflects their competence.

Lastly, one teacher pointed out that having *two* grades would potentially be more beneficial for Lower Secondary teachers when pupils change schools, as well as for Upper Secondary teachers when interpreting an individual pupils' competence level from Lower Secondary based on grades:

> I had a pupil transfer from a different school [this year], which was given the grade 6 [in English], but … On what foundation? Based on what I have seen, he isn't even close. We might see something similar in the transition from Lower Secondary to Upper Secondary School: What is the foundation for the grades? I know why I have given this grade, but they [the Upper Secondary School teachers] don't. This means little for those at the higher end of the scale, those given 5 or 6 … but for those who are not there, we often see that one proportion [either oral or written production] pull them up or down, and this would be good to know … You get a more nuanced overview of [a pupil's] competence. What are they good at? What do they need to work more on? … In many ways, it is easy with one grade. But I do not think it gives the pupil an equally good understanding of strengths and weaknesses.

4.3 Impact on Practice, Collegial Collaboration & Subject View

This theme is concerned with the study's four contributing teachers' reflections and experiences relating to teaching and assessment practices, collegial collaboration, and view of the English subject when transitioning from two grades to one. At its essence, this theme touches on the experience of being an English teacher while navigating curriculum changes.

4.3.1 Thinking & Planning Differently

In the first interview round, Lisbeth said she did not personally anticipate any big challenges or having to make changes to her assessment practice as a result of the change, but rather reflected that it might lead to a reduction in time spent on assessment. She confirmed the latter when we met again in January and June. Nevertheless, Lisbeth quickly noted: "You need to think differently and plan in a different way". No one noted in the first interview round that they anticipated big changes to their assessment or teaching practices, but there were also no responses indicating that there would be no changes at all. Ragnhild, who was initially quite negative to the change, reflected the following in the first interview round:

> We might be able to show more breadth in the assessment. We need to consider how we design tasks and what we assess. So maybe it will challenge us to maybe do things a little differently than when we had two grades. Maybe we can merge things together, so that pupils can show both written and oral competence at the same time; not separately, but … connected.

Ragnhild's positive and explorative attitude contrasts her initial pessimistic outlook on the change, as she explained herself as "initially very negative". Nevertheless, she later said that "we expect the pupils to be curious and explorative in the subjects, so maybe we

should be as well … We need to work in new ways too". Ragnhild's reflections also contrast Lisbeth's indication of no personal need to make significant changes to her practice. While Lisbeth initially referred to it as "not a problem" and "easy", she emphasised its weaknesses and potential consequences in similar ways to those respondents showing a preference to the two-grade system. By the second interview round, this initial positive attitude was complicated by grading highly complex cases, which consequently meant that indeed, Lisbeth would need to think differently in order to provide an assessment for two pupils (see 4.3.3). Peter experienced few challenges in transitioning from two grades to one, but said he preferred the two-grade system. Furthermore, he expected to work more on assessment practice over time:

> Maybe not this year, but eventually I do expect that we view things a bit differently, like how we organise the school year and how we assess the pupils, what components we assess. At the same time, I do feel like we already assess pupils in different ways, with different types of assessment. But there is no doubt that [LK20] gives us opportunities, and demand that we test pupils in other components and other ways than we have done previously. I feel like we haven't had enough time yet, but I see that over time, this is something we have to look at.

For Ragnhild, this had already materialised in reflections on how she could now give pupils more innovative, open-ended tasks. She first mentioned this in the first interview round, then elaborated on it when we met again in January: "I am more and more intrigued by open tasks, where they have choices. But they get to use the competence they have. They have an option to choose". She shared this interest with Lisbeth, who signalled a wish in the second interview round to use a similar task design after Easter to give pupils a so-called second chance – a completely open assignment, where they would get guidance on which competence to focus on:

> We will talk about it in advance, what would be good to focus on if you have not shown your oral competence … to elevate … 4–5 pupils might need some additional guidance on what to choose, and do, but I will have one-on-one conversations and provide some input then.

Lisbeth calls this 'the unknown task', which is unknown to both the teacher and the pupils prior to execution. She elaborates:

> I have a pupil which has very strong written skills but has shown less of her oral skills. She would benefit, because I know she also has strong oral skills, to choose the proportion [of her competence] which she … uses the least. To manifest that she is equally good in [oral English]. … Since I have not done this before, it could get very interesting.

When we met again just after the final grading in June 2022, it turned out there was not enough time to go ahead with this assignment, despite Lisbeth's enthusiasm. Time constraints is something that was established as a challenge by all four teachers, with particular emphasis on its consequences on planning, collegial collaboration, and assessment. Peter echoed that time goes fast after Easter, and that things do not always go according to plan. This has consequences for assessment:

> The Spring term is very busy; we sometimes have to skip something and admit that we do not have enough time … If we do not have enough formal assessment situations, we are on thin ice. That's a stress factor and a challenge.

On the topic of time, Peter explained how he only had time for two formal assessments in the Spring term, compared to four in the Autumn term. Not being able to go through with all the planned assessment situations was explained due to the low allocation of hours to the English subject, and interference from non-English projects. Similarly, Markus explained how he had not been able to go ahead with everything he had planned:

"That's how it is, they have PRYO, this and that". The respondents all signalled a wish for a higher number of hours allocated to the English subject. Nevertheless, Markus concluded, "there are many subjects that would benefit from more hours". Regardless, all teachers reported that they felt they had sufficient assessment basis overall.

Another interesting finding was that among the three teachers who first indicated a preference to the previous system, none reported significant challenges. For example, Markus, who was perhaps the least convinced about the transition to one grade, said in June that it "had not been particularly difficult this year", although he thought he had spent "the same, if not more" time on assessment. Teachers Ragnhild, Markus, and Peter all displayed over time that their reservations towards the change was not directly corresponding with their experiences. Peter reflected:

> I tried to look at it as a new experience. … How did I assess oral and written competence before, … and how do I do it this time around? Then I tried to give an overall grade based on the work they had done. And without taking an average, I felt that the grades were pretty much given. But you need to think a little differently.

This is similar to Lisbeth's reflection about the need to "think differently and plan in a different way". Markus reflected in June that he did not think the change to one grade had affected how he works as a teacher, but that it has impacted how he views the components of competence which are covered in the English subject. For example, he said that "I try to make sure that the oral component is given as much emphasis as the written component", which is an interesting find considering that he in previous interviews had thought it likely that the oral component would fall behind.

4.3.2 Views on English as a Subject

Reflections and views on teaching practice can be linked to how we view the bigger picture: views on learning, how we understand and decode the curriculum, and school & subject politics. In the initial round of interviews, all teachers reflected on their view on English as a subject, perhaps especially in terms of the number of hours allocated to English in Lower Secondary school. In the first interview round, Lisbeth said the English subject seems deprioritised over time:

> While I don't see a problem with having one grade, practically; what I do see is that the subject is minimised. We see it with the number of hours, the written exam, constantly stealing from the English subject. … My initial thought when they launched the idea of an overall grade was that it questionably reduces the English subject's value. … When you constantly cut some corners, take away one grade, steal a class for an hour, it creates less and less space for the English subject. And when the space is reduced… Well, I am one of many that thinks the subject is being downgraded.

She had not changed her mind when we met again in January and June. In fact, Lisbeth's perspective had been reinforced by the example questions for the written English exam introduced in Autumn 2021, which introduces additional elements such as listening tasks. According to Lisbeth, the questions used for the term tests were "too easy" for assessment purposes, with 2/3 of her 10th-graders achieving the grade 6. All teachers agreed the listening tasks could not be used as basis for assessment. Peter said there was "an inflation in high grades" in the term tests based on questions provided by a textbook publisher, and that his colleagues agreed they could not use it as confirmation of the pupils' competence level. Ragnhild reflected that from her perspective, the grades of her 8th-graders were unrealistically high. However, it was good for those who typically receive lower grades or consider the English subject hard, to get that sense of

accomplishment by getting a higher grade than usual. She emphasised that the pupils had to concentrate, as alternatives were often worded similarly.

Markus reflected that the English subject would benefit from having more time, and that it is "unfortunate" and "a step back" to move away from two grades:

> … And English, that's 3 hours a week. It's … Well, we need that time to say the least. There are so many subjects that would benefit from more hours. But … [providing two grades] demanded a high number of assessment situations. And you have to be creative in the ways you get pupils to speak. The oral aspect might be less prioritised [with one grade], that's the feeling I have. And that would be unfortunate. I think it's a step back, perhaps. But … We just have to make the best out of it.

In this quote, Markus reflects that providing two grades lead to "a high number of assessment situations", yet, interestingly, he preferred this setup. He reflected that it could be a matter of habit and that his words could later be "put to shame", but he reflected honestly on these topics based on gut feeling as well as conversations with other colleagues. On what he thought specifically about the transition, he reflected: "We [my colleagues and I] are a bit … astonished. … We don't see the point, perhaps. I'm sure there's a good reason for it. We just have to deal with it."

4.3.3 Lack of a Shared Understanding

This sub-theme could also have been named "Working without guidelines", which is something that was problematised by all. The reason for this is that none of the four teachers were of the impression that they had received sufficient guidelines about how to transition from two grades to one grade. Both Ragnhild and Peter mentioned that the characteristics of goal achievement could be used as guidelines. Ragnhild said it could also be a beneficial tool if pupils were to formally complain on a grade. Nevertheless, Peter said this is not sufficient:

> We would have liked more support, more guidelines, when they choose to move from two grades to one grade. More concrete feedback, tips, support, guidelines about how to do it. Now, I feel it will be different from school to school, how to assess this, and to get a more shared [understanding]. Going in the same direction in terms of English competence; I think that would be beneficial for essentially all teachers.

Lisbeth explained there was a lack of a "red thread", a guideline to ensure similar assessment across schools. This particular point is reflected in interviews with all the respondents. Ragnhild reflected that her initial thought was that it was potentially "unfair" for the same reason, as well as "unfortunate" for the pupil and "challenging" for the teacher. She emphasised particularly the weighing of competences when deciding grades. When I met Ragnhild again in January, she no longer experienced it as "challenging", but still questioned why teachers had not received any guidelines: "We can be very different yet have plans to help us become more alike". This point on reliable assessments was raised by Markus on several occasions, who noted that individual teachers' autonomy could result in instruction becoming person-dependent, and that guidelines could be beneficial by making it more predictable and similar across schools, regardless of geography. Markus' main concern was the pupils:

> We need to be able to ask questions, to wonder what the goal and the gain is. The goal can be so and so, but the gain I am less certain of. For the pupils. At the end of the day, it's about the pupils.

A potential consequence became clear in the second interview with Lisbeth, when it turned out that she had contacted the Directorate and later the County Governor for

guidance on how to grade two pupils. She explained how the lack of guidelines could result in teachers providing grades not reflecting the pupil's true competence:

> I have two pupils on different ends of the [competence] spectrum … One which refuses to speak English completely … Then another which will speak, but is having problems producing any written work. … So, previously, this has been easier. If you didn't have any oral production, then you wouldn't get an oral grade. If you didn't have any written production, then you wouldn't get a written grade. So I contacted the Directorate to get their perspective on how teachers should act in instances like this … Then they responded that I needed to contact Statsforvalteren. Then… they said it is up to the individual teacher. So … If Statsforvalteren receives a complaint, you have to show that you have tried to solve the problem and give reasons as to how you ended up with the assessment. … The thing is, when there are no guidelines or explanations about this, you can risk that teachers automatically… divide the grade in two, and says his written skills are about a 4, but I have no oral assessment, so he will get a 2. While another thinks the pupil's written production is so good … that he gets a 4 anyway.

Furthermore, all four teachers share the concern that the lack of guidelines and a shared understanding across schools could impact the overall assessment pupils are given. Peter explained that new developments in the English subject "provide us with more opportunities, but also challenges: How should we work? We would like more guidance to achieve a shared understanding". Among the factors that may create inconsistencies, Peter mentions competence, experience, and interpretation, which was echoed by Markus, who emphasised that LK20 in general can be decoded in different ways by different teachers:

> We have decoded LK20 and established a three-year plan, tried to do so in compliance with the overarching goals and the core curriculum, a lot more than this particular topic of one grade. If at all. We have had informal conversations among colleagues. … The problem is that [LK20] can be decoded differently, based on which school you work at. So it can get quite random in some places.

Markus experienced the collegial collaboration relating to the new curriculum as useful. In groups, they had worked on interpreting the core curriculum and presented this to the other teachers across subjects and cohorts. About collaboration with other English teachers, he said:

> We have talked about the grades being merged, but we have not discussed practice. … We have discussed the curriculum a lot, and some are happy they can do as they please … But I don't think you should be able to do as you want, there should be some guidelines.

Markus also explained how they had not had time to discuss topics like assessment of 'smaller' constructs in the English subject, listing mediation and listening as examples. In fact, everyone explained that there had been a significant focus on LK20's bigger changes, but less so in the English subject specifically. Two of the four teachers had problems participating in the English subject meetings as these meetings were often timed so that they had to meet with teachers in another subject, which they had to prioritise because they were fewer teachers in those subjects. Those who had no such structural problems, Markus and Peter, described mostly informal discussions regarding the transition from two grades to one. In June, Peter detailed a meeting in the Spring term where they had discussed how to weigh competences in grading decisions, but ultimately, they had not had sufficient time to secure a shared understanding. None of the other teachers had participated in structured conversations with a goal to negotiate a shared understanding, although one participant indicated more structured talks, but could not remember details as they were far back in time. All four teachers described well-functioning English departments, but pointed to time restraints. Peter emphasised

the need to organise structured conversations in directed time[24] as it was highly challenging to find time collectively outside structured meeting hours, which were already relatively limited and often required focus on time-sensitive matters, such as term papers, rather than more general didactical topics.

4.4 Summary of qualitative findings

This chapter has presented the findings from the study's qualitative data material. For ease of reference, I have added a summary of main findings related to the above themes in this sub-chapter.

Assessing Overall Assessment: Summary of findings

- Main concern is how to weigh competences in cases where the competence level is not coherent
- Clarity is needed concerning what teachers should interpret as the assessment basis
- Characteristics of goal achievement can be a helpful tool, but the teachers indicate a need for clear guidelines for more complex assessment cases
- Three teachers indicate there might be less focus on oral competence with one grade, whereas one says there might now be less written assessment

Consequences for grading and motivation: Summary of findings

- One grade is perceived as beneficial for pupils at risk of not receiving a grade or demonstrating low overall competence
- There is a higher perceived threshold to give the highest grade when the complexity of the assessment basis increases
- School changes or transition to Upper Secondary may cause reliability concerns
- The major concern is cases where there is disparity between oral and written production
  - Receiving a lower overall grade because of lower performance in either oral or written English may decrease motivation
  - The reduced number of grades means it may be harder to motivate those who perform better in one area

Impact on Practice, Collegial Collaboration & Subject View: Summary of findings

- The transition to one grade is not experienced as challenging overall
- No or little systematic collegial collaboration on the transition to one grade
- One grade can open for more integration of competences in assessment, and more open assignments
- Limited time to discuss and gain a shared understanding with colleagues
- Clear guidelines are requested from the national level to ensure reliable assessment across the country

---

[24] Planfestet arbeidstid.

4.5 Presentation of Quantitative Findings

The questionnaire distributed to Lower Secondary English teachers across Norway reveals interesting data complementing the above qualitative findings. In the following, I present relevant findings which are later used in the Discussion chapter.

The teachers participating in the questionnaire were asked to answer how much they agreed to the following statement on a scale from 1 to 5: "It is easier to assess with one grade versus two grades". From a total of 115 responses, the mean response is found to be 2,9, with 15,6% responding they "fully agree" and 16,5% responding they "do not agree".

Table 9. Questionnaire responses to the question "On a scale from 1 to 5, how much do you agree with the following statement? "It is easier to assess with one grade versus two grades." 1 = I do not agree. 5 = I fully agree.

1	2	3	4	5
19	27	28	23	18

The average on a scale from 1 to 5 to whether the respondents agreed with the statement "I have a clear understanding of and overview of how I should weigh the different skills and competences in the pupils' final grade" is 3,08, with 4 being the most chosen alternative (34%).

Furthermore, looking at the questionnaire responses to the question "How have you experienced the change in assessment?", responses show that 48% report that it is more challenging to assess *overall competence*, compared to providing one oral and one written grade, while 30% remark that it is easier. There are varied opinions on most options except flexibility, where only 4 responded that it provided them with less flexibility. 26% reported decreased pressure on varying assessment situations, while 32% experienced increased pressure.

Table 10. Selection of questionnaire responses to the question "How have you experienced the change in assessment?"

More challenging to assess "overall competence" vs oral/written skills	Easier to assess "overall competence" vs. oral/written skills	Decreased pressure on varying assessment situations	Increased pressure on varying assessment situations	More flexibility	Less flexibility
55 (48%)	34 (30%)	30 (26%)	37 (32%)	55 (48)	4 (0,3%)

45% responded that they were "mainly positive to the change", whereas 24% were "mainly negative". 17% found themselves "somewhere in between", and 0 responded that they had no opinion about it. 8 responded "I was negative, but am more positive now", while 8 responded "I was positive, but am more negative now", meaning that 14% had changed their opinion during the transition. On the pupil perspective, 37% of all responses reflect an experience where transitioning to one grade is mainly positive for pupils with a disproportionate skill set, using an example of a pupil who would receive a 6 and a 3 with two grades. 23% responded that it is "mainly negative", whereas less than ten percent said it is either "entirely positive" or "entirely negative".

Furthermore, when asked, on a scale from 1 to 5 whether they believed reducing the number of grades will be beneficial for the majority of pupils, 45% responded 4 or 5, while 21% responded 1 or 2. The statement "I believe reducing the number of grades will have negative effects on certain pupils" received an average response of 3,05, with 38% selecting 4 or 5, and 36% selecting 1 or 2.

