



*Faculty of humanities, social sciences, and education*

***The national test in English: A game of organizational Chinese  
Whispers?***

*A qualitative and quantitative study regarding how the results from the national test in English is used as a mapping- and development tool in a selection of schools in northern Norway.*

*Ishak Moe Haoual & Sverre Ness-Voll*

*Master's thesis in English didactics LER-3902, May 2024*



## Foreword

Our five-year journey towards becoming teachers has finally come to an end. Although time-consuming, we feel a certain amount of pride within ourselves due to our work process, which has been characterized by our ability to enjoy the process instead of being overwhelmed by it.

Firstly, we would like to thank our friends and families for their support and patience, as we endlessly went on about things that were of no concern to them. Secondly, a heartfelt thanks goes out to our fellow students. The ever long coffee breaks were much needed, and the possibility for us to ask our stupid questions whenever we wanted. We also want to thank our supervisor for the assistance that we received.

Tromsø, May 2024

Ishak Moe Haoual & Sverre Ness-Voll



## Sammendrag

Som fremtidige lærere anser vi kontinuitet i skolens kvalitetsutvikling som et svært viktig tema. En stor del av kvalitetsutviklingen i skolen er de årlige nasjonale prøvene, som ble innført etter dårlige resultater på PISA-undersøkelsen i 2000, også kjent som «PISA-sjokket». Etter innføringen har det vært mye kritikk rettet mot de nasjonale prøvene, og senest i 2023 gikk Utdanningsforbundet inn for å avvikle prøvene til fordel for nye læringsstøttende prøver.

Vi har begge jobbet i skolesystemet i flere år, og har gjennom dette opplevd nasjonale prøver som et omstridt tema, preget av usikkerhet og metodefrihet. Et sentralt spørsmål vi stilte oss i starten av vårt masterprosjekt var hvorvidt resultatene fra de nasjonale prøvene ble brukt som det kartleggings- og utviklingsverktøyet det *kan* være, noe som ledet oss til problemstillingen «*i hvor stor grad bruker et utvalg av skoler i Nord-Norge resultatene fra nasjonale prøver i engelsk som en del av skolens faglige utvikling?*» problemstillingen støttes opp med tre forskningsspørsmål:

1. I hvor stor grad erfarer engelsklærere og skoleledere at nasjonale prøver i engelsk er et prioritert kartleggings- og utviklingsverktøy i den faglige utviklingen av elevene?
2. I hvor stor grad påvirker etablerte rutiner anvendelsen av nasjonale prøver i engelsk som et kartleggings- og utviklingsverktøy av lærere?
3. I hvor stor grad anser lærere og skoleledere nasjonale prøver i engelsk som et viktig verktøy i den faglige utviklingen av elevene?

Oppgaven bygger på organisasjonsteorier som teoretisk grunnlag. Dalins fem dimensjoner i skolen som organisasjon, og Senges fem disipliner om skolen som en lærende organisasjon vil være sentral i forskningen. Vi har valgt å se på tidligere forskning fra både Norge og andre land for å gi oss et godt sammenligningsgrunnlag med våre egne data.

For å samle inn data har vi valgt å benytte oss av metodetriangulering, som ga oss et godt sammenligningsgrunnlag i det at vi kunne se kvantitativ data opp mot kvalitativ data. En spørreundersøkelse ble sendt ut til skoleledere og lærere, der 16 skoleledere og

48 lærere deltok. Etter spørreundersøkelsen gjennomførte vi sju intervju med lærere og skoleledere for å utforske trendene som kom fram av spørreundersøkelsen.

Våre funn viser store sprik i hvordan de individuelle lærerne og skolelederne ser på nasjonale prøver i engelsk som et kartleggings- og utviklingsverktøy. Dette påvirker hvordan resultatene brukes som en del av skolens faglige utvikling. Faktorer som manglende informasjonsflyt, kollegialt samarbeid og prioritering har i stor grad innvirkning på hvordan resultatene brukes, og funnene viser at en organisatorisk endring vil kunne være hensiktsmessig for å sikre at nasjonale prøver i engelsk utnyttes til dets fulle potensial som kartleggings- og utviklingsverktøy.

## Abstract

As future teachers, we regard the continuity in school quality improvement as a crucial concern. A substantial aspect of school quality enhancement revolves around the annual national assessments, which were implemented following the disappointing outcomes in the 2000 PISA evaluation, famously dubbed the “PISA shock.” After the implementation, national tests have faced considerably amount of criticism, with the latest instance occurring in 2023 when the Teachers Union advocated for their discontinuation in favor of new learning-supportive assessments.

Having both worked in the educational system for several years, we’ve witnessed national tests emerge as a controversial topic categorized by uncertainty and methodological freedom. At the start of our research process, we asked ourselves a fundamental question: Are the results from national assessments being effectively utilized as the evaluative and development tools they’re meant to be? This inquiry led us to our thesis question: “To what extent are a selection of schools in Northern Norway incorporating the results from national assessments in English into the school's academic development?” This inquiry is underpinned by three research questions:

1. To what degree do English teachers and school leaders view national assessments in English as a prioritized evaluative and development tool in pupils’ academic growth?
2. To what extent do established procedures influence teachers’ utilization of national assessments in English as evaluative and development tools?
3. How do teachers and school leaders perceive national assessments in English as vital tools in pupils’ academic advancement?

The thesis is grounded in organizational theories, with Dalin’s five dimensions of the school as an organization and Senge’s five disciplines regarding the school as a learning organization forming the theoretical backbone of the research. We’ve decided to review prior research from both domestic and international **contexts** to provide a robust comparative foundation alongside our own data.

For data collection, we’ve employed method triangulation, facilitating comprehensive comparison by possessing both quantitative and qualitative data. A questionnaire was

distributed to school leaders and teachers, resulting in 16 school leaders and 48 teachers participating. Subsequent to the questionnaire, we conducted seven interviews with teachers and school leaders to delve into the trends highlighted by the survey.

Our findings reveal significant disparities in how individual teachers and school leaders perceive national assessments in English as evaluative and development tools. This divergence influences the integration of results into the school's academic development. Factors such as information flow, collaborative work, and prioritization significantly shape the utilization of results. The findings underscore the need for organizational change in order to ensure that the results from the national test in English are fully utilized, fulfilling the potential that the national test in English inherits as a evaluative and development tool.



# Table of Contents

1	Introduction.....	5
1.1	Reason for the choice of topic .....	5
1.2	Thesis- and research questions .....	6
1.3	Structure of the thesis.....	7
2	Previous research .....	9
2.1	National research .....	9
2.1.1	Seland, Vibe & Hovdhaugen (2013) .....	9
2.1.2	Werler & Færevaaag (2017).....	9
2.1.3	Parliamentary notice 21 (St. Meld. 21, 2016-2017).....	10
2.2	International research.....	11
2.2.1	Wayman & Jimerson (2014).....	11
2.2.2	Tonich (2021).....	12
2.2.3	Hardy, 2015 .....	13
2.2.4	Beck & Stetz, 1979.....	14
2.3	Chapter summary .....	15
3	National tests .....	16
3.1	Basic documents and guidance materials.....	18
3.2	Teaching to the test.....	20
3.3	Criticism of the national tests .....	21
3.4	Chapter summary .....	24
4	Theory.....	25
4.1	Organizational theory.....	25
4.1.1	School as an organization.....	25
4.1.2	Quality in school.....	26

4.2	Senges' five disciplines .....	27
4.3	Dalin's five dimensions .....	29
5	Methods .....	31
5.1	Interview .....	31
5.2	Interview guide.....	32
5.3	Mixed Methods Design .....	32
5.3.1	The Convergent Design .....	33
5.3.2	The Explanatory Sequential Design.....	33
5.3.3	The Exploratory Sequential Design .....	33
5.3.4	Qualitative interviews .....	34
5.3.5	Mailed- and online surveys or questionnaires.....	34
5.3.6	Transcribing.....	36
5.4	Methodological evaluation .....	37
5.4.1	Validity.....	37
5.4.2	Reliability .....	38
5.5	Gathering, analyzing, and interpreting data .....	40
5.5.1	Sampling .....	40
5.5.2	Quantitative data .....	40
5.5.3	Qualitative data .....	41
6	Research ethical considerations.....	43
6.1	Our role – reflexivity:.....	44
6.2	Strengths and weaknesses of the project .....	46
7	Results .....	48
7.1	Quantitative results.....	48
7.2	Table 1, quantitative results from teachers.....	49
7.3	Table 2, quantitative results from school leaders.....	50

7.4	Qualitative results .....	53
7.5	The national test in English as a mapping-and development tool.....	55
7.5.1	Improvement potential .....	55
7.5.2	Frequency.....	57
7.5.3	Information flow .....	58
7.5.4	Mapping tool .....	59
7.5.5	Traditions.....	59
7.5.6	Development tool.....	60
7.6	Utilizing the results.....	61
7.6.1	Mapping .....	61
7.6.2	Methodological freedom .....	61
7.6.3	Development.....	62
8	Discussion and analysis .....	64
8.1	Traditions.....	64
8.2	Methodological freedom .....	66
8.2.1	Preparation phase .....	66
8.2.2	Mapping- and development phase .....	69
8.3	Improvement potential .....	73
8.3.1	Preparation phase .....	73
8.3.2	Development phase .....	74
8.4	Mapping- and development tool .....	78
8.4.1	Frequency .....	80
8.4.2	Information flow .....	80
9	Answering the research questions.....	82
10	Answering the thesis question .....	87
11	Suggestions for further research.....	89

12	Concluding remarks.....	90
	Works cited.....	91
	Appendix 1, information sheet and consent form.....	93
	Appendix 2, interview guide .....	99
	Appendix 3, survey sent to teachers.....	102
	Appendix 4, survey sent to school leaders .....	107
	Appendix 5, individual graphs, teacher survey .....	112
	Appendix 6, individual graphs, school leader survey .....	115

# 1 Introduction

## 1.1 Reason for the choice of topic

National tests are standardized assessments designed to evaluate pupils' knowledge and skills in the English subject. The tests serve multiple purposes such as offering insight into pupils learning, providing data for educational policy, and tracking trends. The examination and insight that the national test in English, provides teachers and school leaders the opportunity to identify areas for improvement and development, assess the effectiveness of the curricula, assign resources in areas needed, and identifying achievement gaps. However, national testing is often debated regarding teaching to the test and the pressure pupils and educators may face.

The basis for selecting national test in English as our topic consists of several factors. Firstly, we examined different areas to assess research material and during this we discovered a lack of research considering national tests in northern Norway. Therefore, we decided to focus on how national tests in northern Norway is perceived by teachers and school leaders within the region. The reason for us focusing on a northern Norwegian perspective is that we both come from this region of Norway and wish to contribute to the research field.

Furthermore, through our education, we have experienced and undergone national testing in English, leaving us with the impression that there were clear deficiencies regarding the topic. National testing is a significant tool in Norwegian education where the majority have either experienced it or heard about it, which increased our area of interest. We perceived national test as a significant tool, which we felt was utilized differently across each school, which we wished to further examine. National test in English is held at 5<sup>th</sup>- and 8<sup>th</sup> grade consistently, and this led us to wanting to see where the deficiencies are and why it is utilized differently among educators despite its significance. Furthermore, our area of interest increased after numerous articles were published with the idea of phasing out national test in Norwegian education. Two of these articles were published by Utdanningsnytt and NRK, which mentioned the possibility of discontinuing the tool. NRK published an article where several political

parties wished for shaping out of national testing, however one of Norway's biggest political parties wished to keep it due to the need of having knowledge about schools to ensure knowledge in schools (Alnes et al., 2023). The Committee for Quality Development (Utvalg for kvalitetsutvikling) emphasizes their wish of removing it due to the lack of improvement of results and it being a resource-heavy tool not facilitating a desired outcome for neither teachers nor pupils (Ruud, 2023).

## 1.2 Thesis- and research questions

As presented above, standardized assessments is a controversial topic, with issues such as preparation, implementation, and the extent to which schools utilize the gathered results. Therefore, our aim is to explore how the results of the national test in English are implemented as a mapping- and development tool within a selection of schools in Northern Norway.

Previous research indicates that a marginal percentage of teachers have the skills required to analyze the results of national tests (Werler & Færevaaag, 2017, p. 69). This implies that the post-work regarding national tests is influenced of the school's ability to interpret and understand the outcomes (ibid.). Grounded in the context of the national test in English and its potential to be a sufficient tool in the school's development in regard to preparation, implementation, and to what extent schools can leverage the obtained results, we aim to examine how this process is integrated as part of the school's development. Additionally, it will be beneficial to explore the amount of time teachers invest in preparing the pupils for the national test in English.

Based on this, we have chosen the following thesis question: *to what extent do a selection of schools in northern Norway use the results from national tests as part of the school's academic development?*

The thesis question is substantiated by three research questions:

1. To what extent do English teachers and school leaders experience that the national test in English is a prioritized tool in the mapping and academic development of the pupils

2. To what extent does established routines affect the utilization of the national test in English as a mapping and development tool by teachers and school leaders?
3. To what extent do English teachers and school leaders consider the national test in English as a significant tool in academic development of the pupils?

### **1.3 Structure of the thesis**

The thesis is divided into 12 chapters, each with sub-chapters. Presented below is the main content of the chapters.

In chapter 2, previous research is presented. The research regards standardized testing and the use of data, with the addition of teachers' and school leaders' view of standardized testing.

In chapter 3 we present various aspects of the national tests in Norway, including an overview of the purpose of the tests, the expectations of the ones in charge of administering the tests and, the guidance materials provided by UDIR. Additionally, we present criticism of the national tests, and the phenomenon 'teaching to the test'.

Chapter 4 presents relevant theory. The school can be seen as an organization; therefore, we chose organizational theory as the theoretical background for this thesis. Senges' five disciplines of a learning organization and Dalin's five dimensions in school as an organization will be used as the basis for the discussion of our findings.

Chapter 5 involves an overview of the research methods that were chosen for this thesis. We present a general overview of mixed methods design, with the addition of transcribing, validity, and reliability. Furthermore, we present theory regarding the analyzation and interpretation of qualitative and quantitative data.

Chapter 6 regards how our role as researchers could affect the data, and the strengths and weaknesses of our research, involving the collection, analyzation, and interpretation of the data. [In addition, we present theory regarding the insider vs. outsider myth.]

Chapter 7 contains a presentation of the results from the qualitative and quantitative research that was collected for this thesis. Here, we chose to split the results into sub-

chapters that lay the ground for the discussion that is presented in the following chapter.

In chapter 8, we analyze and discuss the data that is presented in chapter 7, using the theories from chapter 4 and the previous research presented in chapter 2 as a foundation. The findings from our research made it natural for us to split the discussion into sub-chapters that are relevant to answer our thesis- and research questions.

In chapter 9, we answer the three research questions using the findings from our discussion in chapter 8, theories from chapter 4, and the previous research presented in chapter 2 as a foundation. This chapter leads us to the following chapter, where we answer our thesis question.

In chapter 10 we answer our thesis question; *to what extent do a selection of schools in northern Norway use the results from national tests as part of the school's academic development?* To do so, we draw conclusions from our discussion in chapter 8 and the answers to our research questions.

Chapter 11 contains suggestions for further research on the topic of national testing and the use of its data for mapping and development of the English subject.

In chapter 12 we present concluding remarks for our master thesis.



## **2 Previous research**

### **2.1 National research**

#### **2.1.1 Seland, Vibe & Hovdhaugen (2013)**

In a study commissioned by the Nordic Institute for Studies of innovation, research, and education (NIFU), Seland, Vibe & Hovdhaugen (2013) points to several findings. One of the main findings of the study is that the way teachers mention the national tests could have an impact on how the pupils regards the tests (2013, p. 24). If the teacher seems worried or negatively tuned towards the test, pressure or fear could be infused in the pupils. If the teachers subjective meaning is that the national testing is of little or no value, the pupils' motivation could decrease.

Further findings show that teachers express frustration due to insufficient information in the follow-up of the pupils after the tests (Seland, Vibe & Hovdhaugen, 2013, 10). The results from the tests are published, and some school leaders' view this as demanding for teachers, pupils, and parents because of the un-nuanced picture it might give of the school. At the same time, the study shows that many school owners and leaders views national tests as sufficient tools that measure what they are intended to measure. The same cannot be said for teachers, that seem more restrained towards national testing, with the view that the tests are not necessarily a good tool in the development of schools (Seland, Vibe & Hovdhaugen, 2013, p. 27). Lastly, it is stated that less than a third of the municipalities that are part of the study checks that the schools have followed the exemption rules given by UDIR (Seland, Vibe & Hovdhaugen, 2013, p. 51).

#### **2.1.2 Werler & Færevaaag (2017)**

In the article "National testing data in Norwegian classrooms: a tool to improve pupil performance?", published in 2017, Werler and Færevaaag (2017, p. 67) points out that official regulations demands that all results from the national tests should be used as part of the quality development of the relevant subjects, even though the tests might not provide the necessary information to do so. Werler & Færevaaag (2017, p. 69) further state that due to the limitations of the information that is provided through the test results, it might be difficult for the teachers to pinpoint which variables they should

change to improve pupils' learning outcome. Based on the observations made through the research, Werler and Færevaag (ibid.) claim that it could be argued that teachers have to *guess* what could be the cause of poor test results in order to improve pupils' learning outcome.

Furthermore, Werler and Færevaag (Werler & Færevaag, 2017, pp. 68-70) state that the national tests have contributed to hold municipalities and schools accountable for each individual pupils' learning. Despite this, the majority of teachers might not be familiar with contextual factors at the respective schools. These factors consist of, among others, pupil demographics and local conditions which play a significant role in the results of the tests. In addition, Pierce and Chick (2011, cited in Werler & Færevaag, 2017, p. 69) state that although it might seem reasonable to assume that Norwegian teachers can read and understand features such as graphs or tables, it is unlikely that the teachers possess the knowledge that is required to compare, contrast and critique multiple datasets.

### **2.1.3 Parliamentary notice 21 (St. Meld. 21, 2016-2017)**

In parliamentary notice 21 (Meld. St. 21), published by the Norwegian ministry of knowledge in 2017, it is stated that the school, like many other areas of society, is governed from various levels. However, central responsibilities, rooms for action and connection lines are not clearly defined (Meld. St. 21, 2016-2017, p. 11). Meld. St. 21 (2016-2017, p. 12) refers to international research that shows that the decentralization of certain decisions, such as the distribution of responsibilities and how to develop the teaching has proven positive when it comes to the pupils learning, given that the local level possesses the necessary knowledge and will to take such responsibilities.

Cooperation between teachers, school leaders and school owners is set forth as key to develop better quality teaching practice, based on skills and experience (Meld. St. 21, 2016-2017, p. 13). Meld. St. 21 (ibid.) states that teachers, school leaders and school owners should have the lead role in the task of developing the quality in the school, and that the state should provide the necessary frames around the local action space. The task of improving the quality of teaching has the best conditions in schools where teachers cooperate, and Meld St. 21 (2016-2017, pp. 26-27) stresses that pupils learning is dependent on that the teachers themselves develop. Although the teachers seem to

have an important role in the development and improvement of quality in teaching, the school leader should communicate and anchor the schools' goals, ambitions, and values, and lay the grounds for a healthy development culture.

The development of cooperation in Norwegian schools is not optimal, and teachers' experiences show that there is relatively little cooperation. Additionally, school leaders and teachers experience little contribution from school owners when it comes to competence development. However, Meld. St. 21 (2016-2017, p. 28) points out that Norway meets the international average in this matter.

## **2.2 International research**

### **2.2.1 Wayman & Jimerson (2014)**

In their article 'Teacher needs for data-related professional learning', Wayman & Jimerson (2014, p. 25) address two questions: (1) What skills do teachers need to use data effectively? (2) How should teachers receive data-related professional learning? Through their study, these questions were addressed using qualitative data from 110 participants. Wayman & Jimerson point out that "research indicates that educators struggle with using data to inform practice, citing issues such as data systems, principal leadership, time and a lack of knowledge about how best to use data to improve instruction". As Wayman & Jimerson (ibid.) points out, teachers report that although they are frequently exposed to professional learning about data use, little of it meets their practical needs. Wayman & Jimerson (2014, p. 26) consider *data* to be "any information that helps educators know more about their pupils and which can be codified in some manner", and *data use* to be "the actions in which educators engage as they collect, organize, analyze and draw meaning from these data in efforts to inform practice".

Collaboration, common understanding, triangulation, and time are skill areas that Wayman & Jimerson (through previous research) suggest are important skill areas for teacher data use. Furthermore, Wayman & Jimerson (2014, p. 27) review research that suggests that "educators benefit from professional learning activities that are: (1) collaborative, (2) engaging, (3) contextual, (4) job-embedded, (5) intense, and (6) coherent". Through the interviews conducted by Wayman & Jimerson (2014, pp. 31-32),

it becomes evident that the knowledge of how to ask appropriate questions of the data, how to analyze and interpret their data, how to link data to practice, and how to efficiently navigate computer data systems is key in teacher data use. According to Wayman & Jimerson, the collaboration aspect is an essential element of data-related professional learning. However, the study showed few structures that ensured consistent collaboration, with educators often citing the importance of collaboration without suggesting the need to learn the necessary skills to ensure effective collaboration. Wayman & Jimerson (2014, p. 33) conclude that many solutions to the presented issues are systematic in nature, and that making changes to organizational practices could be the best way to build capacity in individual teachers.

### **2.2.2 Tonich (2021)**

With the purpose of determining the effect of principals' leadership abilities on school performance, both directly and through their schools' organizational culture, Tonich (2021, p. 54) underwent a qualitative study, comprised of 350 school principals. Tonich (2021, p. 63) argues that principals are of clear importance in organizing school life in order to achieve optimal outcomes, due to them often doing their best to serve both staff and pupils, with knowledge of both their duties and of how to 'set the rhythm' for the school. Leadership is defined by Tonich (ibid.) as "having the ability to use all the available resources in an organization in the best way possible to achieve the stated goals. Citing Robbins (1990), Tonich (2021, p. 65) states that organizational culture is desirable to improve the performance of the school. Furthermore, through being what Peterson (2013) and Singh (2014) calls "the invisible hand guiding people's behavior", strong leadership in an organization is of high importance, because it determines the organizational culture (Yuan & Lee, cited in Tonich, 2021, cited in Tonich, 2021, p. 65). Moreover, Tonich (2021, pp. 65-66) argues that effective leadership is a main precondition for creating a conducive organizational culture that can strengthen systems that may otherwise fail, with his study finding that a good organizational culture fosters high levels of performance and improves the moral of teachers and pupils.