Table 11. Questionnaire responses to the question "What is your general experience with assessing pupils with a disproportionate skill set? For example, a pupil you would give 6 in oral English and 3 in written English with the old system."

My experience is that the change is entirely positive for these pupils	My experience is that the change is mainly positive for these pupils	My experience is that the change is entirely negative for these pupils	My experience is that the change is mainly negative for these pupils	Similar, but I have to work more on adapting to their individual needs	Similar, but I have to work less on adapting to their individual needs
9 (7%)	43 (37%)	8 (7%)	26 (23%)	11 (10%)	2 (2%)

An interesting finding from the questionnaire is the predominance of teachers having now added weight to oral skills when assessing with one grade: 32% had increased their focus on oral skills, whereas 8% had increased focus on written skills. Additionally, the questionnaire results signalled an increased focus on 'understanding' (a total of 65 respondents, or 57%) and more focus on listening (a total of 64 respondents, or 56%).

On questions regarding collegial collaboration, 44% reported informal discussions with colleagues, whereas 15% had engaged in formal discussions only. 34% reported engaging in both formal and informal discussions on the topic of transitioning to one grade. The below table shows reported answers regarding topics covered in these conversations.

Table 12. Questionnaire responses to the question "What have you and your colleagues covered in your discussions about the change from two grades to one grade? Choose all that apply."

Alternative	Number (and %) of responses
Assessment Criteria	64 (56%)
Assessment Culture	50 (43%)
The Relationship between Summative and Formative Assessment	40 (35%)
Creating a Common Assessment Practice	54 (47%)
The Potentially Positive Effects on Pupils	70 (61%)
The Potentially Negative Effects on Pupils	73 (63%)
The Potentially Positive Effects on Teachers	45 (39%)
The Potentially Negative Effects on Teachers	23 (2%)
How to Assess the Overall Competence Versus Oral and Written Skills	93 (81%)
The Potential Challenging Aspects of Assessing Overall Competence	57 (50%)
The Potential Positive Aspects of Assessing Overall Competence	39 (34%)
Clarifying the Term "Overall Competence"	54 (47%)
We have not discussed any of this	6 (0,5%)
Other	6 (0,5%)

A total of 81% reported to have covered how to assess overall competence compared to the previous system. Moreover, 80% had used the new competence aims to aid their discussions, and the examples of exam tasks were used by 49%. The new definition of 'competence' had aided discussions for 47%, and 63% responded that they had used the characteristics of goal achievement to support discussions. 13% reported having used the Common European Framework of Reference for Languages, and nearly half had discussed creating a common assessment practice.

Furthermore, 64% responded that they had not received "any supportive material, training etc. provided by the Directorate of Education and Training, the County Governor, your school or other relevant parties". 15% responded "Not sure", while 11% reported support from the Directorate, and 2% from the County Governor. 6% had received support from their school. Furthermore, 77% said they had not "been informed that you will receive any supportive material, training or similar at some point". Only 5 responses indicated that they had received information, while 18% were "Not sure".

The final question asked was "Which of these thoughts about spending more time on the topic of one grade align the most with yours?". Here, the teachers were offered 5 alternatives. 63% (73 respondents) ticked that "[i]t is worth prioritising to create a common assessment practice". 3% responded "[i]t is not important to me, while alternatives "It is important to me" and "We have so many other topics to be covered" received 13% of responses each. 6% responded "We have already discussed it in detail".

4.6 Summary of Main Quantitative Findings

- 15,6% responded they "fully agree" it is easier to assess overall competence, whereas 16,5% responded they "do not agree". The mean answer is 2,9 on a scale from 1 to 5.
- 41% indicated they fully or almost fully had a clear understanding of weighing of competences for grading purposes.
- Nearly half responded it was more challenging to assess overall competence than oral vs. written skills.
- 45% responded they were "mainly positive" to the change, while 24% were "mainly negative". 0 responded they had no opinion about it. 14% had changed their opinions during the transition.
- 37% indicated that one grade is "mainly positive" for pupils whose competence is widely disproportionate, while 23% responded it is "mainly negative". 38% believe one grade will have negative effects on certain pupils.
- 45% reported that they fully agree or partly agree that the change is beneficial to to the majority of pupils.
- 57% reported an increase in focus on understanding, and 56% reported an increase in focus on listening.
- 32% had increased their focus on oral skills, whereas 8% had increased focus on written skills.
- 47% reported having discussed creating a common assessment practice, and 81% had discussed how to assess overall assessment versus oral and written skills.
- 49% had used the examples of exam questions to aid the transition, and 63% had used the characteristics of goal achievement. 13% reported leaning on CEFR.
- 64% responded they had not received guidelines from the national level, while 15% was "Not sure".

5 Discussion

The aim of this study has been to gain insights into some teachers' experiences whilst navigating curriculum change to identify potential challenges and positive results. In the above analysis, I have presented three themes, with the third having three sub-themes to cover the broadness of findings. In the following, I discuss the experiences of these four teachers as they have been vocalised in interviews, and the subsequent quantitative findings, in light of the thesis' theoretical framework and background research. Rather than structuring this chapter using the above overarching themes as guidance, I have decided to structure it in sub-sections relating to what I have identified, in my opinion, to be good discussion points for policymakers and others who take an interest in the topic. My hope is that by doing so, this chapter could also be a helpful read for teachers who are trying to navigate this curriculum change.

5.1 Understanding the goal and the gain

> **Markus:** We need to be able to ask questions, to wonder what the goal and the gain is. The goal can be so and so, but the gain I am less certain of. For the pupils. At the end of the day, it's about the pupils.

Understanding the goals you are aiming for is an essential pre-requisite for learning (Sadler, 1989, in Wiliam, 2010, p. 14), and the respondents in this study explained how they would often involve the pupils in setting assessment criteria for assignments to ensure that they had a good understanding of what they would be assessed in and how they could work towards their intended competence level. According to Wiliam (2000), criteria should be explicit, pre-determined and general, rather than implicit and developed after work has been completed. Furthermore, he explains that "maxims can be used, not as definitions, but as starting points for negotiating meaning so that the learners come to share the implicit standards of quality" (p. 14). In light of the limited explicit guidelines this study has identified or had reported, it is worth asking whether teachers are set up for success in the operationalising process. The teachers are to varying extents comfortable with the transition from two grades to one, with one saying "[w]e don't see the point, perhaps. I'm sure there is a good reason for it. We just have to deal with it". This, combined with other findings would suggest that the reasons for the change have not been well communicated or validated to them, and that some of the teachers would have preferred a different process where they as teachers had been more included. The "non-specific nature" of the subject curriculum (Hasselgreen & Ørevik, 2020), interpreted as "vague" by both Ragnhild and Markus, combined with a lack of explicit, predetermined criteria for success when navigating what is explained as unwanted change by some, is problematic, and provides teachers with a limited shared 'starting point for negotiating meaning'.

Markus explained how he saw no contradiction between assessing overall competence and providing pupils with one oral and one written grade. The way I interpret his answers is that you can interpret different types of competence yet grade their oral and written production, which can vary greatly. This would imply that other constructs in the English subject are implied constituents of the oral and written components of the subject. With

a continuation of two grades, this would lead to different questions, such as what constructs should be included in those two main focal areas. Nevertheless, Markus does have a valid point: You can only make interpretations of a pupil's competence based on their competent performance[25]. With this understanding, it might make sense to continue grading oral and written performance. Teachers opposing the introduction of one grade in Lower Secondary English on a similar basis would benefit from receiving a thorough explanation of the 'goal and the gain'. The Directorate (2019, unpublished) explained in the unpublished hearing summary that it seems like the previous two-grade system encouraged a mathematical approach where grading decisions were based on a view of separate skills and competences, rather than a more holistic view on competence, but if the point is indeed a move away from such a perceived practice of a mathematical view of competence in the English subject, then guidelines should follow to ensure true practice change.

A second consideration relating to understanding the 'goal and the gain' of the curriculum change is concerning which principles to apply with term grading prior to the final summative grade. Because summative assessment for the English subject is only described after Year 10 and Vg1, Burner (2020, p. 54) establishes that all assessment prior to Year 10 is supposed to be formative. The Regulations to the Education Act §3-14 (2020) states that summative assessment gives information about the pupil's competence at the end of the course, supporting indeed that all term grading is to be understood as formative assessment. For formative assessment, the Regulations to the Education Act §3-10–§3-12 (2020) states that the intention is to promote learning, adjust the instruction and increase competence, yet that term grading should express the competence the pupil has reached based on the expected level. The former point implies that these grading events should follow principles for formative assessment. If we are to use the Assessment Reform Group's definition of formative assessment, to seek and interpret evidence "for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how to best get there" (Broadfoot et. al., 2002, p. 1), grading decisions should perhaps prioritise motivational factors over assessment functions, such as providing a valid grade which reflects a realistic interpretation of the pupil's competence, where needed. However, because teachers are required, per the Regulations to the Education Act §3-12 (2020), to indeed provide a grade which expresses the competence based on the expected level, it could become challenging for teachers to provide valid and reliable criterion-referenced assessment in term grading up until the Spring term of the 10th grade. This is because competence aims and characteristics of goal achievement are provided for 10th grade assessment only, and the task of operationalising these for use in 8th and 9th grade is up to the individual teachers.

In this study, the respondents discussed cases from 8th grade to the Autumn term in 10th grade where they had acted in an ipsative manner, i.e. not norm-referenced or criterion-referenced assessment, but rather by reference to personal characteristics and effort (Bøhn, 2020). The teachers did this to motivate pupils to 'gear up' in the following term or show that they cannot lean on easily-accessible skills, but rather have to work for it (by lowering the grade if they were between grades) or motivate to keep it going in the next term (by providing the higher grade, when between grades). This indicates an understanding of term grading as indeed formative, while all simultaneously spoke about term grading as assessment decisions based on criterions. None of the respondents

---

[25] See p. 19.

indicated confusion or frustration relating to how to understand term grading compared to the final assessment in 10th grade, and in retrospect, it would have been beneficial to have specific questions on this to gain insights into this. Nevertheless, Wiliam (2000) emphasises the potential to integrate the view on summative and formative assessment, and Black & Wiliam (2018) argue that a distinction between the terms formative and summative assessment could be counterproductive. A hard distinction between formative and summative assessment in the English subject could thus also be argued to be counterproductive, and the teachers' responses regarding term grading, which I interpret as the integration of elements from both formative and summative assessment, could be a necessary balance. A potential benefit of making the intention clear would be to assist teachers in finding solutions which mitigate the traditionally viewed differences between summative and formative assessment, and result in a shared understanding of end-of-year grading throughout the Lower Secondary School years.

5.2 Establishing the basis of assessment in the English subject

The Directorate (2022b) states that summative assessment requires that teachers have a good understanding of the curriculum, and that it should be based on the pupils' overall competence. The Directorate does not define or detail the term 'overall assessment' in the English subject curriculum, but in a general description, it is referred to as the ability to 'raise the gaze', requiring teachers to see each competence aim in context of other competence aims. It is also established that the competence aims should be understood in connection with the core elements[26] and certain elements from the core curriculum[27], which are to be understood as a framework for interpreting the competence aims (Norwegian Directorate for Education and Training, 2022b). How I understand this then is that teachers should collect evidence of competence, based on all these competence aims, not individually, but in coherence; in forming a unified whole.

One question worth asking, however, is how this is operationalised and practiced. One of the teachers contributing to this study said it might be old-fashioned of him and his colleagues to mainly use the competence aims as the basis of assessment. He explained that from his understanding, the core curriculum was meant to be a bigger part of the assessment basis. It is understandable that teachers may take this approach as it is somewhat unclear what is and what is not supposed to be a part of the assessment in the 'new' English subject. Nothing is made explicit as to how teachers should interpret the 'About the subject' text against the competence aims, or how this is meant to influence assessment in practical terms. In an article published in Communicare, two Directorate employees clarified that what the pupils should be assessed in, is indeed a bit more complicated:

> The new subject curriculums also include a subject-specific text about formative and summative assessment, developed in accordance with the sections on assessment in the core curriculum and written on the basis (med utgangspunkt i) the core elements, the competence definition and competence aims in the subject. These texts express, as a whole, what the pupils should be assessed in; and pupils should show their competence in several and varied ways. The number of competence aims is reduced to give space for in-debth learning (Norwegian National Centre for English and other Foreign Languages in Education, 2020, p. 4)

---

[26] Communication, Language learning, and Textual Encounters
[27] Based on the text "Relevance and central values" in the English subject curriculum: *(inter)cultural understanding, communication, all-round education, identity development* as well as using English to *learn, communicate and connect with others; promoting curiosity;* and *helping to prevent prejudice.* See p. 25.

From this understanding, then, teachers should make assessment decisions based on the *competence aims* as well as the *sections on assessment in the core curriculum, the core elements, the competence definition*, and *the section on assessment in the English subject curriculum*. This subject-specific text about formative and summative assessment was, prior to its implementation, meant to express what pupils overall should be assessed by[28], be a support for teachers, and make clear the progression and coherence in the curriculum[29]. The text relating to summative assessment in the English subject curriculum for 10th grade (Norwegian Directorate for Education and Training, 2020a) is:

> The grade awarded for coursework shall reflect the overall competence of the pupil in English after completing Year 10. The teacher shall plan and facilitate for the pupils to demonstrate their competence in various ways, including through understanding, reflection and critical thought, and in various contexts. The teacher shall assign one grade for the coursework in English based on the pupil's overall competence in the subject.

When asked about their interpretations of the subject-specific text about summative assessment, Ragnhild pointed out that it might be a support for newly educated teachers that it was explicit in the need for varying contexts where pupils could demonstrate their competence. However, the overall impression from the participants' responses was that they did not agree that the subject-specific text matched the intentions prior to implementation.

Researchers Brevik Flogenfeldt and Beiler (2020) stated that "one grade reflects that different pupils master different skills in different ways", but how teachers are supposed to understand this for grading purposes is unclear. Nevertheless, from the 115 teachers responding to the questionnaire, 41% reported they fully or almost fully had a clear understanding of weighing of competences for grading purposes. Meanwhile, 48% responded it was more challenging to assess overall competence compared to the two-grade system. One point raised by a teacher in this study is when pupils are given the opportunity to show listening comprehension by answering in English. While they show evidence of competence in one area of the subject (listening comprehension), they do not do so in others (oral production, interaction). Although this can easily be justified in a learning context, it raises a problematic point in the context of assessment: Should good reading comprehension be equally weighed as good oral production? If not, how should the different competences in the English subject be weighed against each other in cases where the general level is not easy to define? Would it be possible for pupils who show e.g. excellent listening competence but refrain from producing evidence of oral or written competence to get a good grade?

If the transition from two grades to one is meant to invite practice change in order to indeed reflect that different pupils master different skills in different ways, where pupils' oral and written production is not weighed against each other, it is problematic that no guidelines have been written to assist teachers, as they may interpret the subject curriculum differently. For assessment purposes, this is problematic because it may result in validity and/or reliability problems, such as teachers not assessing what they are meant to assess, assessing something they are not meant to assess, or assess widely different to other teachers. Skar and Hopfenbeck (2021, p. 16–18) establish that, based on the limited research there is, Scandinavian pupils and teachers experience a feeling of uncertainty in understanding what components constitute pupils' final grade and how to get evidence of 'goal achievement'. In the context of the transition from two to one

---

[28] Meld. St. 28 (2015–2016)
[29] Norwegian Directorate for Education and Training (2018c)

grade, teachers and pupils alike could experience this, with potential consequences being that pupils could have trouble understanding the basis of their assessment, and for teachers to interpret it differently than teachers in other schools. One of the studies Skar and Hopfenbeck (2021) refer to (Sivenbring, 2018), show that Swedish pupils found it difficult to understand and make use of assessments given by their teachers due to the overly complicated language, and that they apply certain strategies to be perceived as 'good students' to get the grades they need for Upper Secondary. For pupils to be adopting such strategies rather than working on actual learning goals seems contra-distinct to deep learning, which is one of the goals of simplifying the curriculum (Norwegian National Centre for English and other Foreign Languages in Education, 2020). One way to avoid such development would be by simplifying the language used for assessment purposes. Nevertheless, this would require teachers to fully comprehend indeed what they are to use as the assessment basis, and how to weigh competences in more complex assessment cases.

5.2.1 Identifying constructs in the 'new' English subject

Hasselgreen & Ørevik (2020, p. 375) establish that "the construct[30] against which students' overall achievements in English are measured" is *the English subject curriculum*. The teachers participating in this study have pointed to several issues with this. For example, are they supposed to assess based on the competence aims alone, and if so, how should they be weighed against each other? How should teachers grade if a pupil for example demonstrates extraordinary skills in the oral domain, but for any reason struggles with English writing? And what are the implications for reliable assessment across the country?

While nearly half of the questionnaire responses reflect a positive attitude to the change, the questionnaire findings also reveal that it is perceived as more challenging. One concern voiced in the interviews is if any constructs in the English subject should be emphasised more than others when interpreting the assessment basis. While the M87 Reform specifically pointed to the oral component as the most important on every level in the English subject (Norwegian Ministry of Church and Education, 1987, p. 206), the LK20 English subject curriculum is "non-specific in nature" (Hasselgreen & Ørevik, 2020, p. 384). Directorate representatives have pointed out that "English … have emphasised the core element *Communication* as the most important in the subject" (Norwegian National Centre for English and other Foreign Languages in Education, 2020, p. 5, my translation), although this is not explicitly expressed in the subject curriculum. The lack of clarity makes it potentially challenging for teachers to evaluate what elements should be emphasised, in what manner, and whether any component of the subject curriculum should be assumed to be of more importance in cases where assessment is less straight-forward.

Although I did not ask the teachers for their definition or understanding of *overall competence*, I found it relatively clear from their responses that they had an abstract understanding of the many aspects that constitute competence in the English subject. When I asked about or otherwise attempted to guide the conversation towards assessment of interaction and comprehension, it seemed so 'integrated' in the way they work that perhaps it has become difficult to explain or make explicit. I was told "it is not so difficult" and that "you get a sense of that very easily". My interpretation has been that these teachers do perhaps emphasise the written and oral production more than

---

[30] For an explanation of the term 'construct', see p. 27.

other competences when it comes to grading, but that they assess other constructs more implicitly and in less formalised ways.

Establishing *communication* as the most important element in the English subject corresponds with findings by Bøhn (2015) and Hasselgreen & Ørevik (2020). The CEFR (2020, p. 32) operates with the term *overall language proficiency* to encompass general competences, communicative language competences, communicative language activities, and communicative language strategies, and emphasise that it is implicit that communicative language activities and strategies (reception, production, interaction, and mediation) are adopted in classroom tasks. Perhaps it is this implicit, taken-for-granted approach that has been adopted by these teachers, as well: one where the *communication* construct has been integrated and automatised in their teaching and assessment practice.

Nevertheless, defining *communication* as a construct has perhaps more of a theoretical value than a practical one. The Directorate problematises the potential of a mathematical approach to competence, and emphasises the need for holistic assessment, while two of the four teachers in this study referred to the curriculum as vague. In light of Sivenbring's (2018) findings (see p. 66), showing that overly complicated language makes it difficult for learners to understand what is to be assessed, it might be worth asking whether the application of less comprehensive constructs would be more beneficial for teachers, and whether it could trickle down to the pupils as a more comprehensible basis of assessment. Of course, the teachers had worked extensively on how to operationalise the curriculum, but they expressed a need for further support when experiencing more complex grading situations. In the context of constructs, this relates particularly to the weighing of oral and written production, but also what other components the English subject consists of.

Moreover, the CEFR's term *general competences* consists of *savoir* (declarative knowledge), *savoir faire* (skills and know-how), *savoir apprende* (the ability to learn), and *savoir-être* (existential competence) (Council of Europe, 2020; North, 2022). The latter encompasses e.g. interaction, development of cultural understanding, empathy, motivation, reflection, philosophical considerations, discussions, as well as creating, sharing information, developing and building something through interaction (Langseth, 2009, p. 7). While motivation is explicitly mentioned alongside effort as elements not to be assessed, an interesting point made by Peter in this study is that while it should not be a factor in assessment, it *is* an important factor in the degree to which pupils *demonstrate* their competence. In the 1980s and 1990s, research found that such nonachievement factors was habitually included (Moss, 2013), but it has since received increased focus, and the Regulations to the Education Act §3-3 (2020) states that it should not be included in the assessment basis. The application of savoir-être for assessment purposes is perhaps more concerned with the ability to discuss, share information, interact, reflect and develop cultural understanding, which all can be said to have a link to communicative competence (linguistic, sociolinguistic and pragmatic competence), although the latter could be better defined within the 'content' construct[31]. Nevertheless, my lack of data on assessment of interaction from this study, combined with a lack of signs of the interaction construct in the new characteristics of goal achievement[32], raises the question of how teachers do indeed, and are expected to,

---

[31] Bøhn (2015) found that Upper Secondary English teachers focused on constructs *communication* and *content* (see p. 27).
[32] See table 4, p. 30.

assess interaction. One competence aim directly reflecting interaction is to "ask questions and follow up input when talking about various topics adapted to different purposes, recipients and situations" (Norwegian Directorate for Education and Training, 2020a). Council of Europe (2020, p. 70) defines *interaction* as "two or more parties co-constructing discourse" and explain that interaction is fundamental in learning. Inspired by Peter's points about effort, I wonder whether the interpersonal nature of interaction could perhaps make it confusing to differentiate interaction with effort. Unless constructs are well defined, or the assessment basis is well understood, the *savoir-être* in the English subject could be experienced as interflowing and thus be operationalised differently. Peter's main concern was twofold: How are the pupils to demonstrate their realistic competence if their efforts or motivation varies and affects classroom behaviour? And how should teachers assess competence when it is demonstrated in very varying manners? Ultimately, he concluded, "if they do not show effort, they do not show what they know – which affects grading".

According to Speitz (2020), knowledge about the CEFR may contribute to a better understanding of the subject curriculum. The updated CEFR's descriptors[33] could provide an interesting entrance for discussions regarding constructs in the English subject. For example, my application of terms from the CEFR framework to the analysis of the characteristics of goal achievement[34] could serve as a starting point, with discussions to follow regarding what is missing, what there is an overweight of, and questions such as "How can explicit use of constructs assist with assessments in the English subject?". Of course, teachers may conclude it is not a helpful tool. This could, however, invite other conversations and conclusions. Ultimately, the CEFR had not been used in discussions by any of the four teachers, and only 13% of the 115 teachers responding to the questionnaire. However, teachers would need to have time set aside for this, and the teachers in this study indicated that time restraints complicate collaboration in the English subject.

5.3 Identifying possible consequences on grading

The findings of this study show that transitioning from two grades to one may come with benefits and disadvantages to different pupils. This sub-chapter is focused on the findings from the theme "Consequences for grading and motivation". I address the findings concerning pupils as well as those relating to the teacher's perspective, but aspects regarding validity and reliability are rather discussed in 5.6.

The four teachers highlighted that pupils often demonstrate competence with some consistency in skills and competences. Thus, grading decisions were not substantially affected by the reduction in grades in the English subject. This is also reflected in the quantitative data, where 45% of the 115 teachers report they think the change is beneficial for the majority of pupils. The teachers reported that being able to rely on a broader assessment basis is beneficial for pupils at risk of not receiving a grade in the English subject. Furthermore, pupils who are not motivated by grades might benefit from the reduced focus on grades. Ragnhild pointed out that a reduced focus on grades might support learning for pupils who continuously receive assessment signalling lower competence. This is supported by the Directorate's own comment on the change to two grades being potentially positive for those with reading/writing challenges such as dyslexia, as it increases the freedom to show their competence using different

---

[33] See table 3, p. 29.
[34] See table 4, p. 30.

modalities[35]. Additionally, the Directorate (2019b, unpublished) refers to the School Student Union's hearing response, highlighting that many pupils experience stress when receiving many grades, which ultimately can hamper learning. There is certainly research which suggest that low achievers tend to be overwhelmed by testing and grading (Sivenbring, 2018), and a 2020 report on Norwegian assessment practices establishes that a majority of teachers experience that assessment situations contribute to stress (Rambøll Management Consulting, 2020). Moreover, a Swedish longitudinal study (Klapp, 2015) found that among 8558 pupils born in 1967, graded low-ability pupils had lower odds to finish Upper Secondary education compared with nongraded low-ability pupils. A gender difference was also established, as graded girls had higher odds of finishing compared to ungraded girls as well as graded and ungraded boys. This research was conducted in a period where grading was optional in Sweden. Only one of the four teachers in this study expressed that their pupils were generally motivated by grades, while the three others to varying degrees were reducing the use of grades in formative assessment. This reflects a move towards less grade-based assessment, which is in line with the Assessment for Learning paradigm.

Nevertheless, the conceptualised distinction between formative and summative assessment has led to a distinction between summative assessment and learning (Wiliam, 2000). Based on the findings in this study, there could certainly be aspects of assessment where term grading supports learning, particularly with regards to motivation. The respondents wondered what impact the change could have on motivation for pupils who would acquire a higher grade in one component, but now receive a lower overall grade. One aspect of this is the requirement to provide a fair and valid grade reflecting their true competence level, but another is indeed if pupils' motivation will suffer as a result. The teachers problematised how to motivate pupils who show impressive oral skills or a high classroom participation but are simultaneously demonstrating a lower competence level in written English (or vice versa), now that the option of having them work towards a higher grade in one component has been eliminated. Interestingly, the questionnaire responses lean more towards the opposite view: In cases where a pupil would receive 6 and 3 with two grades, 37% suggest one grade is "mainly positive". However, would the responses differ had the example been 4 and 6? For pupils affected negatively by now receiving a lower overall grade, not necessarily reflecting their true competence level, it is of course unknown whether their motivation in the English subject might decrease. This way, the teachers' reflections would support Wiliam's claim (2000) that we should find ways to mitigate this tension between formative assessment and grading events.

Another interesting finding of this study is the differing experiences with assessing pupils whose competence are interpreted as between grade levels. While Lisbeth expected that more pupils would receive a higher grade when between grade levels, Ragnhild experienced the opposite in at least one instance. This can perhaps be explained by the differing assessment principles applied in 8th grade versus 10th grade, as 10th grade summative assessment can be viewed as a 'ticket' for further education or work (Engen, 1996, p. 57). Moreover, Lisbeth emphasised the aim to give *the highest realistically achievable grade*, but that there was now an increased threshold to give the grade 6. With regards to the above point about motivation, this is worth problematising in teacher discussions.

---

[35] From a text (Norwegian Directorate for Education and Training, 2021a) about quality criteria for the selection of learning materials, not assessment.

Finally, the data material confirms that there are certainly dilemmas in reducing the number of grades. Competence in the English subject is indeed multifaceted, and a separation of oral and written skills under-communicates the variety of constructs in the English subject. However, the introduction of one grade means the grade could give what Lisbeth calls a "wrong summary", as inconsistency in competence is complicated to grade.

5.4 Navigating curriculum change without guidelines

The clearest finding of this study is the varying experiences relating to the teachers' attempts to navigate curriculum change without guidelines. While neither of them found it particularly challenging on a general basis to adapt to the new grading system, they all displayed frustration with not having received guidelines beyond the characteristics of goal achievement. They pointed to two problems in particular:

- Limited time to discuss and gain a shared understanding with colleagues

- The subject curriculum will be understood and operationalised very differently from school to school, leading to reliability problems with assessment

Although all the teachers had worked with the implementation of the LK20 curriculum for a number of years, they had much less time to discuss how to operationalise the changes for the English subject specifically. Lisbeth pointed out that there were so many changes that the transition from two grades to one was small in comparison, while Peter reflected that the time allocated to such discussions was so limited that it would often have to stop mid-discussion, leading to ideas still floating in the air, without conclusion. Meanwhile, Markus said they had only had informal conversations about the topic and that other topics had overshadowed this. Robinson (2018) claims that when the implementation of too many changes is intended for the same time, it leads to superficial development and not lasting change in neither leaders' actions nor teachers' practices. Although the teachers contributing to this study displayed willingness to change, it may appear they have little time to commit to change in this context as the LK20 curriculum requires change in many other aspects. Changes to the centrally administered exam have been introduced simultaneously, as well as the introduction of core elements, overarching themes and new competence aims. According to Hasselgreen & Ørevik, 2020, p. 368), such changes to the subject will inevitably impact assessment:

> Changes in the content of the subject necessarily lead to changes in ways of assessing students' competence, as the subject curriculum valid at any given time determines what is to be assessed. … Interestingly, where LK06/13 required students to *discuss* literary texts, LK10 states that *interpreting* literary texts is a curricular aim, which carries important nuances relevant to testing of competence in the English subject.

Such subtle changes to the curriculum as changing the verb in use with the competence aim for which the pupils are to be assessed by, also requires time set aside for discussion. Both *discussing* and *interpreting* are verbs fitting within the communicative competence framework as well as competences such as interaction, production and reception. Nevertheless, 'interpreting' will inevitably require more subjective involvement as the act of discussing does not require subjective optics, but rather the ability to see [something] from different perspectives. However, how does a teacher assess the ability to interpret, and how will the nuances between 'discuss' and 'interpret' be understood by teachers? Why is this change not explicitly highlighted in subject-specific supportive material to guide teachers to a change *which carries important nuances relevant to testing of competence in the English subject*?

A second 'smaller' change in the competence aims is the introduction of the verb 'explore'. The use of 'explore' was problematised in one of my interviews with Lisbeth, who expressed frustration when asked to reflect on how to grade using the competence aims where the verb 'explore' is used. Lisbeth, who explained how their school had worked extensively with implementing exploration in their teaching practices, asked: "How do you assess a pupil's ability to explore?", and spoke about how it can mean very different things to different pupils – and teachers. Only one of the respondents, Markus, explained how they had specifically worked on how to operationalise the term for practical use, and landed on a definition to use for assessment purposes. The implementation of 'explore' for assessment purposes should perhaps have received more attention, as it is defined within *emotional and social competence* in NOU (2015:8), which is explained as aspects not to be part of the basis of assessment in Meld. St. 28 (2015–2016).

One of the more substantial changes to the English subject is the recent changes made to the written exam, which was mainly experienced in negative terms by the four teachers. The listening and reading tasks incorporated in the example tasks had been perceived as not suitable for assessment purposes by several of the teachers, and the only positive aspect raised was that it was perceived as motivating for pupils who would otherwise rarely get good grades. Changes to how exams are structured may impact teaching and assessment practices, as Sandvik et. al. (2012) has established that we may assume that the content of the educational activities in the English subject is guided by the centrally provided exam, and thus what is used as guidelines for assessment. 49% of the teachers participating in the questionnaire used the exam example questions to aid discussions or individual work related to the transition from two grades to one, and 56% responded that transitioning from two grades to one had resulted in an increase in focus on listening. There may be a link between these findings and the change in the written exam. However, findings based on the interviews identifies the changes made to the exam as predominantly problematic, as it was regarded as something these teachers did not lean on as qualified support for grading for neither term in the 21/22 school year.

Robinson (2018) points out that planned changes often fail because those who make decisions, underestimate the complexity of implementation. The four teachers participating in this study were not satisfied with the level of support from the Directorate on this, and among the teachers responding to the questionnaire, only 11% suggested they had received support from the Directorate and 6% from the school they worked at. 79% reported they had either not received or were not aware of receiving guidelines. Although some of the teachers pointed to the characteristics of goal achievement as somewhat helpful, my interpretation of the qualitative data is that it has had limited practical value for these teachers when dealing with more complex grading decisions, where a pupil's overall competence is not perceived as coherent, but rather very varied.

As a result, the teachers involved in this study have navigated a field where they have had to rely on their experience, colleagues, and professional judgment rather than national guidelines. Lisbeth stated that her long classroom experience was beneficial in identifying ways to navigate without guidelines, but that it would probably be experienced as more problematic for teachers with less experience or subject competence. This is linked to teachers' assessment literacy, which includes the understanding of purpose, content and connection between learning goals, assessment methods, data collection of pupils' learning outcomes, and biases (Stiggins 1995 in Fjørtoft & Sandvik, 2016, p. 18–19), as well as understanding how to assess

competences such as deep learning (Burner, 2020, p. 58). There is, of course, no direct correlation between experience and assessment literacy, and Peter, as the teacher having accumulated the least years of teaching practice out of the four, reflected in similar, qualified ways to the three others with 15–30 years of experience. It is, nevertheless, worth asking how teachers without formal background in English didactics, or teachers with no access to an interpretation community, experience the same. Without guidelines, pupils are graded based on the teacher's impression of what the English subject is, and this can be very different in the eyes of an experienced teacher, a newly educated teacher, and one without English didactic, or any formal, competence.

The EVA20 project's first report stated that with LK20, the authorities had invited to a shared, or joint, project, and that it would inevitably lead to tension, dilemmas and disagreement (Karseth, Kvamme & Ottesen, 2020, p. 16). Based on the interview responses in this study, it is certainly worth asking whether Lower Secondary English teachers feel they have been sufficiently involved in what is referred to as a joint project. Although the 2019 hearing about the English subject curriculum specifically asked about whether a move from two grades to one would be a good arrangement, the unpublished hearing summary does not include an analysis of the responses from the Lower Secondary English teachers specifically, those directly impacted by the change, but rather considered all responses equally. How the question in the Directorate's (2019a) hearing was formulated, could also be problematised:

> We suggest a change to the summative assesment in 10th grade: Pupils will receive one grade expressing pupils overall competence at the end of 10th grade. Is this a good arrangement based on the curriculum and the allocation of hours?

Firstly, this formulation raises questions of how to understand term grading prior to 10th grade, and secondly, what the teachers have responded to. Several hearing responses point out that one grade makes sense based on the allocation of hours today, but that it would be better to increase the time for English, based on the subject's importance. One highlights that it is disproportionately low when compared to the Norwegian subject (Teigar Lower Secondary, 2019), which is particularly interesting in light of how a reduction of term grades in the Norwegian subject was introduced following a four-year trial period. The teachers contributing to this study also reflected about how the English subject would benefit from more time, with Lisbeth concluding she thought the English subject had been downgraded over time, and that going from two grades to one represented another thing "taken away" from the subject.

It is worth noting that a study by Wyatt-Smith, Klenowski & Gunn (2010) found that common assessment materials do not necessarily lead to common practice or shared understandings. Furthermore, Rødnes & de Lange (2012) found that guideline documents made available to teachers had not been used as intended. Nevertheless, the Directorate (2009, p. 26, my translation) has previously noted that "experience shows that introducing new curriculums, regulations or other measures in the sector without following up with competence development leads to a low degree of practice change". Based on the findings in this study, it would undoubtedly be beneficial to consider competence development following the curriculum change as a response to validity and reliability concerns.

5.4.1 Facilitating practice change

The four teachers had different opinions on the likelihood of practice change. Findings from this study indicate a motivation to change certain aspects of their teaching and

assessment practices, as a result of reflecting on what constitutes the English subject, and what should be assessed. These changes are detailed in the theme 'Thinking & Planning Differently' and include Lisbeth and Ragnhild's move towards creating more open-ended tasks, where the pupil involvement is increased. Ragnhild's reported practice change, where pupils combined skills or were given an option in terms of modality, may have benefits to pupils demonstrating a broader range of competence, independence and motivation than Ragnhild's example, which she emphasised would demand significant independence.