According to Tonich (2021, p. 66-67), the principal is responsible for fostering an organizational culture for education in order to improve the school's performance. As

part of this, principals should master and understand all aspects of their role as an educational administrator. Furthermore, leaders who can act as effective agents of change could be part of triggering improvements in organizational culture, further improving employee and organizational performance (ibid.). Tonich's study concludes with him finding that a principal's managerial ability exerts a significant influence over both the school's organizational culture and overall school performance. However, Tonich (2021, p. 68) notes that the optimization of the managerial abilities should be prioritized over improving a school's organizational culture and relates this to the fact that these abilities make a greater direct contribution to the performance of schools, compared to the indirect contribution of organizational culture.

### **2.2.3 Hardy, 2015**

With the intention of drawing upon the experiences of teachers and principals in Queensland, Australia, Ian Hardy (2014, p. 2) intends to reveal how political and policy contexts influences schooling practices and the contested nature of such practices. *NAPLAN*, the equivalent to the national tests in Norway, was initially tested in the Queensland area in 2008, with relatively poor outcomes. Due to these results, Queensland schools were recommended to engage in test readiness activities, with the aim of improving the test results. Citing Ball et. al. (2011), Hardy (2014, p. 3) argues for greater attention to schooling contexts when it comes to enacting policies in school. In addition, Hardy (ibid.) states that the "histories and ethos of schools also matter, as do the results of mediations between schools, governments and local/ regional authorities.". Hardy (2014, p. 4) emphasizes that policies which demands compliance from teachers for reasons of necessity often leaves little room for teacher judgment, with the effect of making them tired and overloaded much of the time, despite their creativeness in their approach to managing and implementing such policies.

Through 55 individual interviews with teachers and principals from three schools in Queensland, one in a rural area and two in metropolitan areas, Hardy (2014, p. 17) found that the teachers and principals in question, despite of strong national and state policy and political pressure to improve *NAPLAN* results, sought to appropriate performative demands in a more general fashion, to further assist them to focus on the educational nature of their work. Hardy argues that "teachers, principals, system personnel and

other policy- 'makers' need to recognize the intrinsic nature [...] of such appropriation capacities to understand how strong policy support for improved test scores may play out in practice, to inform subsequent policymaking.”

#### **2.2.4 Beck & Stetz, 1979**

In 1979, Michael Beck and Frank Stetz underwent a national study regarding standardized testing in the United States of America. The purposes of the study were to explore “1) teachers’ sentiments regarding the amount of standardized testing in their school systems, 2) their uses of standardized achievement test results in their classrooms, and 3) their opinions concerning the usefulness of standardized test results for various purposes” (Beck & Stetz, 1979, p. 2). A questionnaire was developed with four questions that concerned: “1) teacher’s opinions of the amount of standardized testing in their school systems; 2) particular uses made by teachers of standardized achievement test results; 3) their views of the usefulness of a variety of possible application of test results; and, 4) their opinions on various test moratoriums and other test related policies” (ibid.). A national sample of 3300 elementary and secondary school teachers who had administered the Metropolitan Achievement Tests (MAT) in their classrooms responded to the questionnaire. Although the teachers had recently administered the MAT, they were instructed to answer the questions in the survey based on their attitudes toward standardized tests in general. Teachers from small (under 500 pupils per grade) and large (over 500 pupils per grade) public schools and non-public schools answered the questionnaire.

Firstly, 69% of the respondents answered that the amount of standardized testing in their school was “about right”, with teachers in non-public schools generally being more satisfied with the amount of testing, whilst teachers in large schools tended to consider the amount of testing too great (Beck & Stetz, 1979, p. 4). Secondly, about 10% of the teachers made considerable use the results from standardized tests in their classroom, whilst about 50% made “some” use of the test data (ibid.). Lastly, the majority of frequent uses of the test data was 1) diagnosing strengths and weaknesses, 2) measuring growth, and 3) individual student evaluation. Furthermore, the of teachers that answered the survey considered standardized tests useful for helping to plan instruction and for measuring the educational status of individual students.

## 2.3 Chapter summary

In this chapter, we have presented previous research regarding standardized testing, the use of data and how the school leaders could affect the organizational culture of a school. The research shows that important factors of using data for improvement are cooperation and triangulation, and further shows the need for better knowledge and leadership when interpreting data. The results from the national tests are limited, consequently making it difficult for educators to pinpoint what can be done in order to improve the pupils' learning outcome. Whilst the teachers do play an important role in the quality development in the schools, the research shows that school leaders must be clear on the schools' goals and values in order to lay the ground for quality development. Directives implemented from higher authorities, with little room for creativity, could have the effect of teachers feeling over worked and less motivated. Therefore, school leaders should act as agents of change, with a goal of both academic development and the increase of motivation in the teachers. Central responsibilities, rooms for action and connecting lines regarding the national tests are often distributed within the school as an organization. Whilst this could be positive, the local levels must possess the necessary knowledge and will to take such responsibilities. Furthermore, the research shows a discrepancy between how teachers and school leaders view the national tests, with school leaders often being more positive towards the tests. Moreover[?], research conducted in the late '70s found that teachers considered standardized testing as mostly positive, and as a good tool for factors such as the measuring of educational growth within the pupils and planning instruction.

### 3 National tests

The purpose of the national tests is to provide schools with insights into pupils' fundamental skills in reading, mathematics, and English. The information derived from these tests are intended to serve as a foundation for formative assessment and quality development at all levels within the educational system (UDIR, 2022c). Whilst the national tests in reading and calculation are conducted in 5<sup>th</sup>, 8<sup>th</sup>, and 9<sup>th</sup> grade, the national test in English is only conducted in 5<sup>th</sup> and 8<sup>th</sup> grade. In the Norwegian primary and secondary school, reading and calculation is regarded as basic skills, whereas English is not. Therefore, the national test in English focuses on competence aims from only one subject, specifically reading- and listening competence, vocabulary, terminology, and grammar (ibid.).

Teachers are expected to use the results to monitor and provide ongoing assessment for their pupils', in addition to individualized instruction. Municipalities and schools are to utilize the results as a basis for enhancing the quality of education. The tests provide information regarding individual pupils, groups, grade levels and schools, which teachers and school leaders require for the ongoing development of their school. The result from the national tests gives a restricted view of the skill and competence each pupil possesses. Therefore, it is necessary to consider the results in conjunction with other available information about the school, municipality, or the pupils.

The assessments are electronic and consists of various elements such as texts, images, and tasks with questions. It takes up to 60 minutes to complete a national test in English. The general rule regarding national tests is that all pupils are expected to participate. However, pupils with the right to special education or specialized Norwegian language instruction may be exempted from these assessments (UDIR, 2022b).

Subject experts from universities, colleges and national center collaborate with teachers and experts at the Norwegian Directorate for Education in developing the tests. These subject experts base their work on a framework outlining the content and technical specifications of the assessments. It takes several years to develop an assessment. Tasks are tested multiple times to ensure they function as intended, and the assessment as a whole measures what it is supposed to measure (UDIR, 2023c, p. 3) The Norwegian



Directorate for Education is responsible for ensuring that the assessments are developed in accordance with the quality requirements outlined in the framework for national assessments and that they undergo sufficient quality assurance.

Sjøberg (2014) explores the impact of the Programme for International Student Assessment (PISA) on Norwegian education policy, particularly in the aftermath of the PISA survey in 2000, more known as the “PISA shock”. His article delves into the transformative influence of PISA results on the Norwegian educational landscape, instigating substantial policy changes. It highlights the ensuing critiques of the Norwegian educational system and the subsequent introduction of a national quality assessment system, incorporating standardized national tests as a pivotal component.

The PISA test, conducted triennially by the Organisation for Economic Co-operation and Development (OECD), gauges the proficiency of a sampled cohort of pupils in reading, mathematics, and science. As argued by Sjøberg (2014a, p. 30), the PISA-project shapes the public view of the schools, and sets the premises for educational policies, as shown by the effects of the 2000 PISA results, widely publicized as the “PISA shock” in Norwegian media, that catalyzed a shift in the country's educational policies. Notably, the middling performance of Norwegian pupils compared to their OECD counterparts prompted severe criticism and an urgent call for educational reform.

Eivind Elstad (Sivesind & Elstad, 2010, p. 100) underscores PISA's evolving role as a central reference framework for assessing the quality of the Norwegian educational system . International assessments, particularly PISA under the auspices of the OECD, significantly influencing Norwegian educational policies. The OECD's perspective regards education as an investment in human capital, fostering various values and productivity essential for economic growth. Svein Sjøberg (2014, p. 196) contends that the PISA project is more about politics, emphasizing globalization and market dynamics, rather than pedagogy. In the neoliberal, globalized market economy, education is regarded as a competitive advantage, leading to an emphasis on measurable outcomes, notably national tests.

Sjøberg (2014a) argues that the framework for national tests underscores their role in assessing fundamental skills in reading and mathematics, emphasizing that these tests are not subject-specific but aligned with the overarching principles of the curriculum

(UDIR, 2020a). In the domain of reading, national tests evaluate pupils' ability to extract information from texts, interpret it, and synthesize the information into a reflective form (UDIR, 2020b).

While national tests in English are subject-specific, aligning with curriculum objectives, they are also designed to assess fundamental skills in listening and reading. The results are categorized into various proficiency levels, with different delineations for 5th, 8th, and 9th grades. The tests underwent a redesign in the fall of 2022, introducing new score thresholds for the proficiency levels. To facilitate year-to-year comparisons, scalar score thresholds were established to indicate the proficiency level, with lower levels denoting lower proficiency (UDIR, 2023c).

Lastly, Sjøberg (2014a) concludes his article with an exploration of the post-assessment phase, emphasizing the role of the Norwegian Directorate for Education and Training (UDIR). UDIR advocates for school owners and leaders to deliberate on the reasons behind the results, formulate strategies for follow-up actions, and implement necessary interventions. Additionally, UDIR highlights the importance of comparing results for the same student cohort across different years for a comprehensive assessment of educational progress (UDIR, 2022b).

### **3.1 Basic documents and guidance materials**

As part of the preparation for implementing the national tests, teachers are advised to familiarize themselves with the guidance materials provided by UDIR. The guidance materials are divided into five main parts, with each part containing further information and links to useful information. The guidance materials can be accessed on UDIR's public web page without any log-in information.

A guide is provided regarding how the pupils should be prepared for the national tests. here, UDIR (UDIR, 2023b, p. 1) states that although the pupils are not intended to practice for the national tests, they should be prepared for the tasks. According to UDIR (ibid.), it is beneficiary to look at example tasks with the pupils or to look at task sets from previous years. The ones responsible for preparing the pupils for the national tests (i.e. the teachers) should take care as to how the pupils could experience the implementation of the national tests in the best way possible, and UDIR (ibid.) stresses

that since the tasks are of varied academic levels, the pupils should be made aware that they might not be able to answer all tasks. Nevertheless, UDIR (ibid.) states that the pupils should be encouraged to answer all tasks as good as they can, emphasizing that wrong answers do not entail minus points. If the pupils are unsure of what is the correct answer, they are meant to answer what they *think* is the correct answer. As stated by UDIR (2017, p. 2), the results from previous tests show that the average pupil answers 50-60 percent of the tasks correctly. In addition to preparing the pupils for the national tests, UDIR stresses that the parents should be informed of 1) when the national tests are to be carried out, 2) in which grade levels and for which competencies the national tests are conducted, 3) what the purpose of the national tests are, and 4) the results from the national tests (UDIR, 2023b, p. 2).

Secondly, an overview of how the national tests should be administered has been developed by UDIR. Divided into three parts, this guide elaborates on important matters such as preparation and registration, implementation, and reviewing results. Here, all responsibility areas are listed, divided into school owners, school leaders and teachers. Summarized, it is the school owners' responsibilities to oversee that all practical areas of the tests are in order, such as making sure that all pupils are registered in the correct systems, making sure that there is adequate human- and technical resources, and monitoring that the implementation of the national tests is done in a correct manner.

The school leaders have the responsibilities of making sure that enough time is set aside for both the preparation for the tests, ensuring that the school's computers and network undergo necessary testing and preparation, and the responsibility of informing the pupils' parents of both the implementation- and results from the tests. In addition, the school leaders are responsible for facilitation for pupils that have special needs and facilitating the implementation of the national test results by the teachers in the classroom.

The teachers are responsible for the preparation of the pupils towards the test, including setting aside time for this. Although the pupils shouldn't receive any assistance in solving the tasks, the teacher should provide the pupils with support and assurance during the test. Additionally, one of the responsibilities that the teachers have is to use

the test results as an active part of the academic feedback given to both the pupils and their parents, with the aim of promoting further learning (UDIR, 2023a, pp. 4-5).

As part of the guidance materials, UDIR has developed examples of how the schools could work with the teaching of the pupils, giving examples for each of the grade levels and competencies that the national tests measure. The examples are based on the core elements in the curriculum LK20, and involves specific focus areas that can be used in teacher-pupil conversations, examples of activities that can enhance oral communication amongst pupils (in the classroom), and what UDIR refers to as “guided reading”, an activity that gives the pupils a chance to work at their own academic level, and to get help with developing their individual learning processes.

To help the teachers and school leaders understand what the national test in English measure, and what the proficiency levels entail, UDIR has developed a guide explaining these factors. In-depth paragraphs for each of the competence aims (from LK20) provide the information needed to understand what the tests are designed to measure. Included in these paragraphs are also elaborations regarding certain areas of the tests that measure multiple competencies simultaneously (UDIR, 2017, p. 2). To explain what the different proficiency levels entail, a guide has been developed to give clear instructions regarding this. For instance, proficiency level 5 for pupils in the 9<sup>th</sup> grade entails that the pupil is able to understand long and complex sentences, and that he or she is able to utilize reading- and listening strategies that are appropriate (UDIR, 2022a, p. 5).

### **3.2 Teaching to the test**

In the context of what standardized tests such as the national test in English concerns, it is necessary to explain the concept “teaching to the test”. McMillan (2000, cited in Volante, 2004, p. 1) states that when standardized tests are equipped appropriately, it helps teachers identify strengths and weaknesses within the pupils, furthermore, strengthening its purpose of measuring learning outcomes and skillsets. Mitchell (1997, cited in Volante, 2004, p. 1) emphasizes that standardized tests is perceived as the most important measure of student performance, moreover, expressing a concern regarding how politicians, school personnel, administrators and teachers have begun to employ practices that Mitchell (1997) clearly perceives not being in the best interest of the pupils. Furthermore, he explains his vision by stating that the practices are too focused

on test content and therefore eliminating other curricular content which is seen as important. Test contents are usually released in order to prepare the pupils and teachers of its format (Volante, 2004, p. 1).

Providing teachers and pupils with general information regarding the tests format is not questioned, however, giving away test items to ensure good results is (Volante, 2004, p. 2). Teaching to the test involves a great amount of classroom time due to the great number of worksheets, drills, practice tests being utilized. People who view teaching to test as negative, often state the concern of only using part of the curriculum, whereas basic-skill subjects and high-order thinking skills are negatively prioritized (Herman, 1992, cited in Volante, 2004, p. 2). Despite teaching to the test possibly enhancing test results, research suggests that the learning outcome will not change. Furthermore, an important aspect of its criticism is that teaching to the test reduces the depth of instructions given in specific subject, while narrowing the curriculum further aggravating non-tested disciplines such as creative and physical education (Volante, 2004, p. 2). Lastly, teaching to the test concerns a great amount of priority, which may affect the time of other subjects and important areas.

Teachers spend a large amount of time preparing for standardized tests, focusing directly and indirectly on techniques that are aligned with “teaching to the test”. To develop and become more including in the aspect of teaching, it is dependent on various factors. Firstly, administrators and school personnel need to be skeptical regarding results that are outside the norm for the pupils and schools (Volante, 2004, p. 3). It is not reasonable that a student with poor results, receives a high score of a standardized test. Secondly, school administrators should receive appropriate training to further develop their leadership and have the ability to convey this knowledge to their teachers. Lastly, teachers should receive adequate training in teaching within the curriculum, which requires them to direct their instructions to a specific set of skills (Popham, 2001, as cited in Volante, 2004, p. 3).

### **3.3 Criticism of the national tests**

National tests have become an integral part of Norway’s educational system and are often used as a tool to evaluate pupils’ knowledge and performance across regions,

counties, cities, and schools. Despite their widespread use in Norwegian society, national tests have been the subject of considerable criticism from teachers, pupils, and educational experts.

The public disclosure of the national test results has generated significant debate across various platforms, particularly in the media, where it has received extensive coverage. Furthermore, the public disclosure of individual schools' results in the media could also lead to the stigmatization of schools, as well as pressure on each school to direct teaching towards the "teaching to the test" phenomenon to achieve desired outcomes of the tests, rather than focusing on learning itself. A consequence of this is that the curriculum could be narrowed to fit the national tests (Nusche et al., 2011, p. 55). The results of national tests are published online and are accessible to the public, which can lead to media outlets ranking schools and portraying them in a one-sided manner (Tveit, 2007, p. 33). According to Willie (2010, p. 75), ranking based on national tests scores has no learning-enhancing effects for pupils. The results from national tests serve as a basis for providing an overview of the student population, but they reveal little about the quality of individual schools, particularly given contextual factors such as region, city, etc. (Tveit, 2007, p. 35)

Standardized testing remains a contentious issue in education today, and many argue that it weakens creativity. Scores generated by state assessments are used for political purposes to compare pupils, institutions, and teachers. Standardized testing has always had a major impact on education, but it now impacts an area in which pupils have opportunities to display creativity in their education (Longo, 2010, p. 55). Even exceptional teachers "teach to the test" without even realizing it. Excellent teachers satisfy the requirements of state assessments without spoon-feeding the content. By using an inquiry approach, educators can combine both content and process skills, thereby preparing pupils for standardized testing while still maintaining creativity in the classroom (Longo, 2010, p. 56).

One of the most prominent criticisms of national tests is that they can promote a narrow understanding of education, where academic achievements are measured by test scores, while important skills such as creativity, critical thinking, and social abilities are often overlooked (Tveit, 2007, p. 36). This can lead to a teaching culture where teachers feel

pressured to “teach to the test”, potentially reducing the quality and variety of instructions. Publicly disclosing the results can also lead a school with poor outcomes to aim for improvement, with the skills targeted by national tests receiving higher priority among teachers and the schools’ administration. However, the test results themselves provide no insight into the overall quality of a school’s instruction (ibid.). National tests have also been criticized for addressing only the educational mission of schools, neglecting the broader purpose of fostering holistic personal development. “Schools have both an educational mission and a mission of personal development. They are interconnected and mutually dependent. The principles for working with learning, development, and character formation are intended to help schools fulfill this dual mission” (UDIR, 2018).

The preparations that schools undertake in connection with the national tests are believed to significantly impact the results, as some schools spend considerable time specifically practicing for the national tests, consequently deprioritizing many other aspects (Marsdal, 2011). Teachers are expected to prepare pupils for standardized assessments while still providing creativity, and this is seen as a challenging task (Longo, 2010, p. 54). Furthermore, pupils’ express concerns with the unknown of what will face them in the assessment, resulting in teachers often being criticized for “teaching to the test” and enabling pupils. State assessments such as national tests can affect the most experienced teachers, steering them towards a “teaching to the test” method in order to ensure good results (ibid.). Kvaale (1970, cited in Engelsen, 2012, p. 125) points out that assessment has a power function, allowing for the control and influence of other people. Eggen suggests that assessment indicates what constitutes important knowledge, similar to the way curricula do (Eggen, 2009, p. 92). This, in turn, gives national tests a position of power that can influence what individual teachers prioritize in their teaching.

Furthermore, national tests can promote stress and pressure among pupils, and those who do not perform well may experience reduced motivation and personal mastery. National tests can create unfair comparisons between schools, as they often do not consider differences in resources or student composition. This can result in some schools being unfairly criticized or stigmatized, which may impact teacher morale and the school’s reputation.

In order to maintain creativity while preparing the pupils for state assessments, inquiry learning models facilitates this. The implementation of an inquiry learning model can stimulate creativity in the classroom, while still preparing pupils for high-stakes state assessments (Longo, 2010, p. 54).

### **3.4 Chapter summary**

The national tests are conducted annually in 5<sup>th</sup>, 8<sup>th</sup> and 9<sup>th</sup> grade, with the exemption of the national test in English which is conducted in 5<sup>th</sup> and 8<sup>th</sup> grade and serve the purpose of providing schools with insight into pupils' fundamental skills in reading, mathematics, and English. The tests should serve as a foundation for formative assessment and quality development and were implemented after the "PISA shock" in 2000. UDIR has developed guidance materials for all phases of the national tests, and teachers and school leaders are advised to familiarize themselves with the guidance materials. Although teachers are not advised to practice for the tests, even exceptional teachers "teach to the test" without even realizing it. A consequence of this could be that the test results become less valid. The national tests have been the subject of criticism, with one of the most prominent criticisms being that the tests could promote a narrow understanding of education, where academic achievements are measured by test scores, while skills such as creativity, critical thinking, and social abilities often being overlooked.



## 4 Theory

### 4.1 Organizational theory

#### 4.1.1 School as an organization

Compared to landscapes by Dalin (2005, p. 29), organizations are complex and varied. As he points out, organizations- like landscapes, are complex and varied, and change with our observational perspective. Organizations could be seen as 1) an organism in the way that they are dependent on their environment and 2) as a brain in the way that they inherit the ability to process data and that they 'learn to learn', or as culture- where values, norms and rituals are administrated (Morgan, 1988, cited in Dalin, 2005, pp. 29-30). Knut Roald (2012, p. 115) states that both professional literature and educational policy documents takes it for granted that one can view schools within organizational theoretical perspectives, and points to the fact that organizational perspectives has evolved into a more central part of professional literature and educational research (Roald, 2012, p. 116).

Referred to as *loose connections* or *loose couplings* ("løse koblinger in Norwegian) by organizational sociologists, the decisions regarding what is to be taught and which methods that are used, are taken by the respective school leaders and teachers, with little or no influence from the state or municipality (Fevolden & Lillejord, 2005, pp. 100-101). There is, however, a national educational policy (referred to by Fevolden & Lillejord as "guidelines") that encompass general questions such as the age at which school starts, the distribution of responsibilities and what the pupils should generally learn at school (ibid.). Roald (2012, p. 117) states that these loose connections implies that important decisions are made by those responsible for the practical implementation, i.e. the teachers. For instance, the organization, execution, evaluation, and development of teaching is done mostly by teachers (ibid.). In effect, the only realistic opportunity for change as proposed by the government is by appealing to what is described by Fevolden & Lillejord (2005, p. 158) as volunteering. As Dalin (2005, p. 45) points out, most schools are characterized by a number of units (i.e. classrooms) that are isolated from each other, by initiatives that have no practical consequences and by guidelines that are not followed. An important question put forth by Dalin (ibid.) is

whether or not this is good or bad. However, Fevolden & Lillejord, points to two consequences regarding this; firstly, changes that are made often match the personal values and attitudes. Secondly, said changes are more or less disconnected from the system of the school or the school system as a whole. This, according to Roald could explain why 1) the majority of changes made at an overall level have little or no practical effect in schools, 2) why national development programs have little or no documented effect in schools (2012, p. 117). However, Dalin (2005, p. 45) argues that the relative freedom provided by loose connections (or loosely coupled systems, as he calls it) might help certain units adapt to the requirements of their environments, whilst the rest of the organization remains stable.

#### **4.1.2 Quality in school**

According to Møller & Ottesen (2011, p. 15), one third of the population in Norway is somehow connected to the Norwegian educational system, either as pupils or as employees. This, with the added factor of education being one of the largest expense posts in public administration, makes the public school system something that affects most, if not all inhabitants. In order to secure future welfare, develop a strong democratic society and securing integration and the individuals change of life fulfilment, Møller & Ottensen (ibid.) state that a good quality school is key. There is, however, no single answer when it comes to what can be done to secure that the educational system is of good quality, and international surveys such as the PISA test has shown that there is a lack of quality in Norwegian schools (Ibid.).

As Fevolden & Lillejord (2005, p. 9) points out, the Norwegian school system constantly faces reforms and changes. In a society that is increasingly based on knowledge and scientific insight, an important prerequisite for the educational system is that school owners, school leaders and teachers adapt so as to keep in line with the changes in society (Fevolden & Lillejord, 2005, p. 9; Roald, 2012, p. 119). As well as being able to adapt to the changes in society, the schools must possess the knowledge to observe strengths and weaknesses in their own organization, so that they see where change is needed when it comes to their own development and improvement. Furthermore, school owners and school leaders, as well as the teachers themselves, should work within the schools to create an environment where there is a culture of continuous

learning, meaning that all individuals that work within the school should have knowledge about how competence can be developed, and what it means to possess knowledge about something (Fevolden & Lillejord, 2005, p. 9). Developing a consciousness about these factors could be key to securing good quality schools and education.

## 4.2 Senge's five disciplines

Peter Senge presents five core disciplines of a learning organization in his book "The Five Disciplines", first issued in 1990. Here Senge argues that problems caused by modern organizations often stem from a lack of systematical thinking, combined with the need for a general and mutual view of values and visions for those within the organization (Senge, 1997, pp. 48-51). Dalin (2005, p. 51) claims that Senge's systematic perspective derives from the humanistic perspective of organizations (i.e. the perspective that is concerned with the individual's contribution in organizations), but that he goes beyond it. Dalin (2005, p. 63) also states that Senge's work is important "because it has something to say about the important relationships of the processes that develop a 'learning organization', which should be of particular importance for schools."

Firstly, Senge (1997, p. 49) emphasizes the need for *systems thinking* in the organization. The ability to understand connections and patterns in an organization (what Senge calls "seeing *wholes*") implies "seeing where actions and changes in structures can lead to significant, enduring improvements". According to Dalin (2005, p. 51), the ability to see an organization as a whole could sway the organization in the direction of meaningful change. To do this, Senge calls for the need of circular thinking, with a view of an organization as a system where cause and effect are woven together, as opposed to the more traditional linear, casual view (Roald, 2012, p. 129; Senge, 1997, p. 48). As part of this, Roald (2012, p. 129) states that the individuals within an organization needs to see themselves as part of both the problem and the solution, so that they can take part in the development of the organization rather than putting blame on others.

The second discipline presented by Senge is *personal mastery*, which concerns the ability to be both self-aware and realistic at the same time (Roald, 2012, p. 129; Senge, 1997, p. 50). Citing Senge, Roald (2012, p. 130) points to the argument that the process of learning in an organization should be based on rational thinking and intuition, and that

the main goals and contexts of the organization should be more important than details. Senge argues that a prerequisite for fostering personal mastery in an organization is “a corporate culture must be established that rewards investigation and enquiry and which empowers people to experiment.” (Senge, 1997, p. 50).

Thirdly, *mental models* are presented as part of the work towards change in an organization. Senge argues that mental models are a crucial part of organizations because they affect the individual’s ability to learn and adapt. When accurate and flexible, mental models enable the individual to navigate complex situations effectively. Senge argues that mental models cannot be changed, but that they can be managed. Through dialogue and honesty, individuals within the organization could acknowledge that mental models exist on both sides, and that both ways of thinking should be discussed and tested, so that the organization could build new mental models or maintain the existing ones. This, according to Senge, could have the result of a position that is “more likely to reflect the true situation” (Senge, 1997, p. 50). Roald argues that in some organizations, even successful measures have no consequence for how the organization implements change over time because the mental models in the organization prevents change and development over time (Roald, 2012, p. 130), making the questioning of assumptions and openness towards alternative perspectives important in order to improve the way the organization thinks, makes decisions and solve problems.

The fourth discipline presented by Senge is *shared vision*. Senge differ between *vision* and *shared vision* in that a vision is often viewed as the imposition of one person’s vision on the organization as a whole, whereas a true shared vision attracts the commitment of all participants (Senge, 1997, p. 51). According to Senge (ibid.), “a successful corporate vision will include a coherent picture of the future of the organization, a clear reason for wanting to reach this goal, and the core values needed to achieve it.”. Roald argues that the development of a common principle in an organization often implies a higher tolerance for new practices, and a higher tolerance for possible faults that might occur in the testing of said practices (Roald, 2012, p. 130).

Lastly, the fifth discipline, or ingredient as Senge calls it, *team learning* is “the process of aligning and developing the capacity of a team to create the results its members desire,

building on their shared vision.” (Senge, 1997, p. 51). Though it is presented as a collective discipline, Senge calls for individual understanding and knowledge in order to undergo fruitful dialogue and discussion (Roald, 2012, p. 130). Citing Senge, Roald (ibid.) differs between dialogue and discussion in that dialogue can be seen as listening and investigative reflection with the goal of finding new insight, whereas discussions could be seen as a more political process that focuses on power, wrongs or rights, etc. Through dialogue, the group could explore complex issues from many points of view, and Senge argues that in this context, “conflict within the team becomes a source of dynamic energy rather than a stumbling block.” (Senge, 1997, p. 51). Dalin (2005, p. 52) connects this to the Greek term *dia-logos*, which signifies a free exchange of opinions in group work, enabling the group to discover new insight that no individual acting alone could achieve .

### **4.3 Dalin’s five dimensions**

As part of seeing schools as more complex systems, where different actions within the schools affect each other whilst at the same time interact with society, Dalin presented his theory about the five dimensions in school as an organization (Roald, 2012, p. 118). In this theory, it is emphasized that it could be more productive to study what *actually* happens in the schools (i.e. from an empirical point of view) rather than discussing the development in schools from normative viewpoints (Roald, 2012, p. 119). Roald stresses that the development of the school as an organization is often seen from a perspective of change, but that an equally important perspective is that of stability. In order to keep in touch with traditions of high academic, social, esthetical and ethical standards, the continuous development of content, organizational- and structural methods in schools is paramount (ibid.) With this as a starting point, Dalin (cited in Roald, 2012, p. 119) presented five dimensions that are mutually dependent on each other; values, structures, relations, strategies and surroundings. In other words, no single dimension stands above the others, and changes in one dimension could change the others. Additionally, loose connections within the organization could cause the school to shield itself from challenges or incidents that occurs within certain parts of the organization.

*Values* are presented as the basic understandings, shown by the ideological and philosophical foundation of the school. This dimension surrounds both expressed goals

and unformal norms, and Dalin (cited in Roald, 2012, p. 120) notes that due to the wide specter of values that are often found in schools, conflicts may occur between formal goals and the values that are realistically represented. In addition, the recognition and acceptance of various attitudes and norms within the schools is key when it comes to giving professional freedom for individuals or groups within the school (ibid.).

*Structures* are the formal frames for organizing employees, pupils, time, material- and economical resources in the school. If good structures are present, routines and traditions are appreciated and continued, but Dalin (cited in Roald, 2012, p. 121) emphasizes that the structures must also be flexible enough for change and renewal to occur (ibid.).

Seen as informal relationships that have a deep impact on problem solving in schools, *relations* concern interpersonal relationships, such as cooperation, commitment, conflict, power, motivation, trust, and support (Roald, 2012, p. 121). Dalin (cited in Roald, 2012, p. 121) argues that individual and organizational learning happens through interaction, making the quality of a school dependent on the quality of the interpersonal relationships within the school. Although often visible through conflicts in interpersonal interactions, Dalin (ibid.) argues that such conflicts could stem from unfit structures or circumstantial issues. Constructive communication is therefore put forth as an important part of organizational development (Roald, 2012, p. 121).

The fourth dimension is *strategies*, which Dalin defines as the methods and tools used in the development in schools (Roald, 2012, p. 121). Problem solving, decision making, delegation etc. are aspects of strategies that leaders in schools could find challenging but is also something that is necessary to impose balance and dynamics between goals, structures, relations and surroundings (ibid.).

Lastly, the *surroundings* of a school, play a large part of the internal life in the school. Roald states that the school has a formal relationship with municipalities, local politics, and departments, but that the interaction between the school and other public and local institutions that work with the upbringing and education of children could be equally important. Dalin (ibid.) emphasizes that schools often have opportunities to interact with their surroundings, giving them opportunities that are often missed by school leaders and teachers (ibid.)

## 5 Methods

This chapter outlines research method selection and the rationale behind our choices. We present a general overview of common research methods, grounded in what we wanted to explore through our research. Through qualitative and quantitative research methods, we wish to explore how the results from the national test in English is used as part of the academic development in schools in northern Norway. To some extent, our research builds upon the research of Seland, Vibe & Hovdhaugen (2013), who sought to acquire comprehensive knowledge about the national tests, in order to utilize the knowledge across multiple levels of the educational system in Norway.

With us having comprehensive exposure with the school system and the national tests, both as pupils and through working in various schools, our personal experiences regarding the tests both as pupils and teachers will be part of the research. Consequently, we are faced with the need for evaluating our role as independent researchers, and how we perceive and interpret the data that is collected. This aspect will be further explored throughout this chapter.

### 5.1 Interview

Interviews, as articulated by Gleiss and Sæther (2021, p. 78) offer access to individuals' perspectives and serve as a well-suited method for delving into the depths of human thoughts, experiences, and perceptions. Gleiss & Sæther (2021, p. 78) emphasizes the need to distinguish between research interviews and everyday conversations, emphasizing the detailed preparation and systematic approach required in the three phases of the interview process: preparation, execution, and post-processing.

The preparatory phase, according to Gleiss and Sæther (2021, p. 78), involves two key considerations. Firstly, researchers must decide on the type of interview they wish to conduct. Secondly, formulating interview questions in advance is crucial to guide the interview effectively. Within the methodological framework of interviews, distinctions are made between individual interviews and those with multiple informants simultaneously. Further categorizations include structured, unstructured, and semi-structured interviews. Structured interviews involve pre-formulated questions asked in a consistent order, facilitating comparison of responses across informants. This method

is applicable in both qualitative and quantitative research, with the latter incorporating answer options (Gleiss & Sæther, 2021, p. 79).

## **5.2 Interview guide**

Utilizing an interview guide is common and advantageous in interviews as a method, providing an overview of the questions to be posed to informants (Gleiss & Sæther, 2021, p. 82). In contrast to structured interviews where questions are read word for word, semi-structured interviews often utilize the guide as a memory aid for addressing specific themes and questions. The formulation of open and closed questions is a crucial factor in obtaining comprehensive responses from informants, as qualitative research aims to gain insight into their experiences, observations, and knowledge on the relevant topic (Gleiss & Sæther, 2021, p. 82-85).

Executing interviews presents challenges, as not everything can be detailed planned in advance (Gleiss & Sæther, 2021, p. 86). The dynamic relationship between the researcher and informant influences the knowledge developed, necessitating reflection on how this relationship is established. Furthermore, the dynamics of pauses, silence, and follow-up questions can significantly impact the process of knowledge development in interviews.

## **5.3 Mixed Methods Design**

According to Creswell and Guetterman (2021, p. 595), a mixed methods research design is “a procedure for collecting, analyzing and ‘mixing’ both quantitative and qualitative methods in a single study or a series of studies to understand a research problem”. With a basic assumption that combination of these methods could provide a better understanding of the research question than either method by itself, Miles, Huberman and Saldaña (2014, p. 44, cited in Creswell and Guetterman, 2021, p. 595) state that the combination of quantitative and qualitative data provides the researcher with “a very powerful mix”. Creswell and Guetterman (2021) discuss three basic and three complex designs, where we have chosen to focus on the three basic designs due to their relevance for our research.



### **5.3.1 The Convergent Design**

Also referred to as parallel or concurrent mixed methods design, the convergent design serves the purpose of simultaneously collecting, merging and comparing the results from both quantitative and qualitative data (Creswell & Guetterman, 2021, p. 601). Creswell and Guetterman (ibid.) claim that the convergent design is based on the core assumption that qualitative and quantitative data provides different results, and that these results can be used to check one another. When using the convergent design method, the two sets of data are collected and analyzed separately, before a comparison of the two data sets is done. This comparison is subsequently used to interpret whether the results support or diverge (Creswell & Guetterman, 2021, p. 601). Although the combination of qualitative and quantitative research methods should be seen as a strength, the researcher still has to determine how he or she should merge the two data sets, and how to assess results that diverge (Creswell & Guetterman, 2021, p. 603).

### **5.3.2 The Explanatory Sequential Design**

Referred to by Creswell and Guetterman (2021, p. 603) as the most popular form of mixed methods approach in educational research, the explanatory sequential design consists of two phases where data sets are collected one after the other. In this design, the quantitative data is collected first, with a following qualitative data set being collected to help explain or elaborate on the quantitative results (ibid.). Creswell and Guetterman (2021, pp. 603-604) explains that the further elaboration on the quantitative data is necessary regardless of whether or not the results from the quantitative data contains any unexpected data, seeing how both expected and unexpected data needs to be further investigated. A difficulty of using the explanatory sequential design is that the researcher needs to determine which parts of the quantitative data to investigate further, and which questions to ask to get the desired information.

### **5.3.3 The Exploratory Sequential Design**

As opposed to the explanatory sequential design, the exploratory sequential design entails collecting qualitative data first, then using quantitative data to explain the findings. By collecting qualitative data first, the researcher can explore a given

phenomenon before determining which questions to ask, which variables to measure and which individuals could serve as useful informants in a following quantitative survey (Creswell & Guetterman, 2021, pp. 604-605). Although this approach allows the researcher to identify measures that are grounded in the data obtained from study participants, Creswell and Guetterman (2021, p. 605) points out that both the time and the extensive data that is required for this process should be seen as a disadvantage in the exploratory sequential design method.

#### **5.3.4 Qualitative interviews**

Asking open-ended questions and recording the answers is regarded as one of the more popular methods in qualitative research. After obtaining the data, the researchers transcribe and analyses the data using a computer program. When asking open ended questions, the respondents are put in a situation where they can share their experiences without constrictions made by past research findings or assumptions made by the researcher. Furthermore, open-ended responses allow the respondents to create the options for responding, in contrast to being forced into response possibilities (Creswell & Guetterman, 2021, pp. 251-252). Although time consuming, one-on-one interviews are popular in educational research. This form of data collection process consists of the researcher asking questions from one respondent at a time and recording their responses. Creswell & Guetterman state that one-on-one interviews are ideal for interviewing participants who are not hesitant to speak and who can share ideas comfortably (ibid.). Using web-based programs such as Microsoft Teams or Zoom addresses the problem of geographic distance, allowing the researcher to interview a participant that is based anywhere in the world at virtually no cost. Recording the interview can be done through the software that is used, or by using digital recorders. Although this form of interviewing opens up the possibilities of interviewing virtually anyone, Creswell & Guetterman (2021, p. 254) points to poor internet connection as a drawback, calling it a frustrating experience that ultimately could make the data unreliable since bits of the conversation could be lost.

#### **5.3.5 Mailed- and online surveys or questionnaires**

A questionnaire is a form of survey that allow participants to complete and return their answers to the researcher. Using surveys in the form of online questionnaires entails

that extensive data can be collected quickly, with the additional advantage of knowing that the form of collecting data is well known, and extensively used in many fields. In addition to this, employing online surveys may allow for effective and economical surveys of large populations. Mailed surveys is a form of data collection where the survey is mailed to members of the sample and are considered a convenient way to reach a geographically dispersed sample of a population. As with online surveys, mailed surveys facilitate quick data collection. However, as Sills & Song points out, issues such as low response rate and the interchangeability of e-mail addresses amongst respondents could cause difficulties when it comes to drawing inferences to a general population (2002, cited in Creswell & Guetterman, 2021, p. 436-437). In addition to this, the lack of personal investment may cause the individuals that receive mailed surveys not to return any form of response to the survey. This, with the added disadvantage of not having any means to probe for additional responses entails the researchers to weigh the advantages and disadvantages with using both online- and mailed surveys.

In quantitative research, surveys contribute to data collection from a larger sample compared to qualitative research (Gleiss & Sæther, 2021, p. 143). Surveys offer the advantage of generalization, allowing insights into a broader population. Designing a survey is challenging and requires thorough preparation. Initial steps involve operationalizing theoretical concepts to concrete questions and determining the formulation and sequence of questions. Operationalization involves translating theoretical concepts into specific questions with corresponding answer options. Attention to fundamental research ethical principles, such as informed consent, confidentiality, and avoiding negative consequences for informants, is paramount in survey research (Gleiss & Sæther, 2021, p. 157). The article details the authors' approach in developing informational sheets and consent forms, ensuring adherence to ethical principles in the research process. Principles such as informed consent, confidentiality, and the avoidance of negative consequences for participants are central to our study. In our research, we have developed information sheets and consent forms that each participant must complete before further engagement is undertaken. Participants are explicitly informed of their right to withdraw at any point, and details regarding the secure storage and handling of data are through 'nettskjema', ensuring deletion upon submission of the master's thesis.

### 5.3.6 Transcribing

Transcription is the most common method for preparing recorded interviews for analysis (Gleiss & Sæther, 2021, p. 97). This involves converting spoken statements in the interview into written text by listening to the recording and then writing down what is said, essentially showing the process of transcription. Although it is possible to analyze recordings without transcriptions, many researchers find that the analysis process is more straightforward with a written text to work with. Transcription is time-consuming, with a 30-minute interview potentially taking one to two hours to transcribe, depending on prior experience and the quality of the audio recording. If the quality of the audio recording is poor, it is more challenging to analyze what is said and to understand the context.

The advantage of transcription is that it allows a more in-depth engagement with the data compared to relying solely on the interview itself, providing direct findings and insights that may emerge from this process. In work involving large tasks, transcription can be considered the first step in a more systematic analysis process (Gleiss & Sæther, 2021, p. 97). Furthermore, Gleiss & Sæther (2021, p. 97) point out that master's pupils are often uncertain about the transcription process when dealing with interview material for the first time due to their lack of experience. Additionally, a range of decisions must be made during the transcription process (Gleiss & Sæther, 2021, p. 98).

One of these decisions involves considerations for anonymization and how the informant is represented in the final written work, ensuring that the meanings attributed to the informant are consistent with the transcription process. Another critical aspect involves discussing the choices made and their justification in relation to the research questions. Decisions must also be made about whether to place periods in the transcriptions, as they are absent from the verbal recording but can indicate where a natural break occurs (Gleiss & Sæther, 2021, p. 98).

## 5.4 Methodological evaluation

### 5.4.1 Validity

Cohen, Manion & Morrison (2018, p. 245) states that validity is key in effective research, calling invalid research «worthless». Whilst some versions of validity regard it as a «demonstration that a particular instrument in fact measures what it intends, purports or claims to measure» (Winter, 2000, cited in Cohen, Manion & Morrison, 2018, p. 245), other definitions state that «validity is the extent to which interpretations of data are warranted by the theories and evidence used (Ary *et al.*, 2002 cited in Cohen, Manion & Morrison, 2018, p. 245). To ensure the validity of the data collected, the researchers must be aware of socially constructed knowledge, and that the meaning and interpretation of the results of the data collection and instrumentation are sound (p. 246).

In quantitative research, features such as replicability, consistency, predictability, and controllability should be kept in mind, and are features that the researcher should be faithful to. In qualitative research, principles such as inductive analysis, respondent validation, member checking and catching agency, meaning and intention are essential, and according to Cohen, Manion & Morrison, triangulation could enhance the validity of the data (2018, pp. 247-249). Furthermore, Fielding & Fielding (1986, cited in Cohen, Manion & Morrison, 2018, p. 247) notes that “the data that is selected must be representative of the sample, the whole data set and the field”.