However, all teachers pointed to the limited time offered for planned, formal collegial collaboration to gain a shared understanding as well as the natural restriction in the low allocation of hours to the English subject. The respondents had all spent substantial time decoding the curriculum, but perhaps not enough time was spent to thoroughly discuss the smaller details of the subject curriculum. The benefits of offering systematic collaboration and discussions on summative assessment practices are established by the Directorate (2022a). Although the Directorate (2022a) insists on the value of interpretation communities and that teachers evaluate their practices with their colleagues, the four teachers established that changes to grading had not been prioritised in formal meetings. Rather, they had spent substantial time on general curriculum changes, and in the English subject, changes to the exam were prioritised. Two of the four teachers also struggled to participate in the meetings in the English department as they were required to go to meetings in one of their other subjects, where there were fewer teachers. A third teacher emphasised the need to organise structured conversation in directed time[36] as it was highly challenging to find time collectively outside structured meeting hours, which were already relatively limited and often required focus on time-sensitive matters. As a result, they had spent little time on this topic specifically, and had not had sufficient time to land on a shared understanding.

The questionnaire data showed that 44% of the 115 teachers had only discussed the change informally, while 34% had discussed it formally and informally. 15% responded that they had only discussed it formally. Among these, 8 in 10 had discussed how to assess overall competence versus oral and written skills. In retrospect, I wish I had added a question asking for an estimate of how much time they had been allocated to formal discussions of this topic, and in what way these had been structured. Furthermore, had these discussions resulted in a shared understanding, or had they, similarly to Peter's explanation, only led to questions "floating in the air"? Nevertheless, 57% reported an increase in focus on understanding, and 56% reported an increase in focus on listening. Furthermore, 32% had increased their focus on oral skills. This could signal that there has been some practice change among those 115 teachers.

5.5 Quality Concerns with the Curriculum Change

Validity and reliability concerns have been established throughout the Discussion chapter thus far, but I aim with this sub-chapter to concentrate these findings to address quality aspects of assessment in the English subject in a coherent manner. The term 'quality' comprises validity and reliability (Black et. al. 2011), which refer to whether teachers assess what they should assess, and whether assessments would produce the same results had they been evaluated by other people, respectively (Moss, 2013). Neither validity nor reliability are mentioned explicitly in the curriculum, but are clear markers of assessment quality. In the following, I discuss concerns of validity and reliability relating

---

[36] Planfestet arbeidstid.

to the transition from two grades to one in Lower Secondary English, based on this study's findings.

Validity and reliability concerns are vital to secure fair, dependable and well-founded assessments. One way to improve validity and reliability is to establish a shared assessment culture, where competence aims, criteria, and performance at different grade levels are discussion topics (Bøhn, 2020, p. 308–309). Moreover, the Directorate (2022a) explains that a 'valid, reliable and fair' summative assessment "requires a systematic collaboration and joint discussions to develop a shared understanding (an interpretation community) and to tune their own practice to others". Certainly, the teachers in this study have been explicit in wanting to establish a shared understanding, but have lacked the time set aside to formally discuss topics related to the transition from two grades to one. Rather, formal discussions have mainly focused on decoding more general aspects LK20 curriculum, and comparatively, the English subject curriculum has received much less focus. Ultimately, the transition from two grades to one has not received much emphasis, which they explained is due the grand scope of the new curriculum, as well as the low allocation of hours to the English subject, and limited time to discuss with other English teachers.

In fact, this study establishes a tension between teachers' limited time in the English subject and the formal assessment requirements of teachers. When planned assessment situations for any reason are not completed, the pupils lose out of an attempt to demonstrate their competence. Because the grade is to reflect a pupil's competence at the end of the school year, it could be problematic when teachers have more formal assessments from the Autumn term than they do in the Spring term. Although all teachers in this study reported that they felt they had sufficient assessment basis overall, several pointed out that the time after Easter was particularly busy and detailed how they had more formal assessment situations in the Autumn term compared with the Spring term, despite their best efforts to even this out.

For teachers, there is a risk that their assessment basis becomes flawed when planned assessment situations are missed. This has potential consequences for grading and particularly in light of documentation requirements needed in case of formal complaints from pupils on the formalities of summative assessment. Markus said it might now be necessary to consider the potential for complaints more so than before because it was more complicated compared to providing two grades. One concern is that what is regarded as not following the applicable regulations for summative assessment[37] in the English subject is complicated by the lack of guidelines. If a teacher receives a complaint, they will have to document that the final grade is based on an overall assessment and not based on a selection of competence aims, or a mathematically derived average of grades from formative assessment situations (Norwegian Directorate for Education and Training, 2022c). Based on Lisbeth's report on feedback from the County Governor, "it is up to the individual teacher. So if the County Governor receives a complaint, you have to … give reasons as to how you ended up with the assessment". The response was in line with the principle of local autonomy, relying on local interpretation and documentation in case of a complaint. If a teacher for any reason cannot establish that the assessment basis is sufficiently holistic, pupils could complain that they have not demonstrated all

---

[37] Pupils have a right to complain on the final grade given in 10th grade if the pupil or the pupil's parent is of the opinion that the applicable regulations for grading have not been followed. Pupils may also ask for a justification of the grade (Norwegian Directorate of Education and Training, 2022c).

aspects of their competence. Without further guidelines on how to assess overall competence in the English subject, there is a risk that validity and reliability concerns arise, which could complicate the complaint process.

An interesting study looking into validity concerns with summative assessment is Black et. al.'s (2011) KOSAP study, which showed a lack of consistency between teachers' summative assessment practices and their beliefs about validity. In fact, assessment purposes rarely matched the assessment practices. However, Black et. al. (2011, p. 466) established guidelines which "would improve the quality of the internal summative assessments of many schools". This included asking the teachers to analyse the validity of their assessments, then evaluate and negotiate these with other teachers, followed by designing common assessment tasks and establishing procedures for intra- and interschool moderation. Although teachers established that they experience good collegial collaboration, including the option to have others give their perspectives on particular assessment cases, this could match what I have interpreted as a wish for the development of a shared understanding, which ultimately, would lead to more valid and reliable assessments.

In terms of reliability issues, the four teachers were very clear: Without guidelines, the subject curriculum can be operationalised in very different ways. More importantly, the assessment basis for one grade can be interpreted very differently. Given how little time these teachers report having had to formally discuss the transition from two grades to one, and the minimal support received, we need to ask: Will summative assessments produce the same results if they are scored by different people (Cameron, 2001)?

6 Conclusion

In this final chapter, I look at what conclusions can be drawn, and possible implications. First, I find it necessary to comment on the scope of this thesis, which initially did not factor in complexities with regards to the planned implementation of the curriculum change in the English subject. Nevertheless, the lack of trial projects, guidelines and supportive material on this topic explains the increase in the thesis' scope, in the sense that it required me to look deeper than anticipated to be able to provide a considerable understanding of the phenomenon. Combined with a lack of response to relevant questions from the Directorate for Education and Training[38], this has necessarily impacted the work on this thesis, as it has left some important questions unanswered. Moreover, it has also raised others: Why has this change been implemented for the English subject after only being raised in one question in a hearing while a recent reduction in grades in the Norwegian subject was preceded by a four-year trial project and three subsequent reports prior to implementation? What does this mean with regards to the status and content of the English subject? The change means the English subject now shares more similarities with foreign language instruction than the Norwegian subject, both in the allocation of hours and number of grades.

One of this thesis' main objectives has been to offer a thorough deliberation on the transition from two grades to one with the aim of assisting teachers, teacher educators, and policymakers to consider the complexities in implementing this curriculum change. Throughout the Discussion chapter, I have raised several questions, which have been left unanswered. For example, I ask why no guideline material has been produced to support teachers in differentiating between "discussing" and "interpreting" for assessment purposes, and how grading decisions should be made in instances where listening competence is demonstrated using the Norwegian language. Moreover, what are the reliability implications of the lack of guidelines? These questions are meant to serve as a thinking tool to the reader to assist reflection or discussion, and the four teachers' reflections and experiences can be seen as helpful contributions to guide conversation on these topics.

In the continuation of this chapter, I will attempt to bring together the different components of the thesis to form a conclusion. Furthermore, I elaborate on implications of this study before I make suggestions for further research.

6.1 Final summary

This thesis has sought to answer the following research question: *How do teachers experience transitioning from two grades to one in Lower Secondary English?* This thesis has found that although experiences vary, there are aspects relating to the curriculum change where the respondents share similar perspectives. All four teachers experienced few challenges relating to deciding grades for the majority of pupils, but similarly, all teachers raised similar concerns about recurring topics. Most importantly, this relates to pupils whose overall competence is not coherent, on the grounds of both validity and motivation. Moreover, this thesis has emphasised the teachers' experiences with and

---

[38] See pp. 14–15 for details.

reflections regarding weighing competences in complex assessment cases, and raised questions such as whether all skills and competences encompassed in the English subject are now equally weighed, since teachers are not supposed to view oral and written skills as separate entities which should be weighed against each other[39].

Other connections across teacher experiences were also found. The teachers agreed that time was limited to discuss and gain a shared understanding with colleagues, and that it would be beneficial to have some national guidelines. Reliability concerns regarding differing interpretations of what constitutes the assessment basis across the country were raised by all four teachers. On other topics, the teachers had deviating experiences. For example, three teachers expected the oral element to receive less focus with one grade, while one said there could be a decrease in written assessments due to these being more demanding. While the qualitative findings are not clear on other constructs in the English subject, there are signs of an increased interest in combining competences in assessment, and looking at more open assessment task designs. Furthermore, the quantitative findings show a reported increase in focus on understanding and listening from 57% and 56% of the 115 participants respectively. Nevertheless, 48% responded it was more challenging to assess overall competence than oral and written skills, while 30% reported it was easier.

This thesis has also raised questions relating to classroom effort and assessment, assessing exploration, and other topics relating to the English subject curriculum. The teachers' experiences and reflections have provided interesting contributions to understand the complex daily life of teachers when navigating curriculum change, and the intricacies of working without guidelines.

6.2 Implications

Validity and reliability concerns raised in this study indicate the need to support Lower Secondary English teachers with guidelines to ensure valid and reliable grading results across the country. The lack of centrally provided guidelines means the individual autonomy with regards to interpreting and operationalising the new English subject curriculum is disproportionately large compared to reliability considerations. This is particularly relevant in terms of defining constructs for assessment, and how to weigh competences in complex assessment contexts. The reason supportive material is emphasised is based on the collected data in this study, but also numerous hearing responses from Lower Secondary English teachers. Multiple responses from Lower Secondary schools request clear guidance or a thorough instruction for the benefit of both teachers and pupils to know how it should be interpreted into practice. In order to support English Lower Secondary teachers in making valid, reliable and fair assessments, it is vital that the authorities look into options for guidelines. These should be used *not as definitions, but as starting points for negotiating meaning[40]* to establish a shared understanding. If not, a possible result could be a great disparity in assessment validity and reliability.

Furthermore, based on the Directorate's (2022a) understanding of assessment communities, this thesis emphasises the need to support Lower Secondary English teachers to establish a shared understanding of what encompasses overall competence in the English subject, and discuss complex assessment situations such as those detailed in

---

[39] Norwegian National Centre for English and other Foreign Languages in Education (2020), quoted on page 14 and 64.
[40] Wiliam (2000, p. 14),

this thesis. This requires time set aside for *formal* conversations to cover this topic. Authorities might want to be inspired by the process preceding the reduction of term grades in the Norwegian subject (2014–2018); or the KOSAP project (2005–2007), where teachers discussed assessment validity with other teachers.

Moreover, it would be beneficial to address the tension between documentation requirements and time restraints in the Spring term. For Lower Secondary English teachers to potentially have to reduce, or plan fewer, assessment situations in the Spring term of 10th grade further complicates potential cases of complaints, particularly in light of the lack of guidelines on how to interpret English competence for grading purposes in the transition to one grade.

6.3 Suggestions for further research

There are natural limitations to a study conducted by a novice researcher, and it is possible a different conclusion would have been drawn by more experienced and connected researchers. Nevertheless, based on the results in this study, I have identified a few suggestions for further research.

For a more thorough review of the transition from two grades to one, it would be interesting to look at whether there has been a change in grades received with the current system compared to when pupils received two grades. A more qualitative approach would be looking at how teachers assess different constructs in the English subject, and how they evaluate their assessment practices in terms of weighing of competences. It would of course be interesting to compare these with previous practices, but this would add a retrospective aspect, which complicates such an approach. Moreover, this thesis could also be used as a starting point for further studies into washback effects and teachers' *assessment practices*. Skar and Hopfenbeck (2021) stated that there are almost no studies looking at actual practices within summative assessment; they are all based on teachers' interpretations. Questions which could be used as starting points include

- What do teachers in the English subject use as assessment basis when deciding on term grades / final grades?

- How do teachers collaborate on summative assessment practices in the English subject?

- How do newly educated teachers in the English subject evaluate their summative assessment practices?

Furthermore, other MA students interested in this topic might want to look at other aspects of assessment in the English subject. For example, my data material reveals that it would be very interesting to look at teachers' experiences with the new written exam. Another interesting aspect is constructs often overlooked – perhaps in the shadow of oral and written production – such as reading, listening, and interaction. During this study, I developed a particular interest in interaction assessment, which this thesis has pointed to as challenging to separate from effort, which is not to be assessed. This leads to my last suggestions for further research in the English subject didactic field, namely how teachers assess interaction without assessing their effort, and what strategies they apply to do so.

References

Aasland, B. (2009). Om å vurdere og bli vurdert. In L. Monsen, H. Bjørnsrud, L. Nyhus, B. Aasland (Eds.), *Kvalitet i skolen* (pp. 196–209). Cappelen Damm.

Brevik, L. M., Flogenfeldt, M. E. & Beiler, I. R. (2020). Livsmestring, mangfold, relevans. Engelskfaget etter fagfornyelsen. *Bedre skole, 32*(1), 46–51.

Brevik, L. M. & Rindal, U. (2020). *Teaching English in Norwegian Classrooms. From research to practice.* Universitetsforlaget.

Brevik, L. Skarpaas, K., & Isaksen, A. (2020). Vocational English: Building vocational orientation and relevance. In L. Brevik & U. Rindal (Eds.), *Teaching English in Norwegian Classrooms. From research to practice* (pp. 65–87). Universitetsforlaget.

Brinkmann, S. & Tanggard, L. (2012). Intervjuet: Samtalen som forskningsmetode. In S. Brinkmann & L. Tanggard (Eds.), *Kvalitative metoder. Empiri og teoriutvikling* (pp. 17–45). Gyldendal.

Broadfoot, P., Daugherty, R., Gardner, J., Harlen, W., James, M. Stobart, G. (2002). *Assessment for learning. 10 Principles. Research-based principles to guide classroom practice* [Brochure]. https://www.hkeaa.edu.hk/doclibrary/sba/hkdse/eng_dvd/doc/Afl_principles.pdf

Black, P., Harrison, C., Hodgen, J., Marshall, B. & Serret, N. (2011). Can teachers' summative assessments produce dependable results and also enhance classroom learning?. *Assessment in Education: Principles, Policy & Practice, 18*(4), 451–469. https://doi.org/10.1080/0969594X.2011.557020

Black, P. & Wiliam, D. (2018). Classroom Assessment and Pedagogy. *Assessment in Education Principles, Policy & Practice*, *25*(6), 551–575. https://doi.org/10.1080/0969594X.2018.1441807

Burner, T. (2020). What the 2020 curriculum reform means for English teachers. In C. Carlsen, M. Dypedahl & S. H. Iversen (Eds.), *Teaching and Learning English* (2nd Ed., pp. 53–61). Cappelen Damm Akademisk.

Bøhn, H. (2015). Assessing Spoken EFL Without a Common Rating Scale: Norwegian EFL Teachers' Conceptions of Construct. *SAGE Open, 5*(4), 1–12. https://doi.org/10.1177/2158244015621956

Bøhn, H. (2016). *What is to be assessed? Teachers' understanding of constructs in an oral English examination in Norway* [PhD]. University of Oslo.

Bøhn, H. (2019a). PhD Revisited: What is to be assessed? Teachers' understanding of

constructs in an oral English examination in Norway. In U. Rindal & L. M. Brevik (Eds.), *English Didactics in Norway – 30 years of doctoral research* (pp. 376–396). Universitetsforlaget.

Bøhn, H. (2019b). Assessing communicative skills. In T. Burner, C. Carlsen & K. Kverndokken (Eds.), *101 Ways to Work with Communicative Skills* (pp. 113–128). Fagbokforlaget.

Bøhn, H. (2020). General perspectives on assessment. In C. Carlsen, M. Dypedahl & S. H. Iversen (Eds.), *Teaching and Learning English* (2nd Ed., pp. 304–321). Cappelen Damm Akademisk.

Cameron, L. (2001). *Teaching Languages to Young Learners.* Cambridge University Press.

Carlsen, C., Dypedahl, M. & Iversen, S. H. (Eds.), *Teaching and Learning English* (2nd Ed.). Cappelen Damm Akademisk.

Chomsky, N. (1965). *Aspects of the theory of syntax.* The M.I.T. Press.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge University Press.

Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume.* Council of Europe Publishing.

Creswell, J. W. (1998). *Qualitative Inquiry and Research Design. Choosing Among Five Traditions* (3rd Ed.). Sage Publications.

Dahler Larsen, P. (2012). Evaluering. In S. Brinkmann & L. Tanggard (Eds.), *Kvalitative metoder. Empiri og teoriutvikling* (pp. 186–206). Gyldendal.

Dobson, S. & Engh, R. (2010). Vurdering i de enkelte fag – likheter og forskjeller. In S. Dobson & R. Engh (Eds.), *Vurdering for læring i fag* (pp. 11–23). Høyskoleforlaget.

Easterby-Smith, M., Thorpe, R. & Lowe, A. (1991). *Management Research. An Introduction.* SAGE Publications.

Eggen, A. B. (2010). Vurdering og validitet. In S. Dobson & R. Engh (Eds.), *Vurdering for læring i fag* (pp. 40–56). Høyskoleforlaget.

Engen, T. O. (1996). Elevvurderingas funksjoner. In Skagen, K. (Ed.), *Karakterboka. Om karakterer og vurdering i ny skole* (pp. 55–79). Universitetsforlaget.