Qualitative and quantitative research methods can both address internal and external validity. To ensure internal validity in quantitative and qualitative research, the researcher must make sure that the findings accurately describe the phenomenon that is being researched (Cohen, Manion & Morrison, 2018, p. 252). Furthermore, the researcher must be aware of various threats to internal validity, such as instrumentalization in quantitative research, where unreliable tests could induce errors in the research (*ibid.*). Another threat to the internal validity in quantitative research could be that the researchers become more experienced throughout the research period, which could cause them to change their scoring procedures. To correct this potential

problem, the use of standardized testing procedures should be used throughout the research (Creswell & Guetterman, 2021, p. 348).

In qualitative research, internal validity can be addressed by using “low-interference descriptors, multiple researchers, participant researchers, peer examination of data and mechanical means to record, store and retrieve data” (LeCompte & Preissle, 1993, cited in Cohen, Manion & Morrison, 2018, p. 253). By utilizing member checking to correct factual errors or misunderstandings between the researchers and the interviewee, as well as giving the interviewees the opportunity to add further information, internal validity in qualitative research could be strengthened (ibid.).

External validity in quantitative research concerns the question of generalizing from sample to population, bringing attention to such factors as generalizing from a given situation to another without taking contextual and casual differences into account, and keeping a constant focus on range of outcomes (Cohen, Manion & Morrison, 2018, p. 254). Furthermore, Cohen Manion & Morrison point to lack of representativeness as a factor that could jeopardize the external validity of the research, arguing that participants might represent an available population without representing the population to which the researcher seeks to generalize (Cohen, Manion & Morrison, 2018, p. 255).

By studying the typical for its applicability to other situations, as well as performing multi-site studies, qualitative research could be generalized (Scofield, 1996, cited in Cohen, Manion & Morrison, 2018, p. 255). However, Bogdan & Biklen argue that “in qualitative research, we are more interested not with the issue of whether the findings are generalizable in the widest sense but with the question of the settings, people and situations to which they might be generalizable” (1992, cited in Cohen, Manion & Morrison, 2018, p. 255).

#### **5.4.2 Reliability**

Used as an umbrella term for dependability, consistency and replicability over time, reliability is relevant to both quantitative and qualitative research methods. If, for instance, participants in a survey are fatigued, nervous or if they misinterpret questions, the data that is collected could become unreliable (Creswell & Guetterman, 2021, p.

188). Cohen, Manion & Morrison (2018, p. 268) argue that research must demonstrate that similar results would be found should the research be carried out on a similar group of respondents. In the search for trends, patterns, predictability and control, Miles and Huberman list three types of reliability in quantitative and qualitative research: stability, equivalence and internal consistency (1994, cited in Cohen, Manion & Morrison, 2018, p. 268). Measuring consistency over time, or consistency over a similar sample implies finding similar results in research conducted over time, or in participants that are closely matched on significant characteristics and finding the results or responses similar to each other (Cohen, Manion & Morrison, 2018, p. 268). Reliability as equivalence consists of devising an equivalent form of test or instrument to demonstrate the reliability of the first test or instrument that has been used. Additionally, reliability as equivalence could be achieved when more than one researcher takes part in the research, making human errors less likely (Cohen, Manion & Morrison, 2018, p. 269). To check for internal consistency, a respondent can be asked the same questions later in the instrument. If the results are the same (i.e. that the respondent answers the questions in a similar manner), they can be viewed as reliable (Creswell & Guetterman, 2021, p. 190).

According to Bogdan & Biklen (1992, cited in Cohen, Manion & Morrison, 2018, p. 270), reliability can be regarded as «a fit between what researchers record as data and what actually occurs in the natural setting that is being researched». Citing Brock-Utne (1996), dependability is brought forth as important by Cohen, Manion & Morrison (2018, p. 271), where member checking, debriefing by peers, triangulation etc. are used to ensure that the results are consistent with the data that is collected. Kleven (1995, cited in Cohen, Manion & Morrison, 2018, p. 271) claims that three questions could be asked to address reliability in qualitative research, each addressing different versions of reliability; The 'stability' version of reliability asks whether or not the results would have changed if the research had been conducted at a different time, the 'parallel form' asks whether or not the results would have been different if other observations and interpretations had been made at the same time, and the 'inter-rater' version of reliability asks if the same observations and interpretations would have been presented by another observer working within the same theoretical framework.

## 5.5 Gathering, analyzing, and interpreting data

### 5.5.1 Sampling

#### 5.5.1.1 Availability selection:

An availability selection involves selecting participants based on their accessibility or ease of access. This selection method does not account for randomness or probability, and participants are often chosen because they are more readily reachable, or data can be collected from them easily. This may lead to a bias in the sample, as it does not represent the entire population in a fair manner.

#### 5.5.1.2 Random sampling:

Random sampling involves giving each individual in the population an equal chance of being included in the sample. This means that participants are selected in a random manner, providing greater assurance that the sample is representative of the entire population. This reduces the likelihood of bias and yields more generalized results.

#### 5.5.1.3 Probability sampling:

Probability sampling is a form of sampling where each element in the population has a known probability of being selected. It can be based on various methods, such as simple random sampling, systematic sampling, or stratified sampling. Probability sampling provides a structured approach that can yield reliable results and is often preferred whenever feasible.

#### 5.5.1.4 Strategical selection:

When collecting data from the national test in English for our master's thesis, it is crucial to assess which sampling method is most appropriate and provides results that are representative of the target population. If feasible, consider employing a form of probability sampling to enhance the validity and generalizability of your findings.

### 5.5.2 Quantitative data

There are several steps involved in the process of analyzing quantitative data. **First**, the data is prepared for analysis. "This involves determining how to assign numeric scores



to the data, assessing the types of scores to use, selecting a statistical program, inputting the data into said program and then cleaning up the database” (Creswell & Guetterman, 2021, p. 205). Here, the researcher would for instance assign the options ‘strongly agree, agree, undecided, disagree, strongly disagree’ numeric values or scores, (i.e. 1,2,3,4,5) (ibid.) The numeric scores should be consistent in continuous scales, whereas categorical scale scores should (or rather, could) be scored ‘to make sense’ (Creswell & Guetterman, 2021, p. 205). In determining the types of scores to analyze, three types of scores are used; *single-item scores* (individual score assigned to each question for each participant in the study), *summed scores* (scores of an individual added over several questions that measure the same variable) and *difference scores* (scores in a quantitative study that represent a difference or change for each individual) (Creswell & Guetterman, 2021, pp. 207-208). When choosing which statistical program to use, academical researchers generally use the programs that are available at their institution. For our part, *IBM SPSS Statistics* will be used.

The second step is the start of the data analysis. In this step one would typically conduct a descriptive analysis, before conducting a more sophisticated inferential analysis to test hypotheses and examine confidence intervals and effect sizes. In the third step, one reports the results that are found, using figures, tables and a discussion of the key results. The fourth and last step consists of interpreting the results from the data analysis.

### **5.5.3 Qualitative data**

In the process of analyzing the qualitative data that will be collected, we find it appropriate to use two analysis methods. The first qualitative analysis is based on phenomenological design, which deals with studying other individuals’ subjective perspective to understand a phenomenon. The second qualitative method, thematic analysis includes discussing the major themes that arise from analyzing the qualitative data, using extensive quotes and rich details to support the themes (Creswell & Guetterman, 2021, p. 315). Cohen, Manion & Morrison (2018, pp. 647-648) describes this as a painstaking process where the data is taken apart, put into main outlines of the phenomena that is being investigated before groups of data is put together in summaries

of what has been found, making them a coherent whole with the intention to move from description to explanation to theory generation.

## 6 Research ethical considerations

To ensure scientific practice, we will consider both our collection and storage of data. Within this, both anonymization and the rights of informants before, during and after data collection will be taken care of. A consent form will be made, making it explicit that informants have the opportunity to withdraw at any moment of the process. All collected data will be anonymized to ensure that the informants cannot be identified based on the content of the thesis. We will use Nettskjema (online survey tool) as the storage location for recordings in accordance with regulations.

The national committees for research Ethics (DNFK) developed general research ethical guidelines in 2014 (De nasjonale forskningsetiske komiteene, 2019). The purpose is not to replace the discipline-specific guidelines but rather to serve as an entry point to research ethical principles and considerations. Withing these principles, DNFK (ibid.) emphasizes four key aspects:

1. Respect: Individuals participating in research, such as informants, should be treated with respect.
2. Good consequences: As a researcher it is essential to ensure that one's activities have positive consequences, and any potential negative or unintended consequences are deemed acceptable.
3. Justice: Research projects should always be designed and conducted fairly.
4. Integrity: Researchers should be open and honest with colleagues and the public, simultaneously acting responsibly and adhering to norms.

Furthermore, DNFK (ibid.) has outlined 14 specific considerations they emphasize, including the pursuit of truth, quality and voluntary informed consent, which will be crucial for us to adhere to. The pursuit of truth involves seeking new knowledge through critical and systematic examination where honesty, transparency, and systematicity are fundamental. Quality pertains to ensuring that the research reflects a high level of academic excellence, with the researcher possessing the necessary competencies.

Voluntary informed consent emphasizes that involved parties are voluntary participants in the research and have given their informed consent to be part of it.

## **6.1 Our role – reflexivity:**

In the process of conducting a master's thesis, ensuring high research quality, and adhering to research ethics are paramount (Gleiss & Sæther, 2021, p. 49). The research process involves numerous conscious and unconscious choices concerning the anticipated consequences of various actions. Consequently, maintaining a consistently critical and inquisitive stance toward one's own research work is crucial, a principle encapsulated by the term "reflexivity" (Gleiss & Sæther, 2021, p. 49). Reflexivity enhances the quality of research by rendering the decisions more justified and easier for the reader to evaluate.

Various factors can influence the researcher's positionality, which is the standpoint from which they perceive the world. Ethnicity, gender, social milieu, and relationships can all impact the research process, making it essential to reflect on one's own positionality by considering aspects of identity and demeanor and how they may affect research processes such as problem formulation, analysis, etc. (Gleiss & Sæther, 2021, p. 49). The ideal of research is objectivity, which inherently conflicts with positionality. Even today, debates persist among researchers regarding the defining characteristics of good research and the criteria useful for assessing the quality of work (Gleiss & Sæther, 2021, p. 51).

Furthermore, ongoing discussions among researchers revolve around the type of relationship with participants that yields the best research outcomes (Gleiss & Sæther, 2021, p. 51). Two myths are distinguished: the insider myth and the outsider myth. The insider myth suggests that researchers with affiliations to the field or social environment are best equipped to understand it. Conversely, the outsider myth points out that individuals with sufficient distance can perceive it more clearly. Relationships can indeed influence the research process, as exemplified by teachers or students conducting data collection in schools where they currently or previously worked. While this approach is possible to do, it necessitates reflection on one's own positionality. A challenge arises in that participants may find it difficult to provide free consent due to

the presence of a personal relationship (Gleiss & Sæther, 2021, p. 51). An integral aspect of a reflexive researcher ideal is contemplating what the research can contribute to the field in which it operates (Gleiss & Sæther, 2021, p. 52).

Regarding whether the insider or outsider position is more appropriate, Gleiss & Sæther (2021, p. 88) highlight that there is not a clear enough distinction between the two. They further state that, in practice, most researchers occupy both an insider and outsider role in the field they are researching. This can be exemplified, for instance, by a student teacher interviewing teachers where the student might understand the daily life of a teacher through practice or work but lack knowledge of each school's culture, working methods, and learning environment (Gleiss & Sæther, 2021, p. 89). Furthermore, both positions carry advantages and disadvantages regarding knowledge development in research. The insider position holds relevant knowledge and experience, which can make it easier to ask questions and gather informants, but being on the inside might cause one to overlook certain aspects. The outsider position may offer an external perspective that highlights things that group members within the insider-position might miss, but can also lead to the researcher asking questions that the informants find irrelevant (Gleiss & Sæther, 2021, p. 89).

When preparing for our research interviews, we made our best efforts to take the presented factors mentioned by Gleiss & Sæther (2021) into account. Although one of us did work with some of the informants or otherwise knew them at a personal level, the other one had no affiliation with said informants whatsoever. We consider this to be a factor that strengthened our research, since one of us could be seen as presenting the insider myth, with him having knowledge of the social milieu (Gleiss & Sæther, 2021), and the other presenting the outsider myth, with sufficient distance to perceive the information given to us by the interviewees more clearly (ibid.). Subsequently, the reflection on both of our positionalities became apparent when discussing and analyzing the data that had been collected, with two views creating the grounds for added depth to our analysis. When interviewing participants from schools that either of us had any affiliation with, we took care to be as objective as possible both during the interviews, and during the interpretation and analyzing of the data. Although we mostly did find ourselves stressing the same factors when interpreting and analyzing the data, we also

experienced that through discussing areas of the research where our individual interpretation differed, we were able to look at the data in an objective fashion. An important aspect of our research was for us to contribute to the field in the sense that there is little research done regarding the national tests in northern Norway, with the added factor of us, through the research, finding that a systematic change in the school as an organization could utilize the full potential of the national tests as a mapping- and development tool.

## **6.2 Strengths and weaknesses of the project**

A potential weakness in the project is our susceptibility to bias such as the previous research we have read, and that we have both worked in educational systems previously exposing us for opinions and impressions. In terms of preconception. We naturally hold opinions and expectations about what the informants may express, which could lead to leading questions and a subjective basis for both qualitative and quantitative data collections. However, we will actively acknowledge this concern throughout this thesis, attempting to ensure the correct presentation of the data.

A general limitation of the project lies in weaknesses within both qualitative and quantitative methods, encompassing issues related to sample selection, data analysis, and interpretation. It is crucial to assess whether the informants' statement align with actual practices (reliability) and pose insightful follow-up questions to safeguard the validity of the data.

A potential strength is the diversity within our selection of schools for the research. We have chosen a variety of both urban and rural schools, aiming to provide a robust and comprehensive foundation for our observations. Another potential strength lies in the mapping of results for further research, particularly in northern Norway where there is a limited body of research in this field.

In the process of developing adequate research, it is essential to be aware of one's own strengths and weaknesses. We are two researchers with little to no prior experience in this field of work, and this is our first time undertaking a project of such scope/extent. Additionally, our lack of established routines and experience may have impacted various

factors, including quality, validity, and decision-making. A key aspect of our weakness lies in the potential sources of error. What are they?

However, our strengths are based on our exploration of previous research related to national tests. Through this work, we have encountered familiar issues and have confirmed previous theories and research findings. Our reliance on established research as a basis for comparison has enabled us to identify consistent themes and connections from a broader perspective identifying what we mean are the real issues.

## 7 Results

In the following chapter, the results from the qualitative and quantitative study will be presented. The data presented in the following chapter has been reviewed by us, and besides coding the qualitative data, no interpretations have been made. Firstly, we will present the data from the quantitative survey, which are split into each of the five sub-categories from the survey. Secondly, the qualitative data is presented through our coding of the qualitative data, where we categorized the data into main codes, each with its own subcodes.

### 7.1 Quantitative results

A survey was made through nettskjema.no, containing 17 questions that was put into sub-categories. The survey was structured as a Likert scale with five values per question. (1-5, with 1 being very low and 5 being very high) and were mandatory to complete the survey. Traditionally, the number of points in a Likert scale can be as few as three or four, however, we opted to increase the number of scale points to five in order to make it closer to continuous scales and normality, in accordance to Wu and Leung (2017, p. 527). The decision was made not to include the option to elaborate on any of the questions. Two identical versions of the survey were sent out: one for teachers and one for school leaders. This was done so that we could compare the answers from teachers and school leaders- due to previous research indicating a difference in how the teachers and school leaders viewed the national tests. Two of the questions in the survey were repeated twice, this was done so that we could compare the answers and calculate the average of the two. The survey was posted in a closed Facebook group for teachers and was also sent via email to all the schools in northern-Norway. In total, 48 teachers and 18 school leaders filled out the survey. Due to the survey being anonymous, there is a possibility that teachers that do not teach English have filled out the survey- despite us stressing that the survey was made for teachers that has English as their main subject. After the survey was closed, all data was put into SPSS Statistics for further investigation. Below are the tables with the results from the surveys sent to the teachers and school leaders, further explained in this chapter. See the appendix for graphs showing the results from the individual questions.



## 7.2 Table 1, quantitative results from teachers

<i>Variable</i>	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
<i>Useful development tool 1</i>	48	1	5	2.83	0.996
<i>Mapping the academic level</i>	48	1	5	3.17	0.907
<i>Explaining the purpose of the tests</i>	48	2	5	3.79	0.874
<i>Guidance materials from UDIR</i>	48	2	5	3.73	0.962
<i>Time for preparation</i>	48	1	5	3.19	0.982
<i>Single tasks, preparation</i>	48	1	5	3.67	1.243
<i>Task-sets, preparation</i>	48	1	5	2.94	1.245
<i>Division of responsibilities</i>	48	1	5	3.56	1.147
<i>Helping the pupils during the test</i>	48	1	5	1.77	1.016
<i>Absence on the day of testing</i>	48	3	5	4.60	0.574
<i>Understanding the results 1</i>	48	3	5	4.00	0.619
<i>Equipped to interpret</i>	48	2	5	3.52	0.922
<i>Using the results for facilitation</i>	48	1	5	3.17	0.930
<i>Knowledge of the guide "nasjonale prøver"</i>	48	1	5	3.48	1.052
<i>Understanding the results 2</i>	48	1	4	2.35	0.812
<i>Using the results to develop the subject</i>	48	1	5	2.79	1.071
<i>Useful development tool 2</i>	48	1	5	2.94	0.998

### 7.3 Table 2, quantitative results from school leaders

Variable	<i>N</i>	Min	Max	<i>M</i>	<i>SD</i>
Useful development tool 1	16	3	5	3.87	0.719
Mapping the academic level	16	3	5	3.75	0.577
Explaining the purpose of the tests	16	1	5	3.44	1.031
Guidance materials from UDIR	16	2	5	4.06	0.854
Time for preparation	16	2	5	3.56	0.727
Single tasks, preparation	16	1	5	3.63	1.088
Task-sets, preparation	16	1	5	2.94	1.124
Division of responsibilities	16	3	5	4.69	0.602
Helping the pupils during the test	16	1	3	1.88	0.806
Absence on the day of testing	16	3	5	4.44	0.629
Understanding the results 1	16	3	5	4.25	0.775
Equipped to interpret	16	2	5	4.00	0.894
Using the results for facilitation	16	3	5	4.00	0.632
Knowledge of the guide “nasjonale prøver”	16	3	5	4.38	0.619
Understanding the results 2	16	1	3	1.87	0.619
Using the results to develop the subject	16	2	5	3.63	0.719
Useful development tool 2	16	2	5	3.69	0.704

The first sub-category of the survey contained questions regarding the teachers’ and school leaders’ attitudes towards the national tests.

When asked to which degree the national test in English was regarded as a useful development tool in the English subject, 18 of the teachers answered to a low- or very low degree, whereas none of the school leaders chose these alternatives. Of the 48 teachers that responded, 18 answered neither-nor, and 12 answered high- or very high. 11 of the school leaders answered high- or very high, with 5 answering neither-nor. The final question in the survey was a repetition of the first question, when asked the same

question for the second time, 17 of the teachers and 1 school leaders answered that the national tests were useful to a little- or very little degree, whilst 15 teachers and 11 school leaders answered that the tests were useful to a high- or very high degree.

Most of the teachers viewed the national test in English as a good mapping tool, with 17 answering high- or very high. The same results came from the school leaders, with 11 answering high- or very high. 11 teachers and 0 school leaders answered low- or very low, whilst 20 teachers and 5 school leaders answered neither-nor. When asked to which degree the teachers and school leaders explained the purpose of the national tests to their pupils the vast majority answered high- or very high, with only 4 teachers and 2 school leaders choosing the lower alternatives.

The second sub-category delved into the preparation phase of the national tests.

When asked how familiar the informants were with the guidance materials provided by UDIR, most of the teachers (32) and school leaders (13) replied that they had good, or very good knowledge of the guidance materials. Although the results varied, a small majority of the teachers (20) did feel that there was plenty of time set aside for the preparations before the national tests. 12 of the teachers answered that little or very little time was set aside, whilst only 1 of the school leaders answered that there was little time set aside for the preparations.

33 of the teachers that answered the survey used single tasks often or very often when preparing their pupils for the national test in English, with only 10 stating that they used single tasks to a small- or very small degree. The majority of the school leaders (10) used single tasks often or very often as part of their preparations, whilst 2 answered that they used single tasks to a small- or very small degree. The informants were divided when asked if they used task-sets in the preparation phase, with 19 of the teachers answering little- or very little, and 18 answering often- or very often. 6 of the school leaders answered little- or very little, with 4 answering often- or very often. Most of the informants were familiar with the division of responsibilities regarding the national tests, with 33 teachers and 15 school leaders stating that they were familiar- or very familiar with this.

The third sub-category was aimed at how the teachers and school leaders carried out the tests. Firstly, most of the teachers (40) and school leaders (12) offered little- or no help to the pupils during the tests (as is recommended by UDIR). 3 of the teachers stated that they offered help to some- or a high degree. Second, both the teachers (46) and the school leaders (15) stated that pupils that were absent on the day of testing got the chance to complete the test on a later occasion.

The fourth sub-category was focused on how the informants and their respective schools worked when analyzing, interpreting, and utilizing the results from the national tests.

39 of the teachers answered that they felt well- or very well equipped to understand the results from the national test in English and 13 of the school leaders answered likewise. None of the informants answered that they were ill- or very ill equipped to do so. 9 of the teachers answered that they to a small degree felt equipped to interpret the results. However, most of the teachers felt equipped- or very equipped to do so. 12 of the school leaders answered that they felt equipped- or very equipped to interpret the results, with only 1 stating that he or she was less equipped to do so. This was one of the questions that were repeated in the survey. When asked the same question a second time towards the end of the survey, 31 teachers and 14 school leaders answered that they understood the results from the national test in English good- or very good. None of the school leaders answered that they understood the results to a little- or very little degree, whilst 5 teachers answered that they understood the results to a little degree.

When asked to which degree the informants were able to use the results from the national test in English for facilitation, 19 teachers answered that they were able to do so to a high- or very high degree, whereas 11 teachers answered that they were to do so to a small- or very small degree. 13 of the school leaders answered that could use the results for facilitation to a high- or very high degree. The majority of the informants (30 teachers and 15 school leaders) answered that they had good- or very good knowledge of the guide “nasjonale prøver”.

The fifth and final sub-category focused on the informants understanding of- and thoughts about the national test in English as a development tool.

17 of the teachers answered that the results from the national test in English were used to develop the English subject to a little- or very little degree, whilst 12 teachers answered that they were used to a high- or very high degree. In contrast, 10 of the school leaders answered that the results were used to develop the English subject to a high- or very high degree, whereas only 1 of the school leaders answered that they were used to a little degree.

## 7.4 Qualitative results

### 7.4.1.1 Methodological freedom – preparation phase

*“[...] but mostly to assure that they are prepared for which tasks they are to solve and what it is expected that they should do, so that we can minimize the confusion and uncertainty that is often attached to the execution. By doing so, we can make sure that the results are a bit more valid” L2*

In the preparation phase of the national tests, all of the five teachers that were interviewed focused on preparing their pupils on the form of test rather than the actual tasks in the test. The teachers at hand stated that a goal was to make sure that the pupils were prepared for a long (60 minute) test, since many of them had little experience with tests of that length. In doing so, the teachers would find example tasks from previous tests or from the resource sites that UDIR provides. L5 chose to focus on strategy and problem solving when preparing her pupils, taking special care to making sure that her pupils understood the phrasing of the questions, and comparing it to sports by saying “you can’t run a 100-meter sprint if you’re *actually* running a marathon, and vice versa. You need to know what you’re doing”. This preparation was done by L5 with the entire class, based on the preparation materials provided by UDIR. L3 also based his preparations on the materials provided by UDIR, and did a thorough run-through of the example tasks from the preparation materials with the entire class. L2 mentioned that one or two teaching sessions were put aside for the preparation of the pupils, whilst L4 and L5 said that they themselves chose how much time should be spent on preparing the pupils for the tests. The principals that were interviewed stated that they received necessary information from the school owners, mostly regarding formal aspects such as important dates, reporting etc. R2 emphasized the need to prepare the teachers, since an unprepared teacher could find it difficult to prepare his or her pupils.

*"It wasn't something I missed, and I didn't feel the need for more support. All the information was out there. In some ways, that's fine. I can't remember having any negative experiences regarding the information that came from the school leaders." L4*

Whilst most of the teachers (four of them) didn't specifically mention any form of cooperation with colleagues in the preparation phase, L1 did state that his team (i.e. English teachers from the same grade) spoke together with regards to planning when the preparation for the test should take place etc. but did not specify how this was done or how much time was spent on this. The information given from the school leaders to the teachers at their respective schools was viewed by all the interviewees as sufficient, although L5 stated that she had nothing to do with the school leaders at her school. The other interviewees stated that the information that came from their school leaders was concentrated about practical features surrounding the national test in English such as time frames, reporting pupils that had a right to be exempted from the tests etc. Although L2 stated that the school leaders at his school had a "serious take" on the national tests, he also stated that the introduction to the guidelines and materials provided by UDIR was deficient and felt an expectation that it was his responsibility to familiarize himself with the test. L4 stated that he couldn't remember that the school leaders were involved in the preparation phase leading up to the national tests, but that he didn't feel a need for them to be involved in it either due to his knowledge on how to find and use the guidelines provided by UDIR. The principals that were interviewed both stated that the main part of the work with preparing for the national test in English was done by the individual schools, based on the information from the school owners and from UDIR. R1 noted that her school had developed their own guidelines to be used in the preparation phase of the national tests, and that she hadn't used the guidelines given by UDIR this year. R2 stated that his school used the guidelines given by UDIR every year, and that these were presented and looked at in one of the weekly meetings that the school has. About one hour and thirty minutes was set aside at the start of the autumn semester for each of the teacher teams at the school to look at the guidelines and to lay a plan for how the tests were to be carried out.

#### **7.4.1.2 Methodological freedom – development**

*“No extra time is put aside, but the general idea is that we are supposed to look at the results in our teams and work with them together” L1*

All the teachers that were interviewed pointed out that there was little or no cooperation between teachers when it came to using the national test in English as a development tool. Although L1 did state that the teacher teams did work together in cases where it became apparent that it was necessary to find common methods and pathways for the development of their teachings, he also noted that the heavy workflow and lack of time made it difficult to do so generally speaking. L4 noted that he was positive towards the freedom that he was given when it came to planning and using his own experience, knowledge of the pupils and his own way of thinking in his development process- implying that he didn't feel the need to cooperate with other teachers. L2 stated that the school leaders at his school encouraged the teachers to bring up the results from the national test in English in development conversations with their pupils, but emphasized that apart from that, the teachers did not have any instructions as to how they could or should use the test results as part of their academic development. L5 admitted that she hadn't used the test results at all this year, and that she had looked at the results only once to see if the test results indicated that any of her pupils needed special education. R1 stated that the routines made at her school were very clear, and that the results were analyzed and presented to all the teachers. She also pointed out that the addition of a 'listening part' in the national tests meant that they needed to adapt and develop their teachings to prepare the pupils for that specific part of the test in later years. R2 found it necessary for the teachers to be assisted in the work of analyzing the test results, and that they were aware of how the school had developed over several years.

### **7.5 The national test in English as a mapping-and development tool**

The following part of the results have been divided into sub-chapters based on our coding.

#### **7.5.1 Improvement potential**

### 7.5.1.1 Preparations

*"I think we've had a somewhat inadequate introduction to all the material related to national tests." L1*

Both L1 and L2 experienced technical issues that caused problems on the day of testing. Difficulties with network connections and logins caused what L1 described as 'chaos', whilst L2 drew lines between technical issues and the lack of conclusive results at his school. One of the principals, R2, also pointed to technical issues as a problem, and tended to wait a few days before taking the tests at his school to secure that the servers weren't overloaded. Out of the five interviewees, four of them experienced that the information given to them from the school leaders was insufficient, with L3 and L5 stating that they felt that they were left with the responsibility to prepare themselves, without sufficient instructions as to how they were supposed to do so.

### 7.5.1.2 Results

*"[...] so I did it my way, using them (the results from the national test in English) as a guide for what I should work on in the English subject. But I didn't get any instructions on how to use them, nor links to resources that discuss smart ways to utilize them" L3*

In the work with analyzing and interpreting the results from the national test in English, two factors were highlighted by several of the interviewees: time and prioritization. Firstly, all five of the teachers that were interviewed stated that the national test in English wasn't a tool that was prioritized by the school, with other tests that were carried out throughout the school year being of higher importance. L3 wished for clearer instructions from the school leaders, and L4 experienced that the work that he did with analyzing and interpreting the results was something that he did merely because he "had to". L1 emphasized that his school had no collegial meetings that were focused on systematic work with the test results, and L2 was under the impression that the school viewed the national tests as a status check more than a tool that could be of use for mapping or development. In addition, L5 stressed that the lack of a common system for working with the test results caused uncertainty amongst the teachers with regards to what could be a "right" or "wrong" way of analyzing and interpreting the results. L3 was under the impression that important questions such as "what could we do with these



results?” and “what are your ideas to make our teaching better?” remained unanswered from the school leaders.

Secondly, time was a factor that all five teachers saw as a flaw in the process of utilizing the tests results. L5 experienced that despite her wishes, little or no time was set aside for the teachers to analyze, interpret and utilize the test results, and drew parallels to this and how the school signaled their view of the national test in English by stating that if the schools give teachers little time, they send a signal that they view the national tests as something of little importance. Additionally, L2 and L4 specifically mentioned that they needed more time to work with analyzing, interpreting, and utilizing the test results, without receiving it from their respective schools. R1 pointed out that there are no common routines for the work that is done after the national tests are concluded and stated that there had been little talk of how the results could be used as part of the development of the English subject. Additionally, R2 saw the need for common grounds in the aftermath [?] of the national tests, pointing out that the respective principals were in charge of analyzing and interpreting the test results- with little or no common guidelines from the school owner or UDIR.

### **7.5.1.3 Development**

*“I wish there was more time to sit down and really work it. This would help shift the focus towards using it as a development tool, rather than just an assessment tool.” L2*

The interviews that were conducted revealed that all interviewees viewed the national tests as a mapping tool, and not necessarily a development tool. L2 stated that he could have used the test results for academic development, but that the lack of time and focus on the development aspect of the test made it difficult for him. L5 was clear in her views, stating that she considered the national tests as a mere mapping tool whose potential is not exploited. L1 believed that teachers want to work together when using the test results to develop their teachings, but that the lack of time made it difficult to do so. Both principals that were interviewed acknowledged the need for common routines in the use of the test results for academic development and emphasizing the advantage of having continuous education of the teachers at their respective schools.

### **7.5.2 Frequency**

*“I would have liked it if we had the national test in English annually in the intermediate grades, that is, in fifth, sixth, and seventh simultaneously. It would have been helpful for any school leader, as you can track the trend from fifth to sixth and seventh grades. I mean, we already do this with the student survey, and we find it to be incredibly valuable.” R2*

Looking at the frequency under results – L3 & R1 have no comments to make in regard to the topic. In contrast, L2, L4, L5 & R2 emphasize their wish of having national tests in English more often. They believe having it in 8<sup>th</sup> grade is not enough and wish to implement it in 9<sup>th</sup> grade additionally. L2, L5 and R2 see the benefit of having the tests in 9<sup>th</sup> grade as a significant tool to gaze at the pupil’s development. In addition, L5 emphasizes that having it in 8<sup>th</sup> grade only gives feedback over what they have learned and achieved from primary school, while teachers would benefit from receiving feedback of the pupil’s learning outcomes under their guidance. Furthermore, she expresses a bitter frustration in not being able to compare results over a scale of time. Additionally, L4 suggests that teachers and school leaders would be more engaged in the work by having it more often as it would give more reliable data and show the importance of engaging with it. R2 suggests that it would increase the potential for national test in English as a development tool and give a more detailed assessment of each pupil in forms of graphs and information. In addition, he expresses the long duration of wait for comparing data from 5<sup>th</sup> to 8<sup>th</sup> grade and that it will be done through two different schools. While the majority seek to expand the duration of national tests, L1 admits having an ambivalent relationship towards either increasing or decreasing the amount of times pupil’s will be tested in regard to national tests as there is already enough assessment tools in use.

### **7.5.3 Information flow**

*“But it would have been nice to have a bit clearer instruction [...] Maybe having a more standardized approach to how we are expected to organize the implementation might have been a good idea, I think.” L2*

The information flow is an integral part of a well-working system where the communication and directions are clear for the parties involved. Looking at the information flow from the top to the bottom, L1, L2 and L5 express dissatisfaction in this regard. With all three stating the lack of information flow, L1 further states that the

majority of information is improvised and therefore lacking. In addition, L2 desires stronger clarity in regard to the instructions given, stating they receive vague information resulting in a bigger gap between how each teacher utilize and prioritize the importance of national tests. While L3 and R1 have no comments in regard to the topic, L4 experiences the information flow as being all right, emphasizing that each teacher manages their own time regarding how they wish to utilize it. There is a change of scenery when addressing the answer from R2, he is happy with the information flow. He emphasizes the importance of a teacher being well-prepared, in order to prepare the pupils for the test in a suitable manner. In addition, he states that as a school they retrieve the guidelines from UDIR every year to ensure they are up to date with the information.

#### **7.5.4 Mapping tool**

*“It gives you a sense of how the student is doing. Especially how well they understand the material and to what extent they can handle this type of test, which is important.” L4*

A common view between the interviewees, both teachers and principals, was that the national tests provide a good insight regarding the pupils' academic level. L5 viewed the national tests as a good tool for easily seeing which of her pupils needed special education, emphasizing that some of her pupils lacked proper mapping. L4 stated that the national tests gave proficient insight to what academic level his pupils were at, but also emphasized that the test results gave a limited picture based on an established form of testing. R1 pointed out that the results from the national test in English were insufficient to see whether or not a pupil needed special education, but that they could be seen as an indicator for something that needed further investigation. L4 experienced the national tests as mapping tool that provided sufficient overview of the academic level of his pupils, the type of teachings that they are in need for and which of the pupils that has special educational needs. R2 stated that the school is dependent on mapping in all subjects but emphasized that the English subject is one of the harder subjects to map due to the lesser amount of mapping tools compared to other subjects in the school.

#### **7.5.5 Traditions**

*“I’ve been doing this for a long time, so we’ve kind of established our written routines for how we follow up on the results. So I don’t think we need much more.” R1*

It might be difficult to break a culture if it is accepted and used in a society for a number of years. Switching our focus to the results of traditions, it gives an insight to how each school works, and positives or negatives that may bring. L1 experiences their traditions in regard to the national test in English as just “something we have to do” and nothing more. L2 states that they use the results from the national test in English as an assessment tool and not development tool. Furthermore, L3 states that how to implement and use the results is up to each individual teacher. While L4 has nothing to add, L5 states they do not work collectively, and that they have not done so for many, many years. When interviewing the school leaders, there is a significant change of attitude towards the school’s traditions. Firstly, R1 clearly states that they use the national test in English as a development tool, emphasizing their established routines involving analyzes on both individual- and group-levels. However, she expresses that they do not use the guidelines from UDIR due to their system which she categorizes as “well-working”. Lastly, R2 states they follow a set template in their use of national tests. Furthermore, he visits every classroom to ensure everything is going in accordance, and to show the pupils the importance of the national test in English.

### **7.5.6 Development tool**

*“But for example, in my case, I set aside some time to see how I could use the results to improve my teaching.” L3*

Although most of the teachers that were interviewed didn’t necessarily view the national tests as a development tool, all of them agreed that the tests could show where the teachings needed to be improved. As part of differentiating tasks, L2 viewed the national tests as a good development tool, emphasizing that giving tasks of fitting difficulty to pupils that he didn’t know very well could be challenging. L3 used the test results as a way to enhance areas where the pupils struggled, whilst L1 used the test results in his development conversation with his pupils in order to help them in areas of the English subject where he or she struggled. Both principals that were interviewed saw the purpose of the national tests as a tool to further investigate and develop their teachings, both on individual and group levels. R1 stated that the ministry of education

made it clear that mapping tools shouldn't be used more than necessary, but also stated that mapping tools such as the national tests could be part of developing better teaching plans, especially when used in combination with other testing tools.

## **7.6 Utilizing the results**

### **7.6.1 Mapping**

*"But the really good thing is that during the fall of the first year, you can already see if there are any pupils scoring at level 1, for instance. Then you know there's a problem." L5*

In the pursuit of our master's thesis, it is necessary to examine the utilization of the results. In this regard, we have structured our analysis into three main components, specifically focusing on the utilization of results in assessment, methodological freedom, and development. Firstly, we will present the results in assessment.

The findings indicate that one out of the seven interview participants provided no response regarding their utilization of the results from the national test in English. Concurrently, five individuals assert employing the results of national test in English to assess pupils' academic proficiency. Additionally, one participant states that such assessments may serve as a valuable indicator provided they are utilized appropriately and in alignment with the intended purpose. Furthermore, one participant perceives that the national test in English is not accorded the same degree of significance as other assessment tools in the context of English education. Another participant utilizes the outcomes to gain a comprehensive overview of the areas in which pupils require further development, asserting that assessment forms the basis for educational advancement. Finally, one participant contends that assessment results offer valuable insights into a student's transition from primary to secondary school.

### **7.6.2 Methodological freedom**

*"You're free to do whatever you want with the results, and in a way, analyze them however you like." L2*

In the context of methodological freedom in utilizing results, one participant perceives a lack of directives from school leadership once the results are available, leaving it up to

individual teachers to chart the course for further action. Another participant also emphasizes that it is at the discretion of each teacher to determine how to leverage the results. Three participants mention the abundance of individualized approaches, with one viewing this positively, while another perceives it as varying depending on the teacher. Additionally, one participant desired more collaboration regarding the utilization of results but cites time constraints as a hindering factor. Furthermore, a school leader advocates for more teachers to be qualified to work with the results through ongoing professional development. Two participants make no mention of methodological freedom in the utilization of results.

### **7.6.3 Development**

*“It’s partly the results from the national tests that have led us to make continuing education and professional development a priority. And particularly in English.” R1*

One participant acknowledges their own potential for improvement by becoming more adept at actively using the results as a development tool. While two teachers and two school leaders recognize the rationale for using the results as a basis for development, they appear to approach it from different perspectives. One expresses that utilizing results is a valuable tool for examining the level of differentiation among pupils, while another uses the results as a starting point to address areas where pupils are struggling. The actions to be taken are left to the discretion of the teacher without any directives from the school leaders, and it may be assumed that achieving this depends on individual teachers possessing the necessary competencies and knowledge needed for sustainable development. Another participant, who incorporates the results in development asserts that it is up to each individual teacher and that they have autonomy without guidance from leadership.

Examining the response of the two school leaders, one is very clear in stating that they adhere to guidelines provided by UDIR to ensure valid data, emphasizing that this provides a better foundation for development without specifying any concrete actions they undertake. It is conceivable that valid data ensures genuine responses, which in turn is the desirable scenario. The other school leader views national tests as a comprehensive development tool and asserts that they utilize it as such. Furthermore, the school leader mentions using the results to implement further measures, such as

sending teachers for professional development to enhance the quality of their development work. Additionally, they conduct follow-ups on results in joint meeting where subject leaders review the relevant grades to assess the progress from a comparative perspective.

## 8 Discussion and analysis

### 8.1 Traditions

Traditions refer to established practices, rituals, or customs that are regularly observed and passed down through generations within the educational community. These traditions often serve to create a sense of identity, community, and continuity among pupils, teachers, and staff. Traditions play a role in creating a cohesive and vibrant school culture. They offer pupils and staff opportunities to connect, celebrate achievements, and establish a shared sense of belonging.

Looking at how the interviewees perceive tradition at their school regarding the utilization of national tests, there are two significant opinions. The two school leaders are positive when asked about their traditions, and further states that they have established routines such as a 'good analytical framework' and a 'set template'. Furthermore, they emphasize working in groups. In order to assure everything is going as planned, R2 visits every classroom to ensure this and to show the pupils that national tests matter. The other school expresses that they do not use the supervisor/guidance from UDIR due to their own system which she categorizes as 'well-working'. Seland, Vibe & Hovdhaugen (2013, p. 9) states that less than a third of the municipalities involved in their study checks that the schools have followed the exemption rules given by UDIR. Based on this, one may suggest that municipalities are not checking if schools have followed the supervisor/guidance, strengthening R1 being able to follow their own system.

Switching our focus from the school leaders to the teachers, all five teachers express negative feelings towards tradition stating things such as the work is left completely up to every individual, while others state that there is no collaborative work. This shows a significant change of views from school leaders to teachers, indicating a need for a well-structured system in order to further develop all parts involved. This indication is further strengthened by Meld. St. 21 (2016-2017, pp. 26-27) which states that improving the quality of teaching has the best conditions when teachers cooperate. Furthermore, the report states that pupils learning is dependent on teachers developing themselves. Teachers have a vital role regarding national tests and improving its quality,



but school leaders should communicate and anchor the schools' goals, ambitions, and values, and lay the foundation for a productive development culture (Meld. St. 21, 2016-2017, pp. 26-27), suggesting the need for more collaborative work, with systems that ensure that teachers possess the adequate skillset to achieve academic development. The lack of collaborative work is showcased through one of the interviewees stating, "there has not been anything collectively for many, many years". This may indicate a lack of clear guidelines for teachers to follow, furthermore suggesting that the teachers might not be sufficiently equipped to ensure the purpose of the national test in English. This is supported by Wayman & Jimerson (2014, p. 25) who points out that educators' struggle with a lack of knowledge about how to best use data to improve instructions, additionally indicating that teachers struggle with data to inform practice, citing issues such as data systems, principal leadership, and time.

Although there are many solutions to ensure cooperation within the organization, Wayman & Jimerson (2014, p. 33) states that the issues are systematic in nature. Therefore, they point out that changes to organizational practices could be the best way to build capacity in individual teachers. One can suggest that some issues are complex and difficult to find an appropriate solution to, furthermore indicating that a change should be implemented to all participants within the organization. Having a tradition that apparently is not well-working, promoting negative feelings from teachers is destructive for the school leaders, teachers, and pupils. Seland, Vibe & Hovdhaugen (2013, p. 135) mention the effects of how each teacher portrays national tests to their pupils as a vital component for development. Furthermore, if a teacher is negative, worried, or stressed, these feelings may be infused in the pupils. If the teachers subjective meaning is that national tests have little or no value, it could decrease the pupils' motivation.

The development of the school as an organization is often seen from a perspective of change, but equally important is the perspective of stability. In order to keep in touch with traditions of high academic, social, esthetical, and ethical standards, the continuous development of content, organizational- and structural methods in schools is paramount. Based on this as a starting point, Dalin (cited in Roald, 2012, p 120) presented five dimensions that are mutually dependent on each other in order to maintain high standard traditions, which are; values, structures, relations, strategies and

surroundings, further stating that one change to a dimension may affect the others, and that no one stands above another. This suggests that a well-structured system is necessary to maintain a high-quality learning development, further indicating that teachers and school leaders may not necessarily know their own traditions. Dalin (cited in Roald, 2012, p. 121) further suggests that if a well-working structure is present, routines and traditions are appreciated and continued. A vital aspect is that such structures can maintain changes and renewals that may occur.

Another important aspect regarding change in an organization is *mental models*. Senge (1997, p. 50) argues that mental models cannot be changed, but that they can be managed through dialogue and transparency. Furthermore, he points out the trait to acknowledge that mental models exist on both sides, and that either way of thinking should be discussed and tested in order for the organization to build new mental models or maintain the existing ones. According to Roald (2012, p. 130) even successful measures in some organizations can have no consequences because the mental models can prevent change and development. This may suggest that achieving a systematical structure and order is vital for the organization, school, leaders, teachers, and student to utilize the potential of national tests, further indicating it as a game of chess, where the first move lays the foundation for the rest. Depending on the organization and school, it may be a good and well-working system, but findings from existing research and the qualitative and quantitative data collected by us suggests that is necessary to create a superior tradition which facilitates a shared vision for the whole organization.