Engh, R. (2010). Nasjonale standarder og kjennetegn på måloppnåelse. In S. Dobson & R. Engh (Eds.), *Vurdering for læring i fag* (pp. 57–74). Høyskoleforlaget.

Engh, R. (2012). Ask & Embla og satsningen på vurdering for læring i Norge. *CEPRA-Striben, (*12), 26–33. https://doi.org/10.17896/UCN.cepra.n12.88

Eriksen, A. G. (2021). Oral assessment in the English language [Master's thesis]. UiT The Arctic University of Norway.

Eriksen, H. & Elstad, E. (2019). Norsklæreres erfaringer med karakterfritt semester: rom

    for forbedringsinnsats hos eleven og større fokus på læringsfremmende
tilbakemelding fra læreren?. *Nordisk tidsskrift for utdanning og praksis, 13*(2), 4–
20. https://doi.org/10.23865/up.v13.1968

Fenner, A. & Skulstad, A. S. (2020). Introduction. In A. Fenner & A. S. Skulstad (Eds.),

    Teaching English in the 21st Century. Central issues in English didactics, (2nd Ed.,
pp. 7–16). Fagbokforlaget.

Finlay, L. (2014). Engaging Phenomenological Analysis. *Qualitative Research in*

    *Psychology, 11(*2), 121–141. https://doi.org/10.1080/14780887.2013.807899

Fjørtoft, H. & Sandvik, L. V. (2016). Innledning. Vurdering og vurderingskompetanse. In

    H. Fjørtoft & L. V. Sandvik (Eds.): *Vurderingskompetanse i skolen. Praksis, læring
og utvikling* (pp. 17–41). Universitetsforlaget.

Gardner, J. (2012). Assessment and learning. An introduction. In J. Gardner (Ed.),

    Assessment and Learning (pp. 1–7). SAGE Publications Ltd.
https://dx.doi.org/10.4135/9781446250808

Gleiss, M. S.. & Sæther, E. (2021). *Forskningsmetode for masterstudenter. Å utvikle ny*

    kunnskap i forskning og praksis. Cappelen Damm Akademisk.

Hasselgreen, A. & Ørevik, S. (2020). Assessment and testing. In A. Fenner & A. S.

    Skulstad (2020): *Teaching English in the 21st Century. Central issues in English
didactics* (2nd Ed., pp. 365–388). Fagbokforlaget.

Haukås, Å. & Speitz, H. (2020). Plurilingual learning and teaching. In C. Carlsen, M.

    Dypedahl & S. H. Iversen (Eds.), *Teaching and Learning English* (2nd Ed., pp. 62–
80). Cappelen Damm Akademisk.

Hellaway, D. (2006). Inside-out: Analysis of the insider-outsider concept as a heuristic

    device to develop reflexivity in students doing qualitative research. *Teaching in
Higher Education, 11(*4), 483–494. https://doi.org/10.1080/13562510600874292

Hestetræet, T. I. & Ørevik, S. (2020). English Vocational Studies. In A. Fenner & A. S.

    Skulstad (Eds), *Teaching English in the 21st Century. Central issues in English
didactics* (2nd Ed., pp. 311–337). Fagbokforlaget.

Ibsen, E. (2004). *Engelsk i Europa – 2002. Norsk rapport fra en europeisk*

    *engelskundersøkelse om holdninger til og ferdigheter i engelsk ved utgangen av
den obligatoriske grunnskolen.* University of Oslo. http://urn.nb.no/URN:NBN:no-
14504

Johnson, B. & Christensen, L. (2012). *Educational Research: Quantitative, Qualitative,*

    and Mixed Approaches (4th Ed.). Sage.

Karseth, B., Kvamme, O. A. & Ottesen, E. (2020). *Fagfornyelsens læreplanverk. Politiske*

    intensjoner, arbeidsprosesser og innhold. University of Oslo.
https://www.uv.uio.no/forskning/prosjekter/fagfornyelsen-
evaluering/publikasjoner/eva2020-delrapport-1.pdf

Klapp, A. (2015). Does grading affect educational attainment? *Assessment in Education.*

Principles, Policy & Practice. 22(3), 302–323.
https://doi.org/10.1080/0969594X.2014.988121

Kvarv, S. (2014). Vitenskapsteori – tradisjoner, posisjoner og diskusjoner. Novus forlag.

Kvernbekk, T. (2005). Pedagogisk teoridannelse: Insidere, teoriformer og praksis.

Fagbokforlaget.

Langseth, I. (2009). Elevvurdering i fremmedspråk. In S. Dobson, A. B. Eggen & K.

Smith (Eds), Vurdering, prinsipper og praksis. Nye perspektiver på elev- og
læringsvurdering (pp. 256–279). Gyldendal Akademisk.

Langseth, I. (2016). Elevinvolvering i engelskfaget? In H. Fjørtoft & L. Sandvik (Eds.),

Vurderingskompetanse i skolen. Praksis, læring og utvikling (pp. 139–154).
Universitetsforlaget.

NESH (2022). Guidelines for Research Ethics in the Social Sciences and the Humanities.

(5th Ed.) The Norwegian National Research Ethics Committees.
https://www.forskningsetikk.no/globalassets/dokumenter/4-publikasjoner-som-
pdf/guidelines-for-research-ethics-in-the-social-sciences-and-the-humanities.pdf

North, B. (2022). Completing the CEFR descriptive scheme: the CEFR Companion

Volume. [PowerPoint slides]. Council of Europe. https://rm.coe.int/completing-
the-cefr-descriptive-scheme-north/16808ae70c

NOU (2015:8). (2015). Fremtidens skole. Fornyelse av fag og kompetanser. Ministry of

Education and Research.
https://www.regjeringen.no/contentassets/da148fec8c4a4ab88daa8b677a700292
/no/pdfs/nou201520150008000dddpdfs.pdf

Nordahl, T. & Hansen, O. (eds.) (2011). Dette vet vi om vurderingspraksis. Gyldendal

Akademisk.

Nordgren, K. & Odenstad, C. (2017). Betyg och ämnesdidaktisk bedömning. In K.

Nordgren, C. Odenstad & J. Samuelsson (Eds.), Betyg i teori och praktik (3rd Ed.,
pp. 9–20). Gleerups.

Normann, A. (2017). Becoming a phenomenologically skilled researcher – an account of

a journey started. Reflective Practice. 18(5), 613–626.
https://doi.org/10.1080/14623943.2017.1307720

Norwegian Directorate for Education and Training (2009). Sluttrapport. Oppdragsbrev Nr.

6 – 2007 om tiltak knyttet til individvurdering i skole og fag- og yrkesopplæring.
https://www.udir.no/globalassets/filer/tall-og-
forskning/rapporter/2009/5/bedre_vurderingspraksis_sluttrapport_til_kd.pdf

Norwegian Directorate for Education and Training (2013). English subject curriculum.

(No. ENG1-3). http://data.udir.no/kl06/ENG1-03.pdf?lang=eng

Norwegian Directorate for Education and Training (2018a). Observations on the national

Assessment for Learning Programme (2010–2018). Skills development in
networks. Final report 2018.

https://www.udir.no/contentassets/596e196f7b9d459e96e64ecc4b463faa/the-norwegian-assessment-for-learning-programme_final-report-2018.pdf

Norwegian Directorate for Education and Training (2018b, 10th October). *Sentrale begreper i læreplanarbeidet*. https://www.udir.no/laring-og-trivsel/lareplanverket/forsok-og-pagaende-arbeid/Retningslinjer-for-utforming-av-lareplaner-for-fag-/vedlegg/

Norwegian Directorate for Education and Training (2018c, 10th October). *Fagspesifikk omtale av vurdering i fag. Retningslinjer for utforming av læreplaner for fag i LK20 og LK20s.* https://www.udir.no/laring-og-trivsel/lareplanverket/forsok-og-pagaende-arbeid/Retningslinjer-for-utforming-av-lareplaner-for-fag-/prinsipper-for-utforming-av-lareplaner-for-fag/#vurdering

Norwegian Directorate for Education and Training (2019a, 18th March). *Høring – læreplan i engelsk.* [Hearing] https://hoering.udir.no/Hoering/v2/338

Norwegian Directorate for Education and Training (2019b, unpublished). *Høringsoppsummering – læreplan i engelsk fellesfag.*

Norwegian Directorate for Education and Training (2019d, 18th November). *Hva er kjerneelementer?* https://www.udir.no/laring-og-trivsel/lareplanverket/stotte/hva-er-kjerneelementer/

Norwegian Directorate for Education and Training (2020a). *Læreplan i engelsk* (No. ENG01-04). https://data.udir.no/kl06/v201906/laereplaner-lk20/ENG01-04.pdf

Norwegian Directorate for Education and Training (2020c, 16th June). *Skal vi vurdere mer enn faglig kompetanse i framtiden?* https://www.udir.no/eksamen-og-prover/eksamen/vurderinger-og-anbefalinger-fremtidens-eksamen/12-skal-vi-vurdere-mer-enn-faglig-kompetanse-i-framtiden/

Norwegian Directorate for Education and Training (2021a, 27th May). *Kunnskapsgrunnlag i engelsk.* https://www.udir.no/kvalitet-og-kompetanse/laremidler/kvalitetskriterier-for-laremidler/kunnskapsgrunnlag-kvalitetskriterium-engelsk/

Norwegian Directorate for Education and Training (2021b, 15th February). *Kvalitetskriterier for utvikling av eksamen. Rammeverk for eksamen – LK20 og LK20s.* https://www.udir.no/eksamen-og-prover/eksamen/rammeverk-skriftlig-eksamen-i-lk20-og-lk20s/4.-kvalitetskriterier-for-utvikling-av-eksamen/

Norwegian Directorate for Education and Training (2022a, 16th March). *Tolkningsfellesskap om standpunktvurdering.* https://www.udir.no/laring-og-trivsel/vurdering/om-vurdering/tolkningsfellesskap-om-standpunktvurdering/

Norwegian Directorate for Education and Training (2022b, 2nd March): *Hva betyr den samla kompetansen?* https://www.udir.no/laring-og-trivsel/vurdering/sluttvurdering/standpunktvurdering/#a173400

Norwegian Directorate for Education and Training (2022c, 23rd May). *Hva kan eleven eller foreldrene klage på?* https://www.udir.no/regelverkstolkninger/opplaring/Vurdering/behandling-av-

klager-pa-standpunktkarakterer-i-fag/1.-hva-kan-eleven-eller-foreldrene-klage-pa/

Norwegian Ministry of Church and Education (1987). *Mønsterplan for grunnskolen. M87.* Aschehoug.
https://www.nb.no/nbsok/nb/2aef891325a059851965d5b8ac193de5#0

Norwegian Ministry of Church, Education and Research (1993). *Læreplan for videregående opplæring. Engelsk. Felles allment fag for alle studieretninger.*
https://www.udir.no/globalassets/upload/larerplaner/felles-allmenne-fag/5/lareplan_engelsk.rtf

Norwegian Ministry of Church, Education and Research (1998). *Elevvurdering. Rettleiing til Læreplanverket for den 10-årige grunnskolen (L97 og L97S).*
https://www.nb.no/nbsok/nb/e8793198ac2784be99c68abd26e6a780?lang=no#3

Norwegian National Centre for English and other Foreign Languages in Education (2020). Spørsmål og svar om nye læreplaner i engelsk og fremmedspråk i LK20. *Communicare 2019/20,* 4–7.
https://www.hiof.no/fss/om/publikasjoner/communicare/2019/2019-20-communicare-nett.pdf

Nusche, D., Earl, L., Maxwell, W. & Shewbridge, C. (2011, September). *OECD Reviews of Evaluation and Assessment in Education. Norway.* OECD.
https://www.oecd.org/norway/48632032.pdf

Meld. St. 28 (2015–2016). *Fag – fordypning – forståelse – en fornyelse av Kunnskapsløftet.* Kunnskapsdepartementet.
https://www.regjeringen.no/no/dokumenter/meld.-st.-28-20152016/id2483955/

Moss, C. M. (2013). Research on classroom summative assessment. In McMillan, J. H. (2013). *Sage Handbook of Research on Classroom Assessment* (pp. 235–255)*.* SAGE Publications. https://dx.doi.org/10.4135/9781452218649.n14

Myhr, R. O. (2014). Den egenrådige læreren. En drøfting av forholdet mellom tilpasset opplæring og lærerens selvstendighet. I C. R. Haugen & T. A. Hestbek (Eds.). *Pedagogikk, politikk og etikk. Demokratiske utfordringer og muligheter i norsk skole* (pp. 159–172). Universitetsforlaget.

Postholm, M. B. & Jacobsen, D. I. (2018). *Forskningsmetode for masterstudenter i lærerutdanningen.* Cappelen Damm Akademisk.

Prøitz, T. S. & Borgen, J. S. (2010). *Rettferdig standpunktvurdering – det (u)muliges kunst? Læreres setting av standpunktkarakter i fem fag i grunnutdanningen.* [NIFU STEP 16/2010]. http://hdl.handle.net/11250/279131

Rambøll Management Consulting (2020, October). *Vurdering i skolen.* Utdanningsforbundet. https://www.utdanningsforbundet.no/globalassets/var-politikk/publikasjoner/eksterne-rapporter/rapport_ramboll_vurdering-i-skolen_okt_2020.pdf

Regulations to The Education Act. (2020). *Kapittel 3. Individuell vurdering i grunnskolen*

og videregåande opplæring. (FOR-2006-06-23-724). Lovdata.
https://lovdata.no/dokument/SF/forskrift/2006-06-23-724/

Rindal U. & Brevik, L. M. (2019). Introduction. In U. Rindal & L. M. Brevik (Eds.), *English Didactics in Norway – 30 years of doctoral research* (pp. 9–15). Universitetsforlaget.

Rimmereide, H. E. (2020). Multimodal texts in the English classroom. In C. Carlsen, M. Dypedahl & S. H. Iversen (Eds). *Teaching and Learning English* (2nd Ed., pp. 191–208). Cappelen Damm Akademisk.

Robinson, V. (2018). *Færre endringer – mer utvikling.* Cappelen Damm Akademisk.

Rødnes, K. A. & de Lange, T. (2012). *Læreres bruk av og erfaringer med veiledninger til læreplaner for fag og veiledning i lokalt arbeid med læreplaner.* University of Oslo. https://www.udir.no/globalassets/filer/tall-og-forskning/forskningsrapporter/lareres-bruk-av-og-erfaringer-med-veiledninger-til-lareplaner-for-fag-og-veiledning-i-lokalt-arbeid-med-lareplaner.pdf

Sandvik, L. V., Engvik, G., Fjørtoft, H., Langseth, I. D. Aaslid, B. E., Mordal, S. & Buland, T. (2012). *Vurdering i skolen. Intensjoner og forståelser. Delrapport 1 fra prosjektet Forskning på individuell vurdering i skolen (FIVIS).* NTNU, Program for lærerutdanning.

Sandvik, L. V. (2022, 8th November). *Mindre vekt på karakterbruk betyr at lærerne må undervise på nye måter.* Utdanningsnytt. https://www.utdanningsnytt.no/bedre-skole-fagartikkel-karakterer/mindre-vekt-pa-karakterbruk-betyr-at-laererne-ma-undervise-pa-nye-mater/334488

Seland, I., Hovdhaugen, E., Lødding, B., Prøitz, T. S. & Rønsen, E. (2018). *Sluttrapport fra evaluering av forsøk med halvårsvurdering med én eller to karakterer i norsk.* [NIFU 2018:9]. http://hdl.handle.net/11250/2503428

Selghed, B. (2011). Ett omöjligt oppdrag? – om lärares bedömningar och betygsättning. In G. Brante, K. Hjort (Eds.). *Dilemman i skolan: aktuella utmaningar och professionella omställningar (*pp. 82–95). Kristianstad University Press.

Sivenbring, J. (2018). Making sense and use of assessments. *Scandinavian Journal of Educational Research, 63*(5), 759–770. https://doi.org/10.1080/00313831.2018.1434827

Skar, G. B. & Hopfenbeck, T. N. (2021). *Kunnskapsgrunnlag om standpunktvurdering.* NTNU (The Norwegian Centre for Writing Education and Research). https://hdl.handle.net/11250/2831020

Skulstad, A. S. (2020). Communicative competence. In A. Fenner & A. S. Skulstad (2020): *Teaching English in the 21st Century. Central issues in English didactics* (2nd Ed., pp. 43–68). Fagbokforlaget.

Smith, J. A. & Osborn, M. (2003). Interpretative phenomenological analysis. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods* (pp. 51–80). Sage Publications, Inc.

Steffensen, B. (2003). *Det fagdidaktiske projekt – almen fagdidaktik set i lyset af viden,*

dannelse og læring. Akademisk forlag.

Speitz, H. (2020). National curricula and international perspectives. In C. Carlsen, M.

Dypedahl & S. H. Iversen (Eds). *Teaching and Learning English* (2nd Ed., pp 40–52). Cappelen Damm Akademisk.

Teigar Lower Secondary (2019). *Høring – læreplaner i engelsk. Uttalelse – Teigar U.*

[Hearing response]. https://hoering.udir.no/Uttalelse/v2/a13dca1e-8ad6-44fe-83a7-8244b5b78bf9?disableTutorialOverlay=True

Thrane, C. (2018). *Kvantitativ metode. En praktisk tilnærming.* Cappelen Damm

Akademisk.

Thurén, T. (2012). *Vitenskapsteori for nybegynnere* (2nd Ed.). Gyldendal Akademisk.

Van Manen, M. & Van Manen, M. (2021). Doing Phenomenological Research and Writing.

Qualitative Health Research, 31(6), 1069–1082.
https://doi.org/10.1177/10497323211003058

Vitello, S., Greatorex, J. & Shaw, S. (2021). *What is competence? A shared interpretation*

of competence to support teaching, learning and assessment. Cambridge
University Press & Assessment.