## **8.2 Methodological freedom**

### **8.2.1 Preparation phase**

When asked if they were familiar with the delegation of responsibilities regarding the implementation of the national tests, the vast majority of the respondents in our survey answered that they were familiar with their tasks. Comparing this to the findings from our qualitative data, one can suggest that some teachers are either unsure, or do not possess the adequate skillset to ensure their area of responsibility in the implementation of the national tests. Furthermore, this may indicate a lack of information, or poor individual preparation.

As teachers are mainly the ones in charge regarding the utilization of the results from the national tests for facilitation, the importance of them being equipped to do so is key. In order to inherit the necessary skills to utilize the test results, school leaders need to exert effective leadership with focus on strengthening systems that may otherwise fail, in accordance to Tonich (2021, p. 66). The quantitative data collected by us indicates a wide spread of answers from teachers when asked to what extent they feel equipped to utilize the results for facilitation, while no school leaders express their concern towards this. This is correspondent with previous research, that show leaders view the national tests as more meaningful compared to teachers, who are more vague in their perception of the national tests (Seland, Vibe & Hovdhaugen, 2013, p. 27). One may suggest that the teachers and school leaders needs to view themselves as both the solution and the problem in order to improve the complexity of the national tests, in accordance to Senge (cited in Roald, 2012, p. 129). Moreover, the school leaders should be aware of how their leadership could be part of fostering high levels of performance, using their “invisible hand” to guide the teachers, as stated by Peterson and Singh (2013; 2014, cited in Tonich, 2021, p. 65). Based on our findings, one may suggest that many teachers and school leaders do not prioritize the national tests more than they have to, often using excuses such as the lack of time to deprioritize the use of the national tests results for facilitation. A valid point of interest here is whether or not the methodological freedom experienced by many teachers plays a part of this.

In the process of preparing for the national tests, all the teachers that were interviewed experienced a high degree of methodological freedom. The informants all stated that they themselves decided how much time they wanted to spend to prepare the pupils for the test, however, they all seemed to prepare their pupils for the test at hand, i.e. prepare them for a long test of a format that the pupils weren't necessarily familiar with.

As shown through the quantitative research, the majority of the informants expressed that they received adequate time in the preparation phase. Nevertheless, 12 teachers answered ‘very little’ and 16 teachers answered ‘neither nor’, showing a wide spread of answers. These findings partially align with the findings from the qualitative data, where all informants expressed the need for more time. An intriguing aspect is that most of the respondents in the quantitative survey believed that adequate time was allocated for preparation, yet many rated their responses as ‘very low’ or ‘low’ on questions relating

to the national tests as development- and assessment tools, familiarity with guidelines, and similar areas. The qualitative data shows that time is perceived by many informants as a significant factor for not prioritizing the utilization of the results from the national test in English, a finding that does not emerge in the quantitative data. One might assume that teachers who express that they are given more than enough time to work with the national tests could mean that they have sufficient time to use the tests as assessment- and development tools, whilst at the same time staying updated on what the guidelines and purpose of the tests are.

The teachers had somewhat varied views on the national tests, with some expressing that the tests were something that simply 'needed to be done', and others that saw them as an opportunity to learn more about their pupils' academic level, with the aim of adjusting their teachings to fit the individual pupils and the class as a unit. The teachers and the principals that were interviewed all pointed out that there was some information from higher levels (i.e. the school owners and the school leaders), but that the information was mostly focused on practical features such as when the tests were to be carried out, and other important features of the preparation such as reporting pupils that for various reasons could be exempted from the tests. Some of the teachers mentioned that there was *some* cooperation between the teachers in the preparation phase of the national tests, but that each teacher chose how to prepare the pupils using their own methods.

As Dalin (Roald, 2012, p. 121) points out, good structures in the school are necessary for routines and traditions to be appreciated and continued. The apparent lack of structures regarding how the teachers prepared their pupils for the national tests could be seen as problematic, since the national test results are dependent on valid data to fulfill its potential as both a mapping- and development tool. However, as Dalin (Roald, 2012, p. 121) argues, structures must be flexible enough for change and renewal to occur, raising the question of whether such methodological freedoms should be seen as positive or negative. The result from our quantitative survey corresponds with the results from our qualitative analysis, where most teachers and school leaders utilize individual tasks in the preparational phase. While the majority utilizes it, 10 teachers and two school leaders use it 'very little' or 'little', indicating inadequate expertise or lack of priority.

Furthermore, UDIR recommends the use of examples task in the preparation phase of the national tests.

As expressed by L4, the possibilities for him to decide what to focus on in the preparation for the national tests made it possible for him to use his experience as a teacher and knowledge about his pupils to prepare them as he saw fit to prepare the pupils as he saw fit. Other teachers, such as L5 and L2 called for more cooperation, with L2 pointing out that a communal view of how the pupils were prepared for the tests could assure that uncertainties and confusions could increase the validity of the test data. The loose connections seen through the lack of information and individual methods used in the preparation phase of the national tests could therefore be seen as both good and bad. Pointing out that a characterization of most schools as individual units that are isolated from each other, Dalin (2005, p. 45) raises the question of whether the tendency of such units operating within their own routines, and not necessarily following guidelines set forth by various directories it should be seen as good or bad, arguing that the relative freedom by such loose couplings could have the positive effect of an ease of adaptation into the environments that the units operate within. However, through the statements made by Fevolden & Lillejord (2012, p. 117) and the data collected through our interviews, it is our view that loose couplings in the preparation phase of the national tests could have the effect of creating uncertainty within the teachers, resulting in some frustration which ultimately has the outcome of less valid data. The establishment of a corporate culture that rewards investigation and enquiry, and which empowers the teachers to experiment, and the inclusion of a corporate vision with a shared view of how the preparation of the national tests should be carried seems to be something that the informants' views as lacking, starting with the limited amount of information from school leaders and school owners.

### **8.2.2 Mapping- and development phase**

Moving on to how the results from the national tests are used as a tool in the academic development of the English subject, it becomes evident that the lack of a shared vision and good strategies affects the use of the national tests as a development tool. With most of the teachers viewing the national tests as a mapping tool, both principals that were interviewed stated that that they used the test results to develop their teachings. This is

in accordance with the previous research conducted by Seland, Vibe & Hovdhaugen (2013), which showed that school owners and leaders viewed the national tests as a good tool that measure what it is intended to measure, whilst teachers had a more restrained view of the national tests and its capability to serve as a development tool. Through the interviews conducted for this thesis, the lack of cooperation and information in the process of using the results from the national tests as a development tool was brought forth as an issue by all teachers. As Meld. St. 21 stresses, the task of improving the quality in the school has the best conditions in schools where teachers cooperate, and school leaders should communicate and anchor the schools' goals, ambitions, and values to lay the grounds for a healthy development culture (Meld. St. 21, 2016-2017, pp. 26-27). As stated by Tonich (2021, p. 63), school leaders are of clear importance in organizing school life in order to achieve optimal outcomes. Thus, the leaders should use their knowledge of their staff to 'set the rhythm' for the school (ibid.). Similar to the informants' views with regards to the preparation phase of the national tests, factors such as cooperation and shared vision becomes an apparent lack in the use of the national tests as a development tool. As Senge (2005, p. 51) points out, the view of an organization as a system where cause and effect are woven together could sway the organization (in this case, the school) in the direction of meaningful change.

Another interesting aspect reveals itself when looking at the quantitative results regarding how teachers and school leaders understand the results from the national tests, as not a single teacher nor leader answers 'very little' or 'little'. This could suggest that all participants inherit the knowledge of how to read the results, without necessarily being able to utilize them. One could imagine that if the teachers and school leaders possess the adequate knowledge of how to understand the results from the national tests, they would be able utilize the national tests as a mapping- and development tool. This finding could indicate that factors such as systems thinking, mental models, information flow, and team learning are not properly enhanced within the school system. When interpreting the quantitative and qualitative data, one could suggest that not all teachers and school leaders possess the sufficient knowledge of how to understand the results from the national tests, although they themselves might believe so. Furthermore, this indicates the importance of a thorough review, ensuring that each teacher and school leaders knows their responsibility to further enhance the

learning outcome of both the pupils and themselves, using the results from the national tests as part of the process.

Five teachers mention the amount of responsibility it is up to each teacher in the context of methodological freedom. Fevolden & Lillejord (2005, pp. 100-101) points out 'loose connections' which concerns what it to be taught and which methods that are used by the respective teacher with little influence from the school or municipality. This may suggest the further increasing gap in how the respective teacher chooses to utilize the results. Further indicating a lack of check-up starting from the state à municipality à schools à teachers. However, one may suggest that without clear directives from the top of the chain, it may be confusing and difficult to engage with these solutions as they are perceived easy to follow. If the solutions that are presented are easy to follow, one could speculate why they have not been enforced yet, suggesting the complexity of the matter starting from the higher levels of the organization and moving down.

Several of the teachers that were interviewed were under the impression that the national tests were viewed by their respective school leaders as something that was of lesser importance compared to other mapping- and development tools. The apparent tendency of viewing the national tests as a mapping- and development tool of lesser importance than other tests did affect some of the teachers, such as L1, who emphasized that the lack of focus on the national tests from her/his leaders had the effect of the teachers deprioritizing the tests. As L1 mentioned, the general idea at his school was for teachers to cooperate with each other in the analyzing and interpreting of the test results, with the goal of using them for both mapping and development of their teachings. However, L1 stressed that no time was set aside for the teachers to do so, rendering the task nearly impossible to do. Similarly, L2 stated that the leaders at his school encouraged the teachers to present and talk about the results in development conversations with their pupils, but that no instructions regarding how the results could- or should be used for mapping and/or development. As part of fostering a culture for improvement in the school, school leaders should master and understand all aspects of their role as an educational administrator, acting as effective agent of change to trigger improvement in the school, as stated by Tonich (2021, pp. 66-67). Through the data we have collected, we see the need for school leaders to use their capabilities of being such agents of change, using their managerial abilities to directly contribute to the

work with using the results from the national test in English as part of the school's academic development. We argue that the school leaders should prioritize setting aside time for collaboration, since doing so could improve the moral of the teachers, creating the possibility for better utilization of the test results.

An interesting observation was that both principals that were interviewed pointed out that the tests were used for development at their schools, and that the lack of other mapping tools in the English subject (compared to other subjects in school) made the national tests stand out as an important tool for them. Again, the research conducted by Seland, Vibe & Hovdhaugen (2013) can be used to show that there is a lack of shared vision between how the teachers and the school leaders view the national tests. As Dalin (2005) argues, problem solving, decision making, delegation etc. are aspects of strategies that leaders in schools could find challenging. However, the necessity for such strategies becomes evident through the research that was conducted for this thesis. As previously mentioned, most of the teachers pointed to a lack of time in an otherwise busy workday made it difficult for the teachers to cooperate in the use of the national test results for mapping and development. L5 specifically mentioned that she wished for more time to cooperate with her colleagues, stating that she didn't know which methods they used in their work with the national test results- whatsoever. There is, however, reason to debate whether the teachers could work together on a more regular basis through their own initiative, without receiving clear instructions from their school leaders. As stated by both principals, the school owners delegate responsibility to the principals, who in turn delegate responsibility to the teachers. Although the teachers show a tendency to view factors such as time, prioritization and information as inhibiting, there is need for all individuals within the school to see themselves as part of both the problem and the solution, so that they can take part in the development of the school and its methods rather than putting blame on others. This, and the need for management of the individuals' mental models seems to be a key in the work of using the national tests as a mapping- and development tool, with Senge (1997, p. 50) stating that accurate and flexible mental models enable individuals within the organization to navigate complex situations effectively.



## **8.3 Improvement potential**

### **8.3.1 Preparation phase**

Through analyzing our qualitative data, we found that the interviewees all stated time as a reason for not being able to properly utilize the national tests as a sufficient tool, yet the findings from the quantitative survey indicated that time is not the sole reason for the dissatisfactory use of the national test results. This may suggest that teachers and school leaders might not prioritize the national test in English as an adequate mapping- and development tool, further indicating that the need for well-structured routines in order to maintain the purpose of the test. One could suggest that if teachers and school leaders does not feel that the national tests are prioritized, they are less likely to devote time to them due to the need to complete other work-related tasks. This highlights the importance of having clear routines in place, ensuring a sufficient flow of information, working in teams with a focus on fulfilling the purpose of the national tests, and enhancing them for the benefit of both the pupils and the respective schools. Based on the data that has been collected, one could suggest that the current level of prioritization is questionable, and that a significantly higher emphasis on the national tests from the higher levels of the school as an organization down to the teachers is needed. It could be unreasonable to expect that teachers dedicate significant time to the work with the national tests if they experience a lack of prioritization and significance in their workplace, in accordance with Hardy (2015, p. 4), who found that the implementation of policies could have the teachers feel tired and overloaded.

In the process of analyzing the interviews, one factor in particular was found to be of importance when it comes to improving the preparation phase of the national tests. The need for better instructions from UDIR, the school owners and the school leaders are presented as a deficiency. In many ways, it is our view that the hierarchical structure that the school as an organization is built as creates what could be seen as a form of a Chinese whispering game, where information is lost as it is passed down between the different levels in the school. As stated in Meld. St. 21 (2016-2017, p. 13), the lead role in the task of developing the quality in the school should be held by the teachers, school leaders and school owners, with the state providing the necessary frames around the local action space. Through the interviews conducted by us, it became apparent that the

two principals relied on 1) basic information from the school owners regarding the preparations for the national tests, and 2) their own routines that were specific for their own school. Four of the five teachers experienced that the information given from their school leaders was insufficient, with L3 and L5 stating that they were left with the responsibility to prepare themselves and their pupils for the tests with seemingly no information on how to do so. As Roald (2012) argues, even successful measures could prove insufficient in the development of the school if dialogue and discussion is not present, making it important for the schools to prioritize team learning to align and develop the capacity of the team, to create the desired routines built on a shared vision.

Ideally, all teachers and school leaders should be sufficiently familiar with the purpose of the national tests and the guidelines provided by UDIR and should be able to explain the purpose of the national tests to their pupils. Our findings indicate that this is not the case, further suggesting that it is dependent on several aspects such as information flow, systems thinking and mental models. It is necessary to have a well-structured information flow to keep all parties involved with the national tests up to date with vital aspects of the tests, such as the purpose of the tests, the benefit of implementing them, and how to utilize the test results. Furthermore, one can see the several aspects as part of a bigger picture, where cooperating together is of high importance due to the large gap in how each teacher and school leader perceives the various aspects of the national tests. Additionally, it is noteworthy that two school leaders that responded to our survey do not explain what the purpose of the national test in English is to their pupils, maintaining the large gap between what is desirable and not. One could suggest that the unfamiliarity regarding the purpose of the national tests could affect the learning outcome of the pupils, with accordance to Seland, Vibe & Hovdhaugen (2013, p. 135) who states that if teachers express negative attitudes towards the national tests, the pupils could be infused with pressure or fear towards the tests.

### **8.3.2 Development phase**

A development assessment tool in schools is a structured method used to measure pupils' learning progress, skills, or competencies to inform educational decisions and support student development. These tools are designed to provide educators, administrators, and policymakers with valuable insights to student performance,

enabling them to tailor instruction and interventions to meet the unique needs of each student or group of pupils.

Development assessment tools play a crucial role in schools by helping educators identify strengths and weaknesses, track progress over time, and create personalized learning pathways. They also contribute to school improvement efforts and policy decisions by providing data-driven insights into educational outcomes.

It is noteworthy that 20 teachers replied 'neither nor', and 11 teachers replied 'little' or 'very little' regarding their perception of the national tests as a mapping tool. This does not correlate with the results we found in our qualitative data, where the majority of our interviewees stated that they viewed the national tests as a good mapping tool. As shown through the research conducted by Beck & Stetz (1979, p. 4), the majority of the teachers that responded to their extensive survey viewed standardized testing as a useful tool, with many pointing to factors such as diagnosing strengths and weaknesses and the measurement of knowledge benefiting from the test results. Our findings may suggest a coherence between the lack of focus and prioritization, as the data from the qualitative data clearly shows that the teachers actively use the national tests as a mapping tool. Looking at the results from our quantitative data, it becomes evident that the school leaders are mainly positive towards the national tests as a mapping tool, whilst the teachers show a wider spread. This is strengthened by Seland, Vibe & Hovdhaugen (2013, p. 27), who found that school leaders are more positive towards the national tests than teachers.

While it is satisfactory that school leaders show positive attitudes towards the national tests, it is noteworthy that many teachers show the opposite, further indicating a significant gap regarding how the national tests are viewed by school leaders and teachers. Teachers are the main factor for the pupils' learning outcome, and it is therefore crucial that they understand the benefits of the national tests, seeing how the teachers are responsible for the practical implementation of the tests, not the school leaders. Our findings show that methodological freedom is highly valued, but one may suggest that clear guidelines are necessary to ensure that both teachers and school leaders inherit the same visions and goals regarding the utilization of the national tests.

In the context of utilizing results for development, it is noteworthy that two participants remain silent. This may indicate a lack of clarity in the development aspect if individuals have not received adequate training and do not possess mastery over the material of the results. This is evident from one participant who highlights that this area has not been a focus at their school, consequently finding it challenging to utilize the results as a development tool. Another participant, who incorporates the results in development asserts that it is up to each individual teacher and that they have autonomy without guidance from leadership. Pierce and Chick (2011, cited in Werler & Færevaaag, 2017, p. 69) points out that it is unlikely Norwegian teachers possess the necessary knowledge that is required to compare, contrast and critique multiple datasets. One participant acknowledges his own potential for improvement by becoming more adept at actively using the results as a development tool. While two teachers and two school leaders recognize the rationale for using the results as a basis for development, they appear to approach it from different perspectives. One expresses that utilizing results is a valuable tool for examining the level of differentiation among pupils, while another uses the results as a starting point to address areas where pupils are struggling. The actions to be taken are left to the discretion of the teacher without any directives from leadership, and it may be assumed that achieving this depends on individual teachers possessing the necessary competencies and knowledges to be successful in their work.

When working with the interpreting, analyzation, and implementation of the tests results for use in the academic development, all teachers that were interviewed pointed out that lack of time and prioritization was a factor that needed improvement. As L5 stated, her view was that if the school leaders failed to set aside time for the teachers to work both individually and together with their colleagues, the message that could be perceived by the teachers was that the school leaders considered the national tests as unimportant, making the work with the national tests of a lesser priority compared to other things that needed to be done. L2 and L4 specifically stated that there had been no focus on how the results from the national test in English could be used as part of the school's academic development in in the English subject, with L4 saying that he had never heard of anyone using the test results as part of the development at all. Furthermore, R1 noted that there are no common routines on how the test results should be used, implying that each school had to do this as they saw fit. L1 implied that a

common desire among the teachers was for there to be an opportunity for them to work together in analyzing, interpreting and utilizing the results, whilst L4 stated that despite there being talk of poor test results in collegial meetings, no time was spent addressing how to improve their teachings with the goal of initiating better learning for the pupils—once again implying the lack of a shared vision regarding the utilization of the national test results for academic development. An apparent issue with this becomes evident when R1 later stated that his school decided to prioritize “other things”, with the outcome of them spending virtually no time on the interpreting, analyzing, and utilizing of the tests results. As argued by Wayman & Jimerson, changes to organizational practices could be the best way to build capacity in individual teachers (2014, p. 33), thus, as implied by both teachers and principals in our research, common routines, clear instructions, and collegial cooperation based on a shared vision becomes necessary.

Furthermore, the majority of the teachers and school leaders that answered our survey stated that they were equipped to interpret the results, with only nine teachers and one school leader expressing that they struggled with this. The respondents that suggested that the interpretation of the results from the national tests was a difficult area, further promoted the idea of a better learning area within the workplace, an idea that is further supported by the results from our interviews, where the informants called for more cooperation. In contrast, the ones who expressed that they were equipped to interpret the results could well be equipped to do so, however, one may suggest that this stems from them interpreting the results using a method that they see fit, with little or no cooperation with other teachers. As pointed out by Werler & Færevaaag (2017, pp. 68-70), the limitations of the information that is provided through the (national) test could make it difficult for teachers to pinpoint which variables they should change to improve pupils’ learning outcome. Meanwhile, the results from our interviews showed that L5 didn’t know which methods her colleagues used when interpreting the results from the national tests, making her unsure of whether she did it ‘correctly’. Furthermore, R2 stated that a situation where his teachers had to resort to guesswork in their work with interpreting the results from the national tests was undesirable. With these findings as a foundation, one could suggest that an increase of focus regarding collaboration and team learning could have the effect of teachers reaching a common understanding of the results, ensuring a more consistent way of utilizing the results from the national tests.

## 8.4 Mapping- and development tool

Looking at the use of the results from the national test in English as a development tool, the majority of the respondents in our survey viewed the national tests as a development tool of little sufficiency, further aligning with the findings in our qualitative data. This may indicate that teachers' subjective opinion and attitudes towards the national tests could be passed down to the pupils, further suggesting the importance of implementing good attitudes towards the national tests, seeing how the opposite could have a negative impact on the pupils' views of the tests. This is further strengthened by Seland, Vibe & Hovdhaugen (2013, p. 135), who state that if a teacher possesses negative feelings towards the national tests, the validity of the data collected through the tests could be weakened. The majority of the school leaders perceive the national tests as a sufficient tool for development, and this corresponds with previous research that shows that leaders are more positive towards the national tests compared to teachers (Seland, Vibe & Hovdhaugen, 2013, p. 27). Furthermore, the answers indicate a clear discrepancy between the perspective of teachers and school leaders, suggesting a lack of emphasis on the national tests as a development tool. This observation further aligns with the findings from our qualitative data. Additionally, one may suggest that further education for teachers could ensure greater competence in the development of the national tests, strengthening the views with R1, who empathizes the importance of having teachers who are qualified.

In the context of utilizing results for development, it is noteworthy that two participants remain silent. This may indicate a lack of clarity in the development aspect if individuals have not received adequate training and do not possess mastery over the material of the results. This is evident from one participant who highlights that this area has not been a focus at their school, consequently finding it challenging to utilize the results as a development tool.

Our findings indicate that one out of the seven interview participants provided no response regarding their utilization of assessment results, while five individuals expressed that they utilize it to assess pupils' academic development. Teachers are expected to use the results to gain an extensive overview and provide ongoing

assessment for the pupils, and the research of Beck & Stetz (1979, p. 4) shows that when prioritized, standardized tests could be a useful tool for not only the mapping of individual pupils' academic status, but also a useful tool for helping to plan instruction for the entire classroom. The reason behind municipalities and schools to utilize the results as a basis is due to enhancing the quality of education. Furthermore, providing relevant information which can be utilized by school leaders and teachers for the sufficient development needed.

Through our research, it becomes evident that a common view of the national tests in English is that the tests are a useful mapping tool, but that its potential remains unfulfilled. As previously mentioned, the research conducted by Werler & Færevaaag (2017, pp. 68-70) show that limitations of the information that is provided through the test results might make it difficult for teachers to pinpoint which variables they should change to improve pupils' learning outcome. Furthermore, Wayman & Jimerson (2014) point out that educators' struggle with using data to inform practice, with issues such as principal leadership, time, and lack of knowledge being some of the factors that contribute to this. Consequently, Wayman & Jimerson suggest that collaboration, common understanding, triangulation, and time are important skills for teachers to possess in order to link data to practice. The teachers that were interviewed as part of our research all pointed to the lack of collaboration, common understanding and time as factors that prevented them from utilizing the full potential of the national tests. Although the majority of the teachers that were interviewed seemed to have clear meaning as to what could have been done better in all phases of the work with the national tests, L3 did mention that he hadn't really thought too much about it before, and stated that "now that we've talked about it, I really do wish that more time had been set aside for us to work with the national test results". Additionally, teachers experience little contribution from school owners regarding developing their competence, further indicating that the development of cooperation is not optimal in Norwegian schools, despite Norway meeting the international average seen from an international perspective.

### **8.4.1 Frequency**

Looking at the results within frequency, four out of seven interviewees wish to increase the yearly number of national tests, with three of them seeing the benefit of having it in 9<sup>th</sup> grade. As mentioned previously, the results from the national tests are categorized into various proficiency levels, with different delineations for 5<sup>th</sup> and 8<sup>th</sup> grade (UDIR, 2022c). UDIR states that in order to achieve a comprehensive assessment of educational progress, it is necessary to compare results for the same student cohort across different years (UDIR, 2022b). While, L5 experiences having the tests with a three-year gap increases the difficulties of enabling an overall understanding and affiliation regarding national tests. Based on this, one can suggest that a three-year gap between the national test in English is too long to ensure that the test results can be utilized for sustainable development. Therefore, we argue that implementing national test in English in 9<sup>th</sup> grade would strengthen and enable the quality of the work. This is supported by L4 who suggests that the implementation of it in 9<sup>th</sup> grade would engage and motivate school leaders and teachers as it would enable more reliable data and progress. Furthermore, R2 agrees with L4 suggesting that it would increase the potential of national tests as a development tool and create a more detailed assessment of each pupil.

### **8.4.2 Information flow**

Information flow is an essential part of a well-working system in order to achieve the desirable results. Of the seven interviewees, only one stated he was happy with the information flow, while three stated they were not happy. Two declined to answer, and one said it was “alright”. The one who expressed he was happy was a school leader, while the three who expressed dissatisfaction were school teachers. Through Seland, Vibe & Hovdhaugen (2013, p. 27), research shows that school leaders and school owners view national tests as a good tool to measure the intended purpose. In contrast, the same cannot be said for teachers, indicating a more restrained approach towards national tests, further suggesting that it is not necessarily a good development tool. This may indicate that teachers view national tests as a good tool when used in accordance with its purpose and goals. In addition, Seland, Vibe & Hovdhaugen (2013, p. 23) show through their findings that teachers express frustration due to the deficient information regarding the national tests. This may suggest that teachers may not have the necessary



competence to ensure a good quality outcome of the results, emphasizing the need for essential training from their school leaders. One can suggest that teachers without the expected competence, will not be able to utilize the national test in English as UDIR desires. This is supported by L2 who says: “But, we could have used it more actively in our teaching, I think”, suggesting that with increased competence, the aim for national tests would be adhered. The part of the response, “I think”, may further indicate that the teacher is unaware of the potential national tests may have. The clear gap between teachers and school leaders strengthens the importance clarity and defining roles may have. This is backed up by, UDIR, who promotes school owners and leaders to actively deliberate on the reasons behind the results, and to formulate strategies and actions to implement necessary interventions to teachers.

The majority of teachers and school leaders are well-familiar with the guidance materials, but one can suggest that with clearer guidelines, all teachers and school leaders would be familiar with it. In order to improve this, one may suggest that a better and more clear information flow and system thinking is crucial to have everyone aligned with the same vision and to create a consistency within the utilization of the national test in English. This could indicate the need for stronger clarity of each teacher’s role, strengthening the information flow resulting in a more systematic way of working. Furthermore, Senge (1997, p. 48-51) argues that issues caused by modern organizations often stem from a lack of systematical thinking, combined with the need for a general and mutual view of values and visions for those within the organization.

## 9 Answering the research questions

National tests are standardized assessments used across the whole country to measure pupils' skills in core areas such as English. These tests offer valuable insights into pupils' academic levels, allowing teachers, schools, and educational authorities to identify areas for improvement and adapt teaching strategies accordingly. Based on this we aim to answer our three research questions, using the qualitative and quantitative data that has been presented and discussed.

Firstly, we seek to answer: *To what extent do English teachers and school leaders experience that the national tests are a prioritized tool in the mapping and development of the English subject/ academic development of the pupils?*

As shown through our research, most of the informants view the national test in English as a useful mapping tool. However, the view of the national test in English as a development tool for the English subject is varied. This could stem from many factors, but our findings show that a typical trend is that the lack of focus on the national test in English as a development tool could be the main cause, with some of our informants stating that they only used the national test results as part of the conversations with pupils or parents, with little or no focus on how they could use the results for individual- or collective development.

Looking at how the national test in English is perceived as a mapping- and development tool, we asked this in both our quantitative survey and to our interviewees in the qualitative. Our findings from this clearly shows that many informants do not perceive the national test in English as a prioritized tool for mapping- and development. Furthermore, another finding occurred through this discovery, with teachers showing dissatisfaction towards the amount of work being individualized. There is a wish for more collegial work to be underheld within each school, and a greater focus on cooperation, as our informants feel a lack in this department.

Furthermore, the majority of the informants in our quantitative survey state that they receive adequate time in the preparation phase, however, our findings indicate that they

do not utilize this time. In contrast to this indication, our findings from the qualitative research suggests that teachers especially mention the need for more time to appropriately ensure the academic development of the pupils through the results from the national test in English. There is a clear contradiction between the two findings, further suggesting the lack of guidelines and clarity within the school as an organization. This is further strengthened by teachers who state they decide how much time they wish to use in the preparation phase, suggesting that time is not a problematic factor.

In accordance with previous research, our findings show that the school leaders tend to be more positive towards the national tests as a mapping and development tool compared to the teachers (Seland, Vibe & Hovdhaugen, 2013, p. 27). In particular, one of the school leaders that we interviewed stated that her school used the test results as an active part of the schools' academic development, whilst the other school leader stated that his school used the test results as part of "determining the direction of future action". Although our data is limited regarding the specifics of why this is the case, we find that the lack of a shared vision, team learning, structures and strategies play a large part in the continuation of the divergence between teachers' and school leaders views on the national test in English.

Secondly, we will address the research question: *To what extent does established routines affect the utilization of the national tests as a mapping and development tool by the teachers?*

As stated in the previous paragraph, our informants express a lack of cooperation and collegial work. Furthermore, they point out their dissatisfaction with how the administration of the national test in English is left to individual teachers, resulting in significant disparities regarding utilization, which may impact the motivation and pride of those involved further decreasing the quality of the work. In order to create and maintain consistency, it is necessary to have established routines which ensures the purpose of the national test in English and creates clarity in all departments. Our informants point out the need for more collaborative work, and this can be enforced through cooperation and working in subject sections, which our findings indicate.

As seen through our own- and previous research, there is a significant gap in how teachers and school leaders utilize the national test in English as a development tool. Teachers need effective routines, and when this is not the case, it is easy to point towards the school leaders who further points to the school administrators and UDIR. The whole process is too inconsistent due to the vague follow-up of each school and relevant people not possessing the necessary competence in order to utilize the national test in English as a mapping- and development tool. Based on this, the process in regard to the utilization of the results is seen as a game of “Chinese whispers”, due to information seemingly “getting lost” as it is passed down through each department of the organization.

As stated by Senge (1997, p. 51), a true shared vision attracts the commitment of all participants in an organization. Furthermore, team learning is defined by Senge as creating the desired results within an organization as a team (ibid.). When looking at the results from our research, it becomes apparent that increased focus towards these two factors could contribute to better routines in the school as an organization, that could have the effect of an increase in the utilization of the national test in English as a mapping- and development tool. Our findings show that the lack of cooperation between the teachers makes it difficult for some teachers to utilize the results from the national test in English, ultimately making some of the teachers deprioritizing the national test in English as a mapping and development tool. Therefore, we argue that the increase of dialogue between teachers, school leaders and the higher levels of the educational system could have the effect of the entire school as an organization finding a common goal, that could make it easier for all individuals in the organization to use the national test results as they are intended to be used.

Regardless of the level of cooperation there is between teachers when using the national test results for mapping and development, our findings show that there is a gap between how the teachers and school leaders view the national tests as a whole. In order to effectively establish quality routines, it is our view that all participants in the school as an organization needs to see themselves as part of the problem *and* the solution. According to Roald (2012, p. 129), doing so could have the effect of the participants taking part of the development process rather than putting blame on other factors, with

time, information flow, cooperation etc. being factors that were often viewed as inhibiting by many of the teachers that we interviewed.

Looking at how methodological freedom is perceived by teachers and school leaders, their views regarding this are both positive and negative. Through our findings, it is evident that all teachers experience a high degree of methodological freedom within the utilization of national tests. Furthermore, our findings show that the great amount of methodological freedom is perceived different by both teachers and school leaders indicating a widespread of information stemming from the information flow. There is no evident definition of how the information flow shall be utilized, however the varied use indicates a lack of routines and clear guidelines in how each school effectively use their information flow.

Lastly, we delve into our third research question; *To what extent do English teachers and school leaders consider national tests as a significant tool in academic development of the pupils?*

When assessing our findings from how teachers and school leaders consider the national test in English, our quantitative and qualitative data indicate that the majority view it as a significant mapping- and development tool. While they perceive it as a positive tool, there is a discrepancy in how teachers and school leaders utilize it in the academic development of the pupils. Our findings indicate the reason behind the discrepancy is how national tests is utilized and prioritized, and this consists of three factors: information flow, clear guidelines, and relevant competence to analyze, retrieve and utilize data.

The utilization of the results is mainly up to each teacher to ensure the facilitations of their results and use it as a tool for further development, hence the importance of teachers being equipped to do so. Additionally, school leaders could benefit from relevant training to ensure teachers are able to ensure the purpose of national test in English. The result of a lack of knowledge and training is a significant variation in the priorities and choices each teacher and school administrator makes.

This aligns with our findings and previous research, where there is a lack of clear guidelines, and where methodological freedom plays a vital role, resulting in teachers and school leaders having a great deal of latitude without the necessary oversight from higher authorities. Through our findings from the quantitative and qualitative research, the response from the participants indicates a lack of directives from the school administrators strengthening our interpretation.

## 10 Answering the thesis question

The significant disparities in how English teachers and school leaders view the national test in English as a mapping- and development tool, becomes apparent through both our quantitative and qualitative research. Whilst the English teachers tend to view the national test as a mere mapping tool, the school leaders view it part of plotting the course for future teaching. This is in accordance with previous research, that found teachers to be more reluctant towards national tests compared to the school leaders (Seland, Vibe & Hovdhaugen, 2013, p. 27).

Our research shows that this reluctance stems from several variables, with the lack of clear instructions, cooperation, lack of time, and prioritization being significant factors. We found that the teachers we interviewed called for better information in all phases of the national test in English, emphasizing the lack of clear guidelines and the need for more established routines. Our data indicates that this lack of information generates uncertainty amongst the teachers, leading them to use individual methods when preparing and implementing the tests, as well as the work with analyzing and utilizing the test results. We find that the amount of methodological freedom is not beneficial for the school's sustainability regarding their use of the results from the national test in English for mapping and development.

This shows the great range of teachers and school leader's prioritizations due to the difference in responses, further strengthening the case for uncertainty being a central factor. Furthermore, it is not only the teacher's responsibility to ensure that the national test in English is adequately taken care of, but school leaders also need to set a good example for teachers to follow. To develop an effective system, the establishment of clear routines is necessary. If school leaders do not prioritize national tests, it can be argued that teachers will neither.

Our research indicates that a change in the school as an organization could be a decisive factor in using the results from the national test in English for the academic development of the pupils. Working towards a shared vision for all educators within the school as an organization could be beneficiary to ensure that school leaders and

teachers utilize the results from the national test in English for quality improvement and academic development of the pupils.

The answer to our thesis question *to what extent do a selection of schools in northern Norway use the results from the national test in English as part of the school's academic development?* is therefore that the results from the national test in English are used to a varying degree, dependent on how each individual school chooses to emphasize the national test in English as an important mapping- and development tool. A change in the organization's overall structure, where systems thinking is emphasized to understand the broader context of the national test in English, could lead to a more consistent utilization of the results across all schools. This consistency could, in turn, promote sustainable development over time.



## **11 Suggestions for further research**

In order to ensure the sustainability of national test in English, we will present our suggestions for the future growth of it. Firstly, it is necessary to have clear guidelines on how to approach the process before, during, and after the implementation of the national test in English. This includes clear directives on how to use the results as both a mapping- and development tool, which requires a clear and strong leadership and collaborative work amongst teachers and school leaders. Secondly, established routines is essential to make a change in the organizations overall structure where systems thinking is emphasized to understand the broader context of the National Test in English. This could lead to a more consistent utilization of the results across all schools, additionally this consistency could promote sustainable development over time. Lastly, the need for a higher frequency is wished from teachers and school leaders and needed to ensure progressive results for comparison. This could increase the motivation each teacher and school leader have toward the national test in English, and help them to see they broader context of its use. Being able to implement National Test in English in 6<sup>th</sup> and 9<sup>th</sup> grade, would ensure it being a better tool equipped for comparing results which would increase the data the teachers and school leaders possess to further develop it as a mapping- and development tool.

## 12 Concluding remarks

Through conducting the research for this thesis, we now inherit knowledge about the national test in English at a far deeper level. We have learned a lot about not only the test itself, but also about factors that play a central part of the national tests as a mapping- and development tool. As future teachers, we wish to use this knowledge as part of our own practice, maintaining the utilization of the national test results as UDIR defines it. Our aim in doing the research was to increase our knowledge about the national tests, and to direct attention towards what we see as a central factor within the teaching role. Additionally, it has been important for us to direct attention towards northern Norway, since our view is that the amount of research regarding the national tests conducted in this region has been insufficient. We hope that future students, teachers and school leaders might find inspiration in this master thesis, and that they can use our research for their own development. The road towards utilizing the results from the national test in English in a sustainable way might be a difficult one, but as stated by one of our interviewees: “you can’t do a 100-meter sprint if you’re competing in a marathon. You need to know what you’re doing!”

## Works cited

- Alnes, E., Rørslett, K., Kyllingstad, O. R., Nylenna, M., & Pedersen, E. (2023). Frida støtter forslaget om å avvikle nasjonale prøver og eksamen. *NRK*.  
<https://www.nrk.no/norge/kunnskapsministeren-far-rapport-om-nasjonale-prover-i-skulen-1.16635024>
- Beck, M. D., & Stetz, F. P. (1979). Teachers' Opinions of Standardized Test Use and Usefulness.
- Creswell, J. W., & Guetterman, T. C. (2021). *Educational research : planning, conducting, and evaluating quantitative and qualitative research* (Sixth edition. ed.). Pearson Education Limited.
- Dalin, P. (2005). *School development: Theories and strategies. An international handbook*. London u.a: Continuum.
- De nasjonale forskningsetiske komiteene. (2019). *Generelle forskningsetiske retningslinjer*. De nasjonale forskningsetiske komiteene. Retrieved 10.10.2023 from  
<https://www.forskningsetikk.no/retningslinjer/generelle/>
- Enggen, A. (2009). Vurdering, evaluering og læreplaner som kunnskapsområde. I S. Dobson, AB Enggen & K. Smith (red.). *Vurdering, prinsipper og praksis. Nye perspektiver på elev-og læringsvurdering*, 75-93.
- Engelsen, B. U. (2012). *Kan læring planlegges? : arbeid med læreplaner - hva, hvordan, hvorfor* (6. utg. ed.). Gyldendal akademisk.
- Fevolden, T., & Lillejord, S. (2005). *Kvalitetsarbeid i skolen*. Universitetsforl.
- Gleiss, M. S., & Sæther, E. (2021). *Forskningsmetode for lærerstudenter : å utvikle ny kunnskap i forskning og praksis* (1. utgave. ed.). Cappelen Damm akademisk.
- Hardy, I. (2014). A logic of appropriation: enacting national testing (NAPLAN) in Australia. *Journal of education policy*, 29(1), 1-18.  
<https://doi.org/10.1080/02680939.2013.782425>
- Longo, C. (2010). Fostering Creativity or Teaching to the Test? Implications of State Testing on the Delivery of Science Instruction. *The Clearing house*, 83(2), 54-57.  
<https://doi.org/10.1080/00098650903505399>
- Marsdal, M. E. (2011). *Kunnskapsbløffen : skoler som jukser, barn som gruer seg*. Manifest. Meld. St. 21. (2016-2017). *Lærelyst - Tidlig innsats og kvalitet i skolen*. Retrieved from  
<https://www.regjeringen.no/no/dokumenter/meld.-st.-21-20162017/id2544344/?ch=1>
- Nusche, D., Earl, L., Maxwell, W., & Shewbridge, C. (2011). *OECD Reviews of Evaluation and Assessment in Education: Norway 2011*. Paris: Organisation for Economic Cooperation and Development (OECD). <https://doi.org/10.1787/9789264117006-en>
- Ottesen, E., & Møller, J. (2011). *Rektor som leder og sjef : om styring, ledelse og kunnskapsutvikling i skolen*. Universitetsforl.
- Roald, K. (2012). *Kvalitetsvurdering som organisasjonslæring : når skole og skoleeigar utviklar kunnskap*. Fagbokforl.
- Ruud, M. (2023). Utvalg foreslår å avvikle dagens nasjonale prøver. *Utdanningsnytt*.  
<https://www.utedanningsnytt.no/eksamen-karakterer-kvalitetsvurderingssystem/utvalg-foreslar-a-avvikle-dagens-nasjonale-prover/380387>
- Seland, I., Vibe, N., & Hovdhaugen, E. (2013). *Evaluering av nasjonale prøver som system*.
- Senge, P. M. (1997). The Fifth Discipline. *Measuring business excellence*, 1(3), 46-51.  
<https://doi.org/10.1108/eb025496>
- Sivesind, K., & Elstad, E. (2010). *PISA : sannheten om skolen?* Universitetsforl.

- Sjøberg, S. (2014). Pisafisering av norsk skole. En suksesshistorie fra OECD. In K. A. Røvik, T. V. Eilertsen, & E. M. Furu (Eds.), *Reformideer i norsk skole : spredning, oversettelse og implementering* (pp. 195-225). Cappelen Damm akademisk.
- Tonich. (2021). The Role of Principals' Leadership Abilities in Improving School Performance through the School Culture. *Journal of Social Studies Education Research*, 12(1), 47-75.
- Tveit, S. (2007). *Elevvurdering i skolen : grunnlag for kulturendring*. Universitetsforl.
- UDIR. (2017). *Kva måler nasjonale prøver i engelsk?* <https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/mestringsbeskrivelser-og-hva-provene-maler/kva-maler-nasjonale-prove-i-engelsk/>
- UDIR. (2018). *Overordnet del – verdier og prinsipper for grunnopplæringen*. <https://www.udir.no/lk20/overordnet-del/>
- UDIR. (2020a). *Core curriculum - values and principles for primary and secondary education*. <https://www.udir.no/lk20/overordnet-del/?lang=eng>
- UDIR. (2020b). *De yngste barna i skolen*. <https://www.udir.no/laring-og-trivsel/tilpasset-opplaring/de-yngste-barna-i-skolen/>
- UDIR. (2022a). *Mestringsbeskrivelser for nasjonale prøver i engelsk*. <https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/mestringsbeskrivelser-og-hva-provene-maler/engelsk-mestringsbeskrivelse/>
- UDIR. (2022b). *Nasjonale prøver*. <https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/>
- UDIR. (2022c). *Rammeverk for nasjonale prøver*. Utdanningsdirektoratet. <https://www.udir.no/eksamen-og-prover/prover/rammeverk-for-nasjonale-prover2/>
- UDIR. (2023a). *Administrere nasjonale prøver*. <https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/administrere-nasjonale-prover2/>
- UDIR. (2023b). *Forberede gjennomføringen av nasjonale prøver*. <https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/forberede-gjennomforingen/>
- UDIR. (2023c). *Kva er nasjonale prøver?* Utdanningsdirektoratet. <https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/om-nasjonale-prover/>
- Volante, L. (2004). Teaching to the test: what every educator and policy-maker should know. *Canadian journal of educational administration and policy*(35).
- Wayman, J. C., & Jimerson, J. B. (2014). Teacher needs for data-related professional learning. *Studies in Educational Evaluation*, 42, 25-34.
- Werler, T., & Færevaaag, M. K. (2017). National testing data in Norwegian classrooms: a tool to improve pupil performance? *Nordic journal of studies in educational policy*, 3(1), 67-81. <https://doi.org/10.1080/20020317.2017.1320188>
- Wille, T. S. (2010). *Vurdering for læring i klasserommet* (2. utg. ed.). Gyldendal akademisk.
- Wu, H., & Leung, S.-O. (2017). Can Likert Scales be Treated as Interval Scales?-A Simulation Study. *Journal of social service research*, 43(4), 527-532. <https://doi.org/10.1080/01488376.2017.1329775>

## Appendix 1, information sheet and consent form

Vil du delta i forskningsprosjektet

### Skolers arbeid med nasjonale prøver i engelsk

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å undersøke i hvor stor grad skoleledere- og engelsklærere bruker resultatene etter nasjonale prøver som et verktøy i utviklingen av engelskfaget. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

#### Formål

Masterprosjektets formål er å undersøke skoleledere- og læreres bruk av resultatene fra nasjonale prøver i skolens engelskfaglige utvikling. Herunder vil vi se på hvordan skolen/lærerne jobber med forberedelse, gjennomføring og tolkning/anvendelse av nasjonale prøver, og resultatene etter dem. Nasjonale prøver skal i følge UDIR brukes som et utviklingsverktøy, men tidligere forskning viser til et sprik når det kommer til den praktiske anvendelsen av resultatene.