Westera, W. (2001). Competence in Education: A confusion of tongues. *Journal of*

Curriculum Studies. 33(1), 75–88. https://doi.org/10.1080/00220270120625

Wiliam, D. (2000, November). *Integrating formative and summative functions of*

assessment. [Keynote address]. European Association for Educational
Assessment, Prague, Czech Republic.
https://discovery.ucl.ac.uk/id/eprint/10001151/1/Wiliam2000IntergratingAEA-E_2000_keynoteaddress.pdf

Wyatt-Smith, C., Klenowski, V. & Gunn, S. (2010). The centrality of teachers' judgement

practice in assessment: a study of standards in moderation. *Assessment in
Education: Principles, Policy & Practice, 17*(1), 59–75.
https://doi.org/10.1080/09695940903565610

Appendices

Appendix A. Master's Agreement

Masteravtale/hovedoppgaveavtale

Sist oppdatert 11. november 2020

Fakultet	**Fakultet for samfunns- og utdanningsvitenskap**
Institutt	**Institutt for lærerutdanning**
Studieprogram	**MDID**
Emnekode	**DID3930**

Studenten

Etternavn, fornavn	**Langfjæran, Karen Elise Wiik**
Fødselsdato	**10.12.1991**
E-postadresse ved NTNU	**karenel@stud.ntnu.no**

Tilknyttede ressurser

Veileder	**Anita Normann**
Eventuelle medveiledere	
Eventuelle medstudenter	

Oppgaven

Oppstartsdato	**01.09.2021**
Leveringsfrist	**25.11.2022**
Oppgavens arbeidstittel	**Transitioning from Two Grades to One Grade in Lower Secondary English Education**
Problembeskrivelse	**Starting Autumn 2021, the new school reform LK20 introduces a new strategy for summative assessment in Lower Secondary English, abandoning the two grade system (one grade to represent oral skills and another for written skills) in favour of one grade to reflect a broader competence. UDIR explains the reason for this being the number of hours assigned to English in Lower Secondary as well as good experience from VGS, but some of the hearing feedback from relevant parties such as schools and teachers, reflect on possible negative consequences. This research project will research the implications that this change in summative assessment has on teachers' assessment and teaching practice as well as the perception of assessment culture. Using data from three rounds of semi-structural interviews and surveys following teachers through the first year of introducing the change, the project will aim to answer the main research question as following: In what ways does the new summative assessment in Lower Secondary English impact teachers' assessment and teaching practice?**

Risikovurdering og datahåndtering	
Skal det gjennomføres risikovurdering?	Nei
Dersom «ja», har det blitt gjennomført?	Nei
Skal det søkes om godkjenninger? (REK*, NSD)**	Ja
Skal det skrives en konfidensialitetsavtale i forbindelse med oppgaven?	Nei
Hvis «ja», har det blitt gjort?	Nei

* Regionale komiteer for medisinsk og helsefaglig forskningsetikk (https://rekportalen.no)

** Norsk senter for forskningsdata (https://nsd.no/)

Eventuelle emner som skal inngå i mastergraden

Retningslinjer - rettigheter og plikter

Formål

Avtale om veiledning av masteroppgaven/hovedoppgaven er en samarbeidsavtale mellom student, veileder og institutt. Avtalen regulerer veiledningsforholdet, omfang, art og ansvarsfordeling.

Studieprogrammet og arbeidet med oppgaven er regulert av Universitets- og høgskoleloven, NTNUs studieforskrift og gjeldende studieplan. Informasjon om emnet, som oppgaven inngår i, finner du i emnebeskrivelsen.

Veiledning

Studenten har ansvar for å

- Avtale veiledningstimer med veileder innenfor rammene master-/hovedoppgaveavtalen gir.
- Utarbeide framdriftsplan for arbeidet i samråd med veileder, inkludert veiledningsplan.
- Holde oversikt over antall brukte veiledningstimer sammen med veileder.
- Gi veileder nødvendig skriftlig materiale i rimelig tid før veiledning.
- Holde instituttet og veileder orientert om eventuelle forsinkelser.
- Inkludere eventuell(e) medstudent(er) i avtalen.

Veileder har ansvar for å

- Avklare forventninger om veiledningsforholdet.
- Sørge for at det søkes om eventuelle nødvendige godkjenninger (etikk, personvernhensyn).
- Gi råd om formulering og avgrensning av tema og problemstilling, slik at arbeidet er gjennomførbart innenfor normert eller avtalt studietid.
- Drøfte og vurdere hypoteser og metoder.
- Gi råd vedrørende faglitteratur, kildemateriale, datagrunnlag, dokumentasjon og eventuelt ressursbehov.
- Drøfte framstillingsform (eksempelvis disposisjon og språklig form).
- Drøfte resultater og tolkninger.
- Holde seg orientert om progresjonen i studentens arbeid i henhold til avtalt tids- og arbeidsplan, og følge opp studenten ved behov.
- Sammen med studenten holde oversikt over antall brukte veiledningstimer.

Instituttet har ansvar for å

- Sørge for at avtalen blir inngått.
- Finne og oppnevne veileder(e).
- Inngå avtale med annet institutt/ fakultet/institusjon dersom det er oppnevnt ekstern medveileder.
- I samarbeid med veileder holde oversikt over studentens framdrift, antall brukte veiledningstimer, og følge opp dersom studenten er forsinket i henhold til avtalen.
- Oppnevne ny veileder og sørge for inngåelse av ny avtale dersom:
  - Veileder blir fraværende på grunn av eksempelvis forskningstermin, sykdom, eller reiser.
  - Student eller veileder ber om å få avslutte avtalen fordi en av partene ikke følger den.
  - Andre forhold gjør at partene finner det hensiktsmessig med ny veileder.
- Gi studenten beskjed når veiledningsforholdet opphører.
- Informere veileder(e) om ansvaret for å ivareta forskningsetiske forhold, personvernhensyn og veiledningsetiske forhold.
- Ønsker student, eller veileder, å bli løst fra avtalen må det søkes til instituttet. Instituttet må i et slikt tilfelle oppnevne ny veileder.

Avtaleskjemaet skal godkjennes når retningslinjene er gjennomgått.

Godkjent av

Karen Elise Wiik Langfjæran
Student

30.09.2021
Digitalt godkjent

Anita Normann
Veileder

30.09.2021
Digitalt godkjent

Mari Linna Mosebekk Aglen
Institutt

24.11.2021
Digitalt godkjent

Master`s Agreement / Main Thesis Agreement

Faculty	Faculty of Social and Educational Sciences
Institute	Department of Teacher Education
Programme Code	MDID
Course Code	DID3930

Personal Information

Surname, First Name	Langfjæran, Karen Elise Wiik
Date of Birth	10.12.1991
Email	karenel@stud.ntnu.no

Supervision and Co-authors

Supervisor	Anita Normann
Co-supervisors (if applicable)	
Co-authors (if applicable)	

The Master`s thesis

Starting Date	01.09.2021
Submission Deadline	25.11.2022
Thesis Working Title	Transitioning from Two Grades to One Grade in Lower Secondary English Education
Problem Description	Starting Autumn 2021, the new school reform LK20 introduces a new strategy for summative assessment in Lower Secondary English, abandoning the two grade system (one grade to represent oral skills and another for written skills) in favour of one grade to reflect a broader competence. UDIR explains the reason for this being the number of hours assigned to English in Lower Secondary as well as good experience from VGS, but some of the hearing feedback from relevant parties such as schools and teachers, reflect on possible negative consequences. This research project will research the implications that this change in summative assessment has on teachers' assessment and teaching practice as well as the perception of assessment culture. Using data from three rounds of semi-structural interviews and surveys following teachers through the first year of introducing the change, the project will aim to answer the main research question as following: In what ways does the new

	summative assessment in Lower Secondary English impact teachers' assessment and teaching practice?

Risk Assessment and Data Management

Will you conduct a Risk Assessment?	No
If "Yes", Is the Risk Assessment Conducted?	No
Will you Apply for Data Management? **(REK*, NSD**)**	Yes
Will You Write a Confidentiality Agreement?	No
If "Yes", Is the Confidentiality Agreement Conducted?	No

* REK -- https://rekportalen.no/

** Norwegian Centre for Research Data (https://nsd.no/nsd/english/index.html)

Topics to be included in the Master`s Degree (if applicable)

Guidelines – Rights and Obligations

Purpose

The Master's Agreement/ Main Thesis Agreement is an agreement between the student, supervisor, and department. The agreement regulates supervision conditions, scope, nature, and responsibilities concerning the thesis.

The study programme and the thesis are regulated by the Universities and University Colleges Act, NTNU's study regulations, and the current curriculum for the study programme.

Supervision

The student is responsible for

- Arranging the supervision within the framework provided by the agreement.
- Preparing a plan of progress in cooperation with the supervisor, including a supervision schedule.
- Keeping track of the counselling hours.
- Providing the supervisor with the necessary written material in a timely manner before the supervision.
- Keeping the institute and supervisor informed of any delays.
- Adding fellow student(s) to the agreement, if the thesis has more than one author.

The supervisor is responsible for

- Clarifying expectations and how the supervision should take place.
- Ensuring that any necessary approvals are acquired (REC, ethics, privacy).
- Advising on the demarcation of the topic and the thesis statement to ensure that the work is feasible within agreed upon time frame.
- Discussing and evaluating hypotheses and methods.
- Advising on literature, source material, data, documentation, and resource requirements.
- Discussing the layout of the thesis with the student (disposition, linguistic form, etcetera).
- Discussing the results and the interpretation of them.
- Staying informed about the work progress and assist the student if necessary.
- Together with the student, keeping track of supervision hours spent.

The institute is responsible for

- Ensuring that the agreement is entered into.
- Find and appoint supervisor(s).
- Enter into an agreement with another department / faculty / institution if there is an external co-supervisor.
- In cooperation with the supervisor, keep an overview of the student's progress, the number of supervision hours. spent, and assist if the student is delayed by appointment.
- Appoint a new supervisor and arrange for a new agreement if:
  - The supervisor will be absent due to research term, illness, travel, etcetera.
  - The student or supervisor requests to terminate the agreement due to lack of adherence from either party.
  - Other circumstances where it is appropriate with a new supervisor.
- Notify the student when the agreement terminates.
- Inform supervisors about the responsibility for safeguarding ethical issues, privacy and guidance ethics
- Should the cooperation between student and supervisor become problematic, either party may apply to the department to be freed from the agreement. In such occurrence, the department must appoint a new supervisor

This Master`s agreement must be signed when the guidelines have been reviewed.

Signatures

Karen Elise Wiik Langfjæran
Student

30.09.2021
Digitally approved

Anita Normann
Supervisor

30.09.2021
Digitally approved

Mari Linna Mosebekk Aglen
Department

24.11.2021
Digitally approved

Appendix B. Approval from the Norwegian Centre for Research Data

Meldeskjema / Transitioning from Two Grades to One Grade in Lower Secondary En... / Vurdering

Vurdering

Referansenummer	Type	Dato
894223	Standard	10.09.2021

Prosjekttittel
Transitioning from Two Grades to One Grade in Lower Secondary English Education

Behandlingsansvarlig institusjon
Norges teknisk-naturvitenskapelige universitet / Fakultet for samfunns- og utdanningsvitenskap (SU) / Institutt for lærerutdanning

Prosjektansvarlig
Anita Normann

Student
Karen Langfjæran

Prosjektperiode
01.09.2021 – 20.06.2023

Kategorier personopplysninger
Alminnelige

Rettslig grunnlag
Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

Behandlingen av personopplysningene kan starte så fremt den gjennomføres som oppgitt i meldeskjemaet. Det rettslige grunnlaget gjelder til 20.06.2023.

Meldeskjema ↗

Kommentar
Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 10.9.2021, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 20.6.2023.

LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake.

Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:
· lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
· formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
· dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
· lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), og dataportabilitet (art. 20).

NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

MELD VESENTLIGE ENDRINGER
Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde: https://www.nsd.no/personverntjenester/fylle-ut-meldeskjema-for-personopplysninger/melde-endringer-i-meldeskjema

Du må vente på svar fra NSD før endringen gjennomføres.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

98

Vurdering

Referansenummer   **Type**   **Dato**
894223   Standard   19.01.2022

Prosjekttittel
Transitioning from Two Grades to One Grade in Lower Secondary English Education

Behandlingsansvarlig institusjon
Norges teknisk-naturvitenskapelige universitet / Fakultet for samfunns- og utdanningsvitenskap (SU) / Institutt for lærerutdanning

Prosjektansvarlig
Anita Normann

Student
Karen Langfjæran

Prosjektperiode
01.09.2021 - 20.06.2023

Kategorier personopplysninger
Alminnelige

Rettslig grunnlag
Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

Behandlingen av personopplysningene kan starte så fremt den gjennomføres som oppgitt i meldeskjemaet. Det rettslige grunnlaget gjelder til 20.06.2023.

Meldeskjema 🗗

Kommentar
Personverntjenester har vurdert endringen registrert 11.1.2022.

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 19.1.2022. Behandlingen kan fortsette.

Det er registrert en ny datainnsamling for utvalg 1. Intervjuguide er lastet opp i meldeskjema.

OPPFØLGING AV PROSJEKTET
Vi vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Kontaktperson: Håkon J. Tranvåg

Lykke til videre med prosjektet!

Vurdering

Referansenummer	**Type**	**Dato**
894223	Standard	13.06.2022

Prosjekttittel
Transitioning from Two Grades to One Grade in Lower Secondary English Education

Behandlingsansvarlig institusjon
Norges teknisk-naturvitenskapelige universitet / Fakultet for samfunns- og utdanningsvitenskap (SU) / Institutt for lærerutdanning

Prosjektansvarlig
Anita Normann

Student
Karen Langfjæran

Prosjektperiode
01.09.2021 - 20.06.2023

Kategorier personopplysninger
Alminnelige

Rettslig grunnlag
Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

Behandlingen av personopplysningene kan starte så fremt den gjennomføres som oppgitt i meldeskjemaet. Det rettslige grunnlaget gjelder til 20.06.2023.

Meldeskjema ↗

Kommentar
Personverntjenester har vurdert endringen registrert 25.5.2022.

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 19.1.2022. Behandlingen kan fortsette.

Det er registrert en ny datainnsamling for utvalg 1. Oppdatert intervjuguide er lastet opp i meldeskjema.

OPPFØLGING AV PROSJEKTET
Vi vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til videre med prosjektet!

100

Vil du delta i forskningsprosjektet

"Transitioning from Two Grades to One Grade in English Lower Secondary Education"?

Dette er et spørsmål om å delta i et forskningsprosjekt hvor formålet er å undersøke hvilke erfaringer lærere gjør seg knyttet til ny vurderingsordning i engelsk på ungdomsskolen, som innføres i tråd med LK20 høsten 2021. I dette skrivet finner du informasjon om målene for prosjektet og hva deltakelse innebærer for deg.

Formål

Ny vurderingsordning i engelsk på ungdomsskolen innebærer å gå fra to karakterer (muntlig og skriftlig) til én samlet karakter. Bakgrunnen for ny ordning er god erfaring fra videregående samt timetallet i faget[1]. Høringsuttalelser fra lærere, skoler, fagorganer, kommuner og andre viser at vurderingen er delt knyttet til hvorvidt ny vurderingsordning er positivt eller negativt[2].

Dette forskningsprosjektet skal følge tre lærere i første år med ny vurderingsordning gjennom intervju i tre runder – i første skolehalvår (oktober), etter halvårsvurdering (januar) og standpunktssetting (mai/juni). Det skal dessuten gjennomføres samtidige, anonyme spørreundersøkelser.

Hovedproblemstillingen for forskningsprosjektet er «Hvordan opplever ungdomsskolelærere at ny vurderingsordning i engelsk påvirker vurderingspraksis og undervisning?».

Dette er en masteroppgave påbegynt høsten 2021, som er planlagt fullført høsten 2022. Oppgaven skrives på engelsk i forbindelse med et masterløp i fagdidaktikk engelsk ved NTNU, Institutt for lærerutdanning.

Opplysningene gitt i intervju skal ikke brukes til andre formål.

Hvem er ansvarlig for forskningsprosjektet?

NTNU er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Jeg henvender meg til deg som engelsklærer i ungdomsskolen. Hvis du takker ja til å bli med på prosjektet, blir du én av tre lærere jeg følger gjennom dette året.

---

[1]UDIR, 2019a: Høring – læreplan i engelsk. Hentet fra https://hoering.udir.no/Hoering/v2/338
[2]UDIR, 2019b: Høringsuttalelser – læreplan i engelsk. Hentet fra https://hoering-publisering.udir.no/338/uttalelser

Hva innebærer det for deg å delta?

Jeg kommer til å gjennomføre intervju ved tre anledninger. Intervjuene gjennomføres fysisk og med lydopptak i oktober 2021, januar 2022 og mai/juni 2022 (med forbehold, telefon-/videointervju kan vurderes ved behov). Vi kan sammen avgjøre når og hvor dette finner sted.

Opplysninger som samles inn, anonymiseres i oppgaven, og ingen personopplysninger vurderes som viktige for oppgaven. Jeg henviser til kjønn, erfaring i skolen (lang/kort), strukturelle betingelser som klassestørrelse, skolestørrelse og kommunestørrelse (stor/liten kommune), men utover dette anonymiseres du fullstendig. Du kan dessuten alltid be om innsyn (se Dine rettigheter under).

Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger blir da slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

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

Disse vil ha tilgang til opplysningene:
- Masterstudent Karen Langfjæran
- Min veileder, førstelektor Anita Normann

Navn og kontaktopplysninger utelukkes så langt det er mulig i opptak og holdes atskilt fra samt anonymiseres i transkribert materiale og øvrige data. Ingen personopplysninger lagres sammen med datamaterialet. Da forskningsprosjektet kun omfatter tre intervjuobjekter er det svært enkelt å holde informasjonen atskilt. Du skal ikke kunne gjenkjennes i publisert materiale.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er november 2022. Personopplysninger og opptak slettes ved prosjektslutt.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

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

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:
- innsyn i hvilke opplysninger vi behandler om deg, og å få utlevert en kopi av opplysningene
- å få rettet opplysninger om deg som er feil eller misvisende
- å få slettet personopplysninger om deg
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger

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

- Masterstudent Karen Langfjæran, █████████████████████████
- NTNU ved førstelektor Anita Normann, ██████████████████
- NTNUs personvernombud: Thomas Helgesen ved NTNU, █████████████████
  ████████████████

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:

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

Med vennlig hilsen

Karen Langfjæran
Masterstudent ved NTNU

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om masterprosjektet «Transitioning from Two Grades to One Grade in English Lower Secondary Education» og har fått anledning til å stille spørsmål.

Jeg samtykker med dette til å delta i intervju og at mine opplysninger behandles frem til prosjektet er avsluttet.

(Signert av prosjektdeltaker, dato)

Appendix D. Information letter for questionnaire participants (per email)

<div align="center">

Invitasjon til deltakelse i forskningsprosjektet

Transitioning from Two Grades to One Grade in Lower Secondary English Education

</div>

Jeg sender deg dette korte informasjonsskrivet med det håp om at du kan videreformidle et spørreskjema til aktuelle lærere ved din skole *(engelsklærere ved 8., 9. og 10. trinn)* som kan bidra til et forskningsprosjekt som ser på læreres erfaring med ny vurderingsordning i engelsk på ungdomsskolen.

Ny vurderingsordning i engelsk på ungdomsskolen innebærer å gå fra to karakterer (muntlig og skriftlig) til én samlet karakter. Bakgrunnen for ny ordning er god erfaring fra videregående samt timetallet i faget[1]. Høringsuttalelser fra lærere, skoler, fagorganer, kommuner og andre viser at vurderingen er delt knyttet til hvorvidt ny vurderingsordning er positivt eller negativt[2].

Prosjektet har gjennom skoleåret 2021/2022 fulgt fire lærere i kvalitative intervju, og tanken er å komplementere disse dataene med kvantitative data om samme tematikk. Masteroppgaven ble påbegynt høsten 2021, og fullføres i november 2022.

Om spørreskjemaet og personvern

Spørreskjemaet gjennomføres gjennom den anonyme løsningen Nettskjema, som ikke registrerer IP-adresse, navn, e-postadresse eller annen identifiserbar informasjon. Spørreskjemaet består kun av lukkede spørsmål, som vil si at det ikke finnes noen mulighet for å etterlate informasjon som forteller noe om lærerens identitet, arbeidsforhold eller tilknytning til bestemte skoler eller kommuner. Opplysningene gitt i spørreundersøkelsen skal ikke brukes til andre formål.

Spørreskjemaet inneholder spørsmål om læreres erfaringer med og tanker om ordningen, og erfaringer fra intervju er at lærere opplever det som positivt å delta fordi det bidrar til økt refleksjon om et tema som kanskje ikke prioriteres i en ellers veldig travel jobbhverdag. Spørreskjemaet kan sånn sett benyttes som en inngang til felles refleksjon knyttet til endringen eller mer generelt om vurdering og vurderingsfellesskap.

Spørreskjemaet kan gjennomføres på
- norsk: https://nettskjema.no/a/270949
- engelsk: https://nettskjema.no/a/217543

Hvis du har spørsmål til forskningsprosjektet, kan du ta kontakt med:
- Masterstudent Karen Langfjæran,
- NTNU ved førstelektor Anita Normann,

Med vennlig hilsen
Karen Langfjæran
Masterstudent i fagdidaktikk
Institutt for lærerutdanning
NTNU

---

[1]UDIR, 2019a: Høring – læreplan i engelsk. Hentet fra https://hoering.udir.no/Hoering/v2/338
[2]UDIR, 2019b: Høringsuttalelser – læreplan i engelsk. Hentet fra https://hoering-publisering.udir.no/338/uttalelser

Appendix E. Interview Guides

<u>**Intervjuguide 1**</u>

1. Innledning / Beskrivelse av arbeidsforhold

Kan du begynne med å fortelle litt om deg selv …

 Hvor gammel er du?

 Hvilke klasser underviser du i? Hvilken bakgrunn har du?

 Erfaring fra skolen, arbeidslengde

 Strukturelle forhold (type skole, skolekultur, kollegiet), vurderingskultur og arbeidskultur, samarbeid/fagteam, elevgruppe, klasseforhold, sammensetning av elever

Deltakelse i høringen

2. Erfaring før innføring av ny vurderingsordning

Hvordan er din erfaring med todelt karaktersetting?

Har du erfaringer med å gi én karakter fra før?

3. Forventninger

Fra to karakterer til én karakter i engelsk

 Kan du huske hva du tenkte om overgangen til én karakter da det ble bestemt?

 Hvordan stiller du deg til å gi én karakter i stedet for to?

Samlet kompetanse

 Hvilke umiddelbare tanker får du når jeg sier «samlet kompetanse i engelsk»?

 Hvordan ser du på det å vurdere en «samlet kompetanse» i én karakter kontra en muntlig karakter og en skriftlig?

  Hvordan vektlegge kompetanser og ferdigheter

Konsekvenser av ny vurderingsordning. Fordeler og ulemper.

Hvilke forventninger har du til ulike elevers reaksjon?

 «Normaleleven»; elever med dysleksi og andre med disproporsjonalt kompetansebilde (sterkere muntlig/skriftlig); elever som ikke vil snakke i timen, spesielt interesserte osv.

4. Gjennomføring og forberedelse

Opplever du at ny vurderingsordning har hatt konsekvenser for hvordan du legger opp undervisningen så langt?

 For eksempel hva du vektlegger i undervisningen, hvilke ferdigheter i fokus

 Er noe annerledes fra da elevene skulle få to karakterer? Endret fokus?

  Hva med underveisvurderingen? På hvilke måter?

Arbeidet dere noe i lærerfellesskapet, i grupper eller på teamnivå for å forberede dere på den nye vurderingsordningen?

Hvordan drøftes vurderingspraksis som følge av ny vurderingsordning? Hvem tar initiativ til endring og utvikling i engelskfaget?

Har dere noe lærersamarbeid

- knyttet til hva dere skal vektlegge i vurderingen?
- for å sikre felles vurderingspraksis?
- med utforming av vurderingskriterier med én karakter kontra to?

På hvilken måte har skoleeier eller skoleleder vært involvert i arbeidet med ny vurderingsform? Har de signalisert at det skal jobbes med dette?

Helt til slutt har jeg noen spørsmål knyttet til forventninger om arbeidet med sluttkarakter
Hva tenker du (på nåværende tidspunkt) kan bli annerledes når du skal vurdere muntlige og skriftlige ferdigheter opp mot sluttkarakter?
Vektlegging av andre ferdigheter med ny vurderingsordning

Intervjuguide 2

1. Refleksjoner rundt setting av terminkarakter samt tentamen
Hvordan synes du det har gått med vurdering av terminkarakter i engelsk?

Hvilke vurderingssituasjoner ligger til grunn for terminkarakteren?
Hvor stor andel av vurderingsgrunnlaget vil du anslå er uformelt, altså ikke basert på innleveringer eller presentasjoner?

Hvordan har du gått frem for å gi vurdering på *samlet kompetanse* sammenlignet med tidligere muntlig og skriftlig karakter?

På hvilke måter er det annerledes enn da du ga to karakterer?

Hvordan *opplevde* du å gi én samlet karakter sammenlignet med det å gi en muntlig og en skriftlig vurdering?

Kan du gi et eksempel på noe som har vært utfordrende i forbindelse med det å gi vurdering til forrige termin?

Hvordan har elevene blitt involvert i prosessen med ny vurdering eller Fagfornyelsen generelt?

Hvis elever nå rent hypotetisk skulle klagd på karakteren og ønsket tilbakemelding på vurderingen, hvordan tenker du at du hadde fremstilt vektleggingen av ferdigheter og den samlede kompetansen?

Hvordan la dere opp tentamen?
Gjennomførte dere noen eksplisitte endringer på årets tentamen (i lys av ny vurderings- og eksamensordning) sammenlignet med tidligere tentamener?
Hvilke konsekvenser hadde det for måten du ga vurdering på?
Jobbet dere med eksempelsettet av eksamensoppgaver i arbeidet med tentamen?

Hvordan har du eller dere som lærere jobbet med de nye kompetansemålene i LK20 i sammenheng med vurderingsgrunnlag og terminkarakter?
Hvordan synes du de skiller de seg fra kompetansemålene i LK06?

2. Samarbeid om vurdering på skolen
Har dere hatt seksjons- eller teammøter om vurdering i sammenheng med tentamen og terminvurdering?

Har dere snakket spesifikt om hvordan dere skal vektlegge ferdigheter i vurderingen?

Har dere, siden forrige gang vi snakket, jobbet noe ellers felles med vurdering?
    Har dere tatt opp dette med én kontra to karakterer i denne omgang?

Hva slags støtte har dere fått fra UDIR eller skoleledelsen om hvordan du skal gi vurdering med én karakter?
        Har dere fått tilgang til noen veiledning, dokumenter eller lignende støttemateriell?
        Har du blitt invitert til eller vært på møter om vurdering i kommunal regi eller gjennom UDIR, Utdanningsforbundet eller lignede aktører?
            Kjenner du til om dette har blitt arrangert?

Føler du at dere som lærere har fått tilstrekkelig veiledning eller retningslinjer fra UDIR eller ledelsen om hvordan dere skal gi vurdering med én karakter?
                Hva hadde du forventet eller ønsket?

Hva tenker du at kompetanseutvikling på området, enten fra ledelsen eller UDIR, kunne bidratt til?

Hvordan tenker du det kan være hensiktsmessig at dere i fagfellesskapet jobber fremover?

3. Veien videre

Vet du om dere kommer til å jobbe noe med vurdering, og spesielt da dette med én karakter kontra to, videre denne våren?

Vet du om dere har noen planer om å jobbe med ny eksamensform denne våren?

Forventer du at noe endrer seg i måten du legger opp undervisningen eller underveisvurderingen i forbindelse med dette med én karakter denne våren?

<u>Intervjuguide 3</u>

1. Erfaringer
Hvordan vil du oppsummere arbeidet med å sette karakterer i engelsk i vår?

Hvilke vurderingssituasjoner/elevarbeider ligger til grunn for standpunktkarakterene du har satt i engelsk?

Kan du gi noen eksempler på elever som har vippet mellom to karakterer, og hva du vektla i den endelige vurderingen?

Kan du si noe om hvordan du tenker generelt om vurdering av en tenkt…
        … elev som strever med den skriftlige delen av faget, men er svært sterk muntlig
        … elev som er svært sterk skriftlig, men med lite muntlig vurderingsgrunnlag

Kan du gi et eksempel på noe som har vært utfordrende?

Kan du gi et eksempel på noe som har gått spesielt bra?

Hvordan har elevene reagert på felles karakter i engelsk? Har du erfart noen endringer i måten de har arbeidet på i faget?

Hvordan har overgangen fra to karakterer til én påvirket måten du har arbeidet med vurdering på?

Hvordan har dere samarbeidet på arbeidsplassen i forbindelse med karaktersettingen?

Har dere jobbet noe med å se på eller diskutere hverandres vurderingsgrunnlag eller enkeltelever?

Har dere, siden sist vi pratet, hatt noe fokus på ny vurderingsordning eller på å utvikle felles vurderingspraksis (enten som signaler fra ledelsen, i team eller seksjonen)?

Har dere brukt noen hjelpemidler i arbeidet?
- De nye kompetansemålene
- De nye kjennetegnene til måloppnåelse
- Læreplanens nye definisjon av kompetansebegrepet
- Nasjonale prøver
- Eksempler på eksamensoppgaver
- Nettressurser fra UDIR, Statsforvalteren eller andre
- Lærebokressurser
- Det felles europeiske rammeverket
    - Deskriptorskalaene/referansenivåene
    - Sjekklistene i den europeiske språkpermen
- Annet

Hvis elever nå (rent hypotetisk) skulle klagd på karakteren og ønsket tilbakemelding på vurderingen, hvordan tenker du at du hadde fremstilt vektleggingen av ferdigheter og den samlede kompetansen?

2. Læreplanforståelse og læringssyn
Hvis du ser for deg en vurderingsskala fra «å tenke helhetlig» til «å vurdere atskilte ferdigheter»: Hvor er du på den skalaen, og hvordan begrunner du det?

UDIR påpeker at det kan hende at lærerne har hatt en litt matematisk forståelse av vurdering når lærere har gitt to karakterer. Hva tenker du om det?

Jeg har på forhånd sendt deg «Omtale av vurdering» fra ny læreplan. Oppfatter du at den
    Er fagspesifikk?[1]
    Synliggjør hvordan underveisvurderingen skal være læringsfremmende?[1]
    Tydeliggjør grunnlaget for både underveis- og sluttvurdering?[1]
    Uttrykker hva elevene samlet sett skal vurderes etter?[2]
    Er en støtte til lærere i underveis- og sluttvurderingen?[2]
    Er i tråd med omtalen av vurdering i Overordnet del**?[2]
    Er utformet med utgangspunkt i kjerneelementene, kompetansedefinisjonen og kompetansemålene?[2]
    Gjenspeiler at vurdering skal støtte opp under elevers dybdelæring, motivasjon og mestring?[2]
    Tydeliggjør progresjon i læreplanen?[2]

Tydeliggjør sammenhenger i læreplanverket, mellom overordnet del, kjerneelementer og kompetansemål? [2]
Hentet fra Meld. St. 28 (2015–2016) [1] og Retningslinjer for utforming av nasjonale og samiske læreplaner for fag i LK20 og LK20s [2].

Jeg har også sendt deg definisjonene på kompetanse slik de er definert i LK06 og LK20. Har dette hatt betydning for din arbeidsplass og hvordan dere har vurdert elevenes kompetanse? *Begrunn svaret.*

En tredje ting jeg kort vil komme inn på fra læreplanen, er begrepet «utforske», som er blant de mest sentrale verbene i kompetansemålene. Jeg har på forhånd sendt deg definisjonen av dette. Har det hatt betydning for hvordan du eller lærerfellesskapet har jobbet med vurdering? *Begrunn.*

3. Tiden fremover
Hva tenker du det kan være hensiktsmessig for lærere å tenke over eller jobbe med videre i overgangen fra to til en karakter i engelsk?

Hvilke forventninger eller ønsker har du fra politikere, Utdanningsdirektoratet, Statsforvalteren, ledelsen eller andre?

Appendix F. Questionnaires

Norwegian Version

Fra to karakterer til én i engelsk på ungdomsskolen

Obligatoriske felter er merket med stjerne *

Fra to karakterer til én i engelsk på ungdomsskolen

Takk for at du deltar i masterprosjektet mitt ved å svare på denne spørreundersøkelsen.

Med LK20 gir engelsklærere i ungdomsskolen nå én karakter i stedet for to (muntlig og skriftlig), og dette masterprosjektet følger lærere gjennom skoleåret 21/22 i kvalitative intervju for å kunne si noe om læreres erfaringer med denne endringen. Formålet med denne spørreundersøkelsen er å hente kvantitative data om tematikken.

Jeg er veldig takknemlig for at du deltar gjennom å svare på denne spørreundersøkelsen.

Hvis du har spørsmål knyttet til spørreundersøkelsen eller masteroppgaven, uavhengig om du svarer eller ikke, kan du ta kontakt på karenel@stud.ntnu.no.

Anonymitet

Alle svar er fullstendig anonyme.

Spørreundersøkelsen består kun av spørsmål med lukkede felter, f.eks. avkrysning, for å utelukke muligheten til å oppgi identifiserende informasjon.

Har du lest og forstått informasjonen? *

○ Jeg har lest og forstått informasjonen

Obligatoriske felter er merket med stjerne *

Din bakgrunn

Jobber du som engelsklærer på ungdomsskolen?

○ Ja, jeg er engelsklærer på flere trinn

○ Ja, jeg er engelsklærer i 10. klasse

○ Ja, jeg er engelsklærer i 9. klasse

○ Ja, jeg er engelsklærer i 8. klasse

○ Nei, men jeg jobbet som engelsklærer i fjor

○ Nei, jeg jobber ikke som engelsklærer på ungdomsskolen

Hvor lenge har du jobbet som engelsklærer på ungdomsskolen?

○ 1–3 år

○ 3–5 år

○ 5–10 år

○ 10–20 år

○ 20–30 år

○ 30+ år

Hvilken aldersgruppe tilhører du?

○ 18-25

○ 25-35

○ 35-45

○ 45-55

○ 55+

Obligatoriske felter er merket med stjerne *

Initielle tanker om endringen

Diskuterte du og kollegaene dine endringen fra to karakterer til én da læreplanen i engelsk var på høring i 2019?

Spørsmålet i høringen: "Vi foreslår endring i standpuntvurderingen etter 10. trinn: Elevene skal ha én karakter som uttrykker elevens samlede kompetanse etter 10. trinn. Er dette en god ordning ut fra læreplanen og timetall i faget?"

Velg … ⌄

Hvis du husker hva skolen/kollegiet svarte: Var dere (som gruppe) positiv eller negativ til endringen?

Velg … ⌄

Hva er din personlige mening om endringen?

Velg … ⌄

Sideskift ···

Obligatoriske felter er merket med stjerne *

Samarbeid i kollegiet

Har du og kollegaene dine diskutert overgangen til en karakter i engelskfaget?

○ Ja, uformelt (uformelle samtaler med kolleger)

○ Ja, formelt (fagteam, seksjonsmøter, osv.)

○ Ja, både uformelt og formelt

○ Usikker

○ Nei

Hva har du og kollegaene dine diskutert om endringen fra to karakterer til én?

Velg alle som reflekterer noe dere har diskutert.