Problemstillingen vi ønsker å besvare er «I hvor stor grad bruker et utvalg av skoler i Nord-Norge resultatene fra nasjonale prøver i engelsk som en del av skolens engelskfaglige utvikling». Problemstillingen støttes opp av to forskningsspørsmål:

1. I hvor stor grad innehar engelsklærere og skoleledere de nødvendige kunnskapene og ferdighetene som skal til for å tolke resultatene fra nasjonale prøver i engelsk, slik at de kan anvendes som pedagogiske verktøy i organisering av engelskundervisningen?
2. I hvor stor grad anser engelsklærere og skoleledere nasjonale prøver som et viktig verktøy i elevenes faglige utvikling.

Innsamlet data vil utelukkende brukes til dette prosjektet, og all innsamlet data vil slettes ved prosjektets slutt (14.05.2024).

### **Hvem er ansvarlig for forskningsprosjektet?**

*UiT – Norges Arktiske Universitet* er ansvarlig for prosjektet.

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

På grunnlag av din profesjon og ditt tidligere arbeid med nasjonale prøver anser vi deg som en gunstig informant til vår masteravhandling. Det ble tidligere sendt ut forespørsel til skole- og fagledere ved din skole, som foreslo deg som en mulig kandidat.

### **Hva innebærer det for deg å delta?**

- Hvis du velger å delta i prosjektet, innebærer det at du deltar i et en-til-en-intervju med to forskere. Intervjuet vil ta ca. 20-30 minutter. Underveis i intervjuet vil det bli tatt lydopptak og notater. Lydopptakene vil i ettertid lagres på Nettskjemas sikre servere.
- Det er ikke planlagt å hente inn noe slags personinformasjon ut over kjønn, alder, rolle på skolen og arbeidserfaring. I tillegg til dette vil all data anonymiseres fortløpende slik at det ikke skal være mulig for noen å se hvem som har vært informant når masteravhandlingen publiseres.
- Det vil samles inn data fra flere personer, både skoleledere og lærere fra andre skoler i Nord-Norge.

### **Det er frivillig å delta**

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

Forskningen og de publiserte resultatene fra den vil ikke påvirke ditt forhold til skolen, ledere og/eller andre lærere.

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

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

- De som vil ha tilgang ved behandlingsansvarlig institusjon er to studenter (forskerne) som gjennomfører prosjektet samt veileder.
- Navnet og kontaktopplysninger som samles inn vil erstattes med en kode som lagres på egen navneliste adskilt fra øvrige data. Innsamlet informasjon vil lagres på Nettskjema og Sikts sikre servere.

### **Hva skjer med personopplysningene dine når forskningsprosjektet avsluttes?**

Prosjektet vil etter planen avsluttes når oppgaven blir levert i mai 2024. For å sikre din anonymitet vil personopplysninger anonymiseres underveis i forskningen ved bruk av kodeskjema som lagres adskilt fra øvrig data. Dette skjemaet vil lagres kryptert og

passordbeskyttet. Lydopptak vil lagres på Nettskjemas egne servere, og vil bli slettet etter prosjektets slutt.

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

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra *UiT – Norges Arktiske Universitet* har Sikt – Kunnskapssektorens tjenesteleverandør 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:

- UiT – Norges Arktiske Universitet ved:
  - Ishak Moe Haoual (student) på 91604508
  - Sverre Ness-Voll (student) på 99379101
  - Annelise Brox Larsen (veileder) – 90631173.
- Vårt personvernombud: Annikken Steinbakk – 77646952 -  
personvernombud@uit.no



Hvis du har spørsmål knyttet til vurderingen som er gjort av personverntjenestene fra Sikt, kan du ta kontakt via:

- Epost: [personverntjenester@sikt.no](mailto:personverntjenester@sikt.no)
- Telefon: 73 98 40 40.

Med vennlig hilsen

*Ishak Moe Haoual (forsker), Sverre Ness-Voll (forsker) og Annelise Brox Larsen (veileder)*

(Forsker/veileder)

-----  
-----

**Samtykkeerklæring**

Jeg har mottatt og forstått informasjon om prosjektet ***Skolers arbeid med nasjonale prøver i engelsk***, og har fått anledning til å stille spørsmål.

Jeg samtykker til:

- å delta i intervju med forskerne
- at mine opplysninger behandles frem til prosjektet er avsluttet

---

(Signert av prosjektdeltaker, dato)

## Appendix 2, interview guide

### Intervjuguide

#### Forberedelser

- Gjøre klar intervjuguide og Nettskjemas diktafon-app

#### 1. Tanker om nasjonale prøver

##### Leder

- I hvor stor grad ser du på nasjonale prøver som et viktig utviklingsverktøy for engelskfaget?
- I hvor stor grad mener du at nasjonale prøver er med på å kartlegge elevenes faglige ståsted?

##### Lærer

- I hvor stor grad ser du på nasjonale prøver som et viktig utviklingsverktøy for engelskfaget?
- I hvor stor grad mener du at nasjonale prøver er med på å kartlegge elevenes faglige ståsted?

#### 2. Informasjon og veiledning

##### Leder

- Hva mener du er formålet med nasjonale prøver?
- *Informasjonsflyt fra skoleeier*
- Opplever du at det er tilstrekkelig opplæring for deg som leder med tanke på hvordan du skal veilede lærerne i gjennomføringen av nasjonale prøver?

##### Lærer

- Hva mener du er formålet med nasjonale prøver?
- *Informasjonsflyt*

#### 3. Forberedelse og gjennomføring

##### Leder

- Hvordan jobber skoleledelsen for å forberede lærerne på gjennomføringen av nasjonale prøver?
- Hvor godt informerer skoleledelsen lærerne om føringene fra UDIR med tanke på gjennomføring av nasjonale prøver?
- Settes det av mye tid til forberedelse og gjennomføring av nasjonale prøver?

#### **Lærer**

- Hvordan jobber du som lærer i forberedelsen på gjennomføringen av nasjonale prøver? Opplever du at støtten fra ledelsen er god i forberedelsesfasen?
- Hvor godt blir lærerne informert om føringene fra UDIR med tanke på gjennomføring av nasjonale prøver?
- Settes det av nok tid til forberedelse og gjennomføring av nasjonale prøver?

#### **4. Bruk av resultatene fra prøvene**

##### **Leder**

- Hvordan er/blir du som leder instruert i hvordan resultatene fra NP skal brukes? Hvor får du evt. disse instruksjonene fra?
- Hvordan jobber skolen (ledere og lærere) med tolkning og analyse av resultatene? Hvilke verktøy brukes til dette arbeidet?
- Settes det av mye tid til å bearbeide resultatene? Er dette evt fellestid eller tid som den enkelte lærer eller team jobber disponerer?
- Hvordan implementerer skolen resultatene som en del av utviklingsarbeidet i engelskfaget?
- Har du noen tanker om hvilke forbedringspotensial nasjonale prøver kan ha?
  - o Hvorfor?

##### **Lærer**

- Hvordan er/blir du som lærer instruert i hvordan resultatene fra NP skal brukes? Hvor får du evt disse instruksjonene fra?
- Hvordan jobber skolen (ledere og lærere) med tolkning og analyse av resultatene? Hvilke verktøy brukes til dette?
- Hvordan implementerer skolen resultatene som en del av utviklingsarbeidet i engelskfaget?

- Har du noen tanker om hvilke forbedringspotensial nasjonale prøver kan ha?  
Hvorfor?

## **5. Fritak**

- Dersom elever ikke har mulighet til å komme på skolen på prøvedagen, hvordan jobber dere for at de skal få gjennomført prøven ved en senere anledning? Er dette noe du selv synes er viktig?
- Er du kjent med UDIRs føringer for fritak og gjennomføring av nasjonale prøver?

# Appendix 3, survey sent to teachers



## Kvantitativ undersøkelse, lærer

### Om prosjektet

Prosjektets formål er å undersøke i hvor stor grad et utvalg av skoler i Nord-Norge bruker resultatene fra nasjonale prøver i engelsk som en del av skolens engelskfaglige utvikling. Funnene fra denne spørreundersøkelsen vil brukes som grunnlag for en rekke intervjuer som skal gjennomføres på et senere tidspunkt.

Resultatene som samles inn vil bli behandlet av to forskere, og all personinformasjon vil anonymiseres underveis. Forskingen gjøres i henhold til de etiske retningslinjene gitt av [De Forskningsetiske Komiteene](#).

Eventuelle spørsmål kan rettes til:

Ishak Moe Haoual, epost [iha170@uit.no](mailto:iha170@uit.no)

Sverre Ness-Voll, epost [svo011@uit.no](mailto:svo011@uit.no)

### Samtykke

Jeg samtykker til at mine svar kan brukes i forskningsprosjektet

Ja

Nei

### Holdninger til nasjonale prøver

I denne delen av spørreundersøkelsen vil du få spørsmål som angår dine generelle holdinger til nasjonale prøver (NP)

#### I hvor stor grad ser du på NP som et nyttig utviklingsverktøy for engelskfaget?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært lav

Lav

Verken høy eller lav

Høy

Svært høy

#### I hvor stor grad mener du at NP hjelper deg å kartlegge elevenes faglige ståsted?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært lav

Lav

Verken høy eller lav

Høy

Svært høy

#### I hvor stor grad forklarer du elevene formålet med å gjennomføre NP?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært høy

Høy

Verken høy eller lav

Lav

Svært lav

### Forberedelse til nasjonale prøver

I denne delen av spørreundersøkelsen vil du få spørsmål som angår hvordan du og skolen din arbeider

med forberedelsene som gjøres før nasjonale prøver (NP) gjennomføres

### Hvor godt kjent er du med veiledningsmateriellet som er utarbeidet av UDIR?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært godt
- Godt
- Verken godt eller lite
- Ikke godt
- Overhodet ikke godt

### Hvor mye tid settes av i forberedelsene til NP?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite tid
- Lite tid
- Verken lite eller mye tid
- Mye tid
- Svært mye tid

### Bruker du enkeltoppgaver fra tidligere nasjonale prøver i forberedelsesfasen?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært ofte
- Ofte
- Verken ofte eller lite
- Lite
- Svært lite

### Bruker du fullstendige oppgavesett fra tidligere nasjonale prøver i forberedelsesfasen?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite
- Lite
- Verken lite eller ofte
- Ofte
- Svært ofte

### Hvor godt kjent er du med ansvar for gjennomføring av NP? (ansvarsfordeling mellom skoleeier-, leder og lærer)

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært godt
- Godt
- Verken godt eller lite
- Lite
- Svært lite

### Gjennomføringsfasen

I denne delen av spørreundersøkelsen får du spørsmål som angår hvordan skolen din og du gjennomfører nasjonale prøver (NP)

**I hvor stor grad hjelper du elevene underveis?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite
- Lite
- Verken lite eller høyt
- Høy
- Svært høy

**I hvor stor grad får elever med fravær på prøvedagen mulighet til å gjennomføre prøven ved en senere anledning?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite
- Lite
- Verken lite eller høyt
- Høyt
- Svært høyt

**Evalueringsfasen**

I denne delen av spørreundersøkelsen vil du få spørsmål som angår hvordan du og skolen din jobber med å evaluere, tolke og forstå resultatene fra nasjonale prøver (NP)

**I hvor stor grad forstår du resultatene fra NP?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite
- Lite
- Verken lite eller høyt
- Høy
- Svært høy

**I hvor stor grad føler du deg rustet til å tolke resultatene fra NP?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært høy
- Høy
- Verken høyt eller lite
- Lite
- Svært lite

**I hvor stor grad kan du bruke resultatene fra NP for videre tilrettelegging?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært høy
- Høy
- Verken høyt eller lite
- Lite
- Svært lite

**Hvor godt har du satt deg inn i UDIRs veileder "Nasjonale prøver"?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*



Svært godt  
Godt  
Verken godt eller ikke godt  
Ikke godt  
Overhodet ikke godt

**Etter nasjonale prøver**

I denne delen av spørreundersøkelsen får du spørsmål som angår din forståelse av- og dine tanker om nasjonale prøver som utviklingsverktøy

**I hvor stor grad forstår du resultatene fra NP?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært lite  
Lite  
Verken lite eller høyt  
Høyt  
Svært høyt

**I hvor stor grad mener du at skolen din bruker resultatene som en del av utviklingen av engelskfaget?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært godt  
Godt  
Verken godt eller lite  
Lite  
Svært lite

**I hvor stor grad ser du på NP som et nyttig utviklingsverktøy for engelskfaget?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært lite  
Lite  
Verken lite eller godt  
Godt  
Svært godt



# Appendix 4, survey sent to school leaders



## Kvantitativ undersøkelse, skoleleder

### Om prosjektet

Prosjektets formål er å undersøke i hvor stor grad et utvalg av skoler i Nord-Norge bruker resultatene fra nasjonale prøver i engelsk som en del av skolens engelskfaglige utvikling. Funnene fra denne spørreundersøkelsen vil brukes som grunnlag for en rekke intervjuer som skal gjennomføres på et senere tidspunkt.

Resultatene som samles inn vil bli behandlet av to forskere, og all personinformasjon vil anonymiseres underveis. Forskingen gjøres i henhold til de etiske retningslinjene gitt av [De Forskningsetiske Komiteene](#).

Eventuelle spørsmål kan rettes til:

Ishak Moe Haoual, epost [iha170@uit.no](mailto:iha170@uit.no)

Sverre Ness-Voll, epost [svo011@uit.no](mailto:svo011@uit.no)

### Samtykke

Jeg samtykker til at mine svar kan brukes i forskningsprosjektet

Ja

Nei

### Holdninger til nasjonale prøver

I denne delen av spørreundersøkelsen vil du få spørsmål som angår dine generelle holdninger til nasjonale prøver (NP)

#### I hvor stor grad ser du på NP som et nyttig utviklingsverktøy for engelskfaget?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært lav

Lav

Verken høy eller lav

Høy

Svært høy

#### I hvor stor grad mener du at NP hjelper deg å kartlegge elevenes faglige ståsted?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært lav

Lav

Verken høy eller lav

Høy

Svært høy

#### I hvor stor grad forklarer du elevene formålet med å gjennomføre NP?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært høy

Høy

Verken høy eller lav

Lav

Svært lav

### Forberedelse til nasjonale prøver

I denne delen av spørreundersøkelsen vil du få spørsmål som angår hvordan du og skolen din arbeider

med forberedelsene som gjøres før nasjonale prøver (NP) gjennomføres

### Hvor godt kjent er du med veiledningsmateriellet som er utarbeidet av UDIR?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært godt
- Godt
- Verken godt eller lite
- Ikke godt
- Overhodet ikke godt

### Hvor mye tid settes av i forberedelsene til NP?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite tid
- Lite tid
- Verken lite eller mye tid
- Mye tid
- Svært mye tid

### Bruker du enkeltoppgaver fra tidligere nasjonale prøver i forberedelsesfasen?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært ofte
- Ofte
- Verken ofte eller lite
- Lite
- Svært lite

### Bruker du fullstendige oppgavesett fra tidligere nasjonale prøver i forberedelsesfasen?

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite
- Lite
- Verken lite eller ofte
- Ofte
- Svært ofte

### Hvor godt kjent er du med ansvar for gjennomføring av NP? (ansvarsfordeling mellom skoleeier-, leder og lærer)

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært godt
- Godt
- Verken godt eller lite
- Lite
- Svært lite

### Gjennomføringsfasen

I denne delen av spørreundersøkelsen får du spørsmål som angår hvordan skolen din og du gjennomfører nasjonale prøver (NP)

**I hvor stor grad hjelper du elevene underveis?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite
- Lite
- Verken lite eller høyt
- Høy
- Svært høy

**I hvor stor grad får elever med fravær på prøvedagen mulighet til å gjennomføre prøven ved en senere anledning?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite
- Lite
- Verken lite eller høyt
- Høyt
- Svært høyt

**Evalueringsfasen**

I denne delen av spørreundersøkelsen vil du få spørsmål som angår hvordan du og skolen din jobber med å evaluere, tolke og forstå resultatene fra nasjonale prøver (NP)

**I hvor stor grad forstår du resultatene fra NP?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært lite
- Lite
- Verken lite eller høyt
- Høy
- Svært høy

**I hvor stor grad føler du deg rustet til å tolke resultatene fra NP?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært høy
- Høy
- Verken høyt eller lite
- Lite
- Svært lite

**I hvor stor grad kan du bruke resultatene fra NP for videre tilrettelegging?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

- Svært høy
- Høy
- Verken høyt eller lite
- Lite
- Svært lite

**Hvor godt har du satt deg inn i UDIRs veileder "Nasjonale prøver"?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært godt  
Godt  
Verken godt eller ikke godt  
Ikke godt  
Overhodet ikke godt

**Etter nasjonale prøver**

I denne delen av spørreundersøkelsen får du spørsmål som angår din forståelse av- og dine tanker om nasjonale prøver som utviklingsverktøy

**I hvor stor grad forstår du resultatene fra NP?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært lite  
Lite  
Verken lite eller høyt  
Høyt  
Svært høyt

**I hvor stor grad mener du at skolen din bruker resultatene som en del av utviklingen av engelskfaget?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

Svært godt  
Godt  
Verken godt eller lite  
Lite  
Svært lite

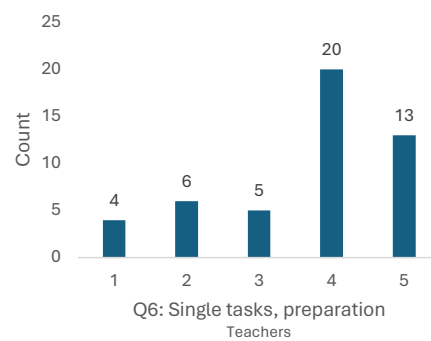
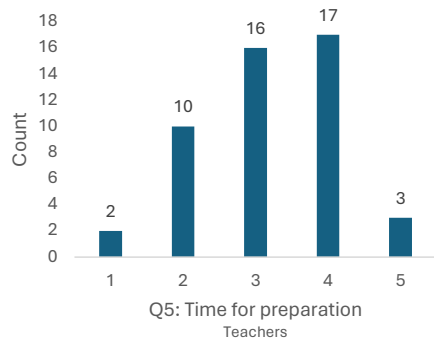
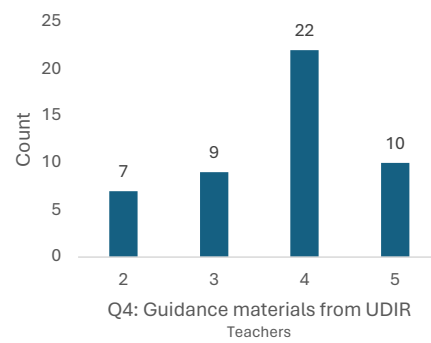
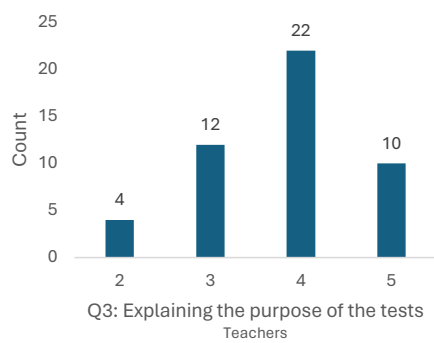
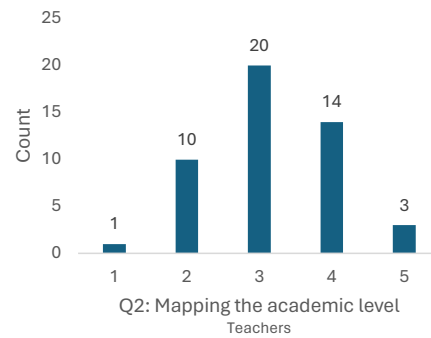
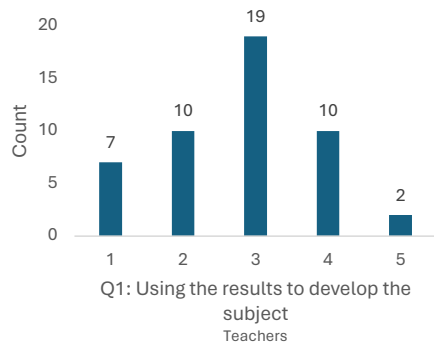
**I hvor stor grad ser du på NP som et nyttig utviklingsverktøy for engelskfaget?**

*Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «Samtykke»*

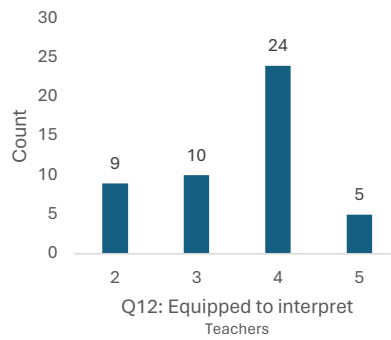