- ☐ Vurderingskriterier
- ☐ Vurderingskultur
- ☐ Forholdet mellom summativ og formativ vurdering
- ☐ Hvordan skape en felles vurderingspraksis
- ☐ De potensielt positive effektene for elever
- ☐ De potensielt negative effektene for elever
- ☐ De potensielt positive effektene for lærere
- ☐ De potensielt negative effektene for lærere
- ☐ Hvordan vurdere samlet kompetanse kontra muntlig og skriftlig
- ☐ De potensielt utfordrende aspektene ved vurdering av samlet kompetanse
- ☐ De potensielt positive aspektene ved vurdering av samlet kompetanse
- ☐ Hvordan definere samlet kompetanse
- ☐ Vi har ikke diskutert noe av dette
- ☐ Annet

Har dere brukt noen hjelpemidler i arbeidet/diskusjonene?

Velg også alternativer du har brukt individuelt.

- ☐ De nye kompetansemålene
- ☐ De nye kjennetegnene til måloppnåelse
- ☐ Læreplanens nye definisjon av kompetansebegrepet
- ☐ Nasjonale prøver
- ☐ Eksempler på eksamensoppgaver
- ☐ Modulen om vurdering i UDIRs kompetansepakke
- ☐ Andre nettressurser fra UDIR, Statsforvalteren eller andre
- ☐ Lærebokressurser
- ☐ Det felles europeiske rammeverket
- ☐ Annet

Har du mottatt støttemateriell, kurs eller lignende fra UDIR, Statsforvalteren, arbeidsplassen din eller andre relevante aktører?

- ○ Ja, fra UDIR
- ○ Ja, fra Statsforvalteren
- ○ Ja, fra skoleledelsen
- ○ Ja, fra andre aktører
- ○ Usikker
- ○ Nei

Har du blitt informert om at du på et tidspunkt kommer til å motta støttemateriell, kurs eller lignende?

- ○ Ja
- ○ Nei
- ○ Vet ikke

Obligatoriske felter er merket med stjerne *

Dine erfaringer

Har du endret måten du vurderer på som resultat av endringen til én karakter (samlet kompetanse) sammenlignet med da du ga to karakterer (muntlige og skriftlige ferdigheter)?

På en skala fra 1 (ingen endring i måten du vurderer på) til 5 (fullstendig restrukturering av måten du vurderer på)

1	2	3	4	5

Verdi [ ]

Har du endret måten du underviser på som resultat av endringen til én karakter (samlet kompetanse) sammenlignet med da du ga to karakterer (muntlige og skriftlige ferdigheter)?

På en skala fra 1 (ingen endring i måten du underviser på) til 5 (fullstendig restrukturering av måten du underviser på)

1	2	3	4	5

Verdi [ ]

Hvis du svarte 2 eller mer: Hva har du endret?

Har du endret måten du underviser/vurderer på, som resultat av overgangen til én karakter (som reflekterer samlet kompetanse) sammenlignet med to karakterer (muntlige og skriftlige ferdigheter)? Velg alle som er aktuelle.

- ☐ Mer fokus på skriftlige ferdigheter
- ☐ Mindre fokus på skriftlige ferdigheter
- ☐ Mer fokus på muntlige ferdigheter
- ☐ Mindre fokus på muntlige ferdigheter
- ☐ Mer fokus på lytting
- ☐ Mindre fokus på lytting
- ☐ Mer fokus på forståelse
- ☐ Mindre fokus på forståelse
- ☐ Mer fokus på spontane muntlige aktiviteter
- ☐ Mindre fokus på spontane muntlige aktiviteter
- ☐ Færre tester eller skriveoppgaver
- ☐ Flere tester eller skriveoppgaver
- ☐ Færre muntlige oppgaver/øvelser
- ☐ Flere muntlige oppgaver/øvelser
- ☐ Generelt mindre fokus på vurdering
- ☐ Generelt mer fokus på vurdering
- ☐ Annet

Obligatoriske felter er merket med stjerne *

Dine erfaringer 2

Hvordan har du erfart denne endringen så langt?

Velg alle som gjelder.

- [ ] Det er mer utfordrende å vurdere samlet kompetanse sammenlignet med å gi to karakterer

- [ ] Det er enklere å vurdere samlet kompetanse sammenlignet med å gi to karakterer

- [ ] Redusert press på å ha varierte vurderingssituasjoner

- [ ] Økt press på å ha varierte vurderingssituasjoner

- [ ] Enklere å drive tilpasset opplæring

- [ ] Vanskeligere å drive tilpasset opplæring

- [ ] Mer fleksibilitet

- [ ] Mindre fleksibilitet

- [ ] Mer tidkrevende

- [ ] Mindre tidkrevende

- [ ] Mer eller mindre som før

- [ ] Annet

Hvordan har du involverte elevene i endringen?

Velg alle som gjelder.

- ☐ Svært lite, da jeg ikke anser endringen som særlig stor
- ☐ Svært lite, da jeg underviser i 8. og de ikke har erfaring med to karakterer
- ☐ Lite, foruten å informere om det
- ☐ Vi har diskutert endringen
- ☐ Vi har diskutert begrepet "samlet kompetanse"
- ☐ Vi har ikke diskutert vurderingsformen, men vi har diskutert kompetansemålene
- ☐ Vi har ikke diskutert vurderingsformen, men vi har diskutert andre aspekter ved LK20
- ☐ Usikker

Hvilke generelle erfaringer har du gjort deg når du har vurdert elever med et disproporsjonalt kompetansebilde?

For eksempel hvis du hadde vurdert en elev til 6 i muntlig og 3 i skriftlig med forrige vurderingsform.

- ☐ Min erfaring er at det er utelukkende positivt for disse elevene
- ☐ Min erfaring er at det er hovedsaklig positivt for disse elevene
- ☐ Min erfaring er at det er utelukkende negativt for disse elevene
- ☐ Min erfaring er at det er hovedsaklig negativt for disse elevene
- ☐ Som før, men jeg må jobbe mer med tilpasset opplæring
- ☐ Som før, men jeg må jobbe mindre med tilpasset opplæring
- ☐ Ingen endring
- ☐ Usikker

Obligatoriske felter er merket med stjerne *

Påstander

På en skala fra 1–5: Hvor enig er du i denne påstanden?

"Det er enklere å vurdere med én karakter"

1 = Ikke enig. 5 = Helt enig.

| 1 | 2 | 3 | 4 | 5 |

Verdi [ ]

På en skala fra 1–5: Hvor enig er du i denne påstanden?

"Å gi én karakter i stedet for to har redusert tiden jeg har brukt på vurdering dette skoleåret"

1 = Ikke enig. 5 = Helt enig.

| 0 | 1 | 2 | 3 | 4 | 5 |

Verdi [ ]

På en skala fra 1–5: Hvor enig er du i denne påstanden?

"Jeg forventer at jeg kommer til å bruke mindre tid på vurdering på grunn av endringen"

1 = Ikke enig. 5 = Helt enig.

| 1 | 2 | 3 | 4 | 5 |

Verdi [ ]

Obligatoriske felter er merket med stjerne *

Påstander 2

På en skala fra 1–5: Hvor enig er du i denne påstanden?

"Jeg har en klar forståelse av og oversikt over hvordan jeg skal vektlegge ulike ferdigheter i avgjørelsen om sluttkarakter"

1 = Ikke enig. 5 = Helt enig.

1	2	3	4	5

Verdi [ ]

På en skala fra 1–5: Hvor enig er du i denne påstanden?

"Jeg tror reduksjonen i antall karakterer kommer til å ha negative konsekvenser for noen elever"

1 = Ikke enig. 5 = Helt enig.

1	2	3	4	5

Verdi [ ]

På en skala fra 1–5: Hvor enig er du i denne påstanden?

"Jeg tror reduksjonen i antall karakterer kommer til å være fordelaktig for majoriteten av elevene"

1 = Ikke enig. 5 = Helt enig.

1	2	3	4	5

Verdi [ ]

Obligatoriske felter er merket med stjerne *

Et siste spørsmål …

Hvilken av disse påstandene om å bruke mer tid på denne tematikken passer best overens med dine tanker?

- ○ Vi har allerede diskutert det i detalj
- ○ Vi har så mange andre tema å dekke i møtene våre
- ○ Det er verdt å prioritere det for å sikre felles vurderingspraksis
- ○ Det er viktig for meg
- ○ Det er ikke viktig for meg

English version

Transitioning from Two Grades to One Grade in Lower Secondary English

Obligatoriske felter er merket med stjerne *

Transitioning from Two Grades to One Grade in Lower Secondary English

Thank you for taking part in my MA research project by completing this survey regarding the change from two grades to one grade in Lower Secondary English.

With LK20, Lower Secondary teachers will now provide one grade, reflecting the overall competence of a pupil, rather than two grades, reflecting oral and written skills. This research project seeks to shed a light on teachers' experiences with the change. I conduct interviews to gather qualitative data throughout the school year 21/22, and the purpose of this survey is to broaden the perspective with quantitative data.

Your participation is very valuable to the project.

If you have any questions about the survey – whether you answer or not, please get in touch on karenel@stud.ntnu.no.

Anonymity

All answers are completely anonymous.

The survey consists of closed questions only, to avoid any chance of obtaining any identifying information.

Have you read and understood the information? *

- ○ I have read and understood the information

Your background

Are you currently an English teacher in a Norwegian Lower Secondary School?

○ Yes, I teach English for a variety of classes

○ Yes, I teach English for 10th grade

○ Yes, I teach English for 9th grade

○ Yes, I teach English for 8th grade

○ No, but I taught English last year

○ No, I am not an English teacher in Lower Secondary School

How long have you worked in Lower Secondary School?

○ 1–3 years

○ 3–5 years

○ 5–10 years

○ 10-20 years

○ 20-30 years

○ 30+ years

What is your age group?

○ 18-25

○ 25-35

○ 35-45

○ 45-55

○ 55+

Obligatoriske felter er merket med stjerne *

Initial thoughts about the change

Did you and your colleagues discuss the move from two grades to one grade as part of the curriculum hearing in 2019?

The relevant question from the hearing: "Vi foreslår endring i standpunktvurderingen etter 10. trinn: Elevene skal ha én karakter som uttrykker elevens samlede kompetanse etter 10. trinn. Er dette en god ordning ut fra læreplanen og timetall i faget?"

Velg … ⌄

If you remember what you responded collectively, were you (as a group) positive or negative to the change?

Velg … ⌄

What is your personal opinion of the change?

Velg … ⌄

Collegial collaboration

Have you and your colleagues discussed the transition to one grade in Lower Secondary English?

○ Yes, informally (informal conversations with colleagues)

○ Yes, formally ("fagteam", "seksjonsmøte" etc.)

○ Yes, both informally and formally

○ Not sure

○ No

What have you and your colleagues covered in your discussions about the change from two grades to one grade?

Choose all that apply.

☐ Assessment criterias

☐ Assessment culture

☐ The relationship between summative and formative assessment

☐ Creating a common assessment practice

☐ The potentially positive effects on pupils

☐ The potentially negative effects on pupils

☐ The potentially positive effects on teachers

☐ The potentially negative effects on teachers

☐ How to assess the overall competence versus oral and written skills

☐ The potential challenging aspects of assessing the overall competence

☐ The potential positive aspects of assessing the overall competence

☐ Clarifying the term "overall competence"

☐ We have not discussed any of this

☐ Other

Have you used any supportive material to aid your discussions?

Please also include any material you may have used individually.

- ☐ The new competence aims
- ☐ The new characteristics of goal achievement
- ☐ LK20s new definition of competence
- ☐ National tests
- ☐ Examples of exam tasks
- ☐ The assessment module in UDIRs competence package
- ☐ Other online resources from UDIR, Statsforvalteren or others
- ☐ Textbook resources
- ☐ Common European Framework of Reference for Languages
- ☐ Other

Have you received any supportive material, training etc. provided by UDIR, Statsforvalteren, your school or other relevant parties?

- ○ Yes, by UDIR
- ○ Yes, by Statsforvalteren
- ○ Yes, by the school I work for
- ○ Yes – other
- ○ Not sure
- ○ No

Have you been informed that you will receive any supportive material, training or similar at some point?

- ○ Yes
- ○ No
- ○ Not sure

Obligatoriske felter er merket med stjerne *

Your experiences

Have you changed the way you assess as a result of transitioning to one grade (reflecting overall competence) rather than two (oral skills + written skills)?

On a scale from 1 (no change in assessment practice) to 5 (complete restructure of assessment practice)

1	2	3	4	5

Verdi [ ]

Have you changed the way you teach as a result of transitioning to one grade (reflecting overall competence) rather than two (oral skills + written skills)?

On a scale from 1 (no change) to 5 (complete restructure)

1	2	3	4	5

Verdi [ ]

If you responded 2 or more in the two previous questions, what have you changed?

Have you changed the way you teach/assess as a result of transitioning to one grade (reflecting overall competence) rather than two (oral skills + written skills)?

Choose all that apply.

- ☐ More focus on written skills
- ☐ Less focus on written skills
- ☐ More focus on oral skills
- ☐ Less focus on oral skills
- ☐ More focus on listening
- ☐ Less focus on listening
- ☐ More focus on understanding
- ☐ Less focus on understanding
- ☐ More focus on spontaneous oral activity
- ☐ Less focus on spontaneous oral activity
- ☐ Fewer tests or writing assignments
- ☐ More tests or writing assignments
- ☐ Fewer oral assignments
- ☐ More oral assignments
- ☐ Less focus on assessment generally
- ☐ More focus on assessment generally
- ☐ Other

Your experiences 2

How have you experienced the change in assessment?

Choose all that apply.

- ☐ More challenging to assess "overall competence" vs. oral/written skills
- ☐ Easier to assess "overall competence" vs. oral/written skills
- ☐ Decreased pressure on varying assessment situations
- ☐ Increased pressure on varying assessment situations
- ☐ More focus on assessing different skills
- ☐ Easier to adapt to pupils' needs
- ☐ More difficult to adapt to pupils' needs
- ☐ More flexibility
- ☐ Less flexibility
- ☐ More time consuming
- ☐ Less time consuming
- ☐ More or less as before
- ☐ Other

How have you involved the pupils with regards to the change in assessment?

Choose all that apply.

- ☐ Very little involvement as I have not considered it a big change
- ☐ Very little involvement as they (8th graders) have no experience with two grades
- ☐ Little involvement besides commenting on it
- ☐ We have discussed the change
- ☐ We have discussed the term "overall competence"
- ☐ We have not discussed the change in grading, but we have discussed the new competence aims
- ☐ We have not discussed the change in grading, but we have discussed other changes in LK20
- ☐ Not sure

What is your general experience with assessing pupils with a disproportionate skill set?

For example, a pupil you would give a 6 in oral English and 3 in written English with the old system.

- ☐ My experience is that the change is entirely positive for these pupils
- ☐ My experience is that the change is mainly positive for these pupils
- ☐ My experience is that the change is entirely negative for these pupils
- ☐ My experience is that the change is mainly for these pupils
- ☐ Similar, but I have to work more on adapting to their individual needs
- ☐ Similar, but I have to work less on adapting to their individual needs
- ☐ No change
- ☐ Not sure

Statements

On a scale from 1 to 5, how much do you agree with the following statement?

"It is easier to assess with one grade versus two grades"

1 = I do not agree. 5 = I fully agree.

1	2	3	4	5

Verdi [ ]

On a scale from 1 to 5, how much do you agree with the following statement?

"Giving one grade rather than two has reduced the time I spent on assessment this school year"

1 = I do not agree. 5 = I fully agree.

0	1	2	3	4	5

Verdi [ ]

On a scale from 1 to 5, how much do you agree with the following statement?

"I expect a reduction in the time I spend on assessment because of the change in assessment"

1 = I do not agree. 5 = I fully agree.

1	2	3	4	5

Verdi [ ]

Statements 2

On a scale from 1 to 5, how much do you agree with the following statement?

"I have a clear understanding and overview of how I should weigh the different skills and competences in the pupils' final grade"

1 = I do not agree. 5 = I fully agree.

1	2	3	4	5

Verdi

On a scale from 1 to 5, how much do you agree with the following statement?

"I believe reducing the number of grades will have negative effects on certain pupils"

1 = I do not agree. 5 = I fully agree.

1	2	3	4	5

Verdi

On a scale from 1 to 5, how much do you agree with the following statement?

""I believe reducing the number of grades will be beneficial for the majority of pupils"
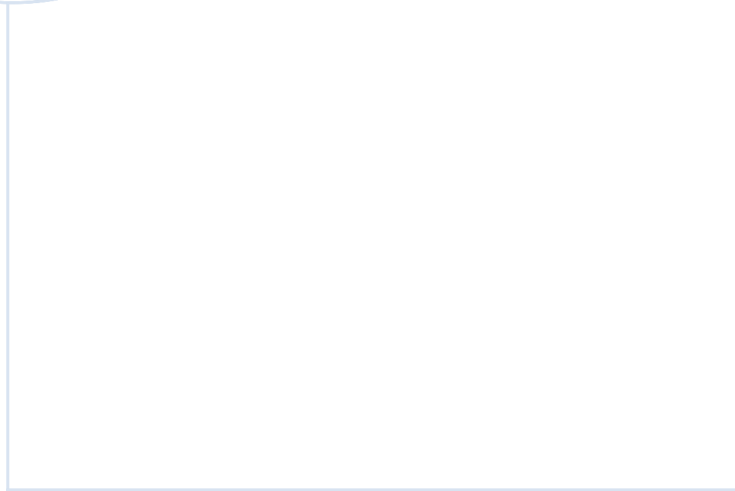
1 = I do not agree. 5 = I fully agree.

1	2	3	4	5

Verdi

One last question …

Which of these thoughts about spending more time on the topic of one grade align the most with yours?

- ○ We have already discussed it in detail
- ○ We have so many other topics to be covered in our meetings
- ○ It is worth prioritising to create a common assessment practice
- ○ It is important to me
- ○ It is not important to me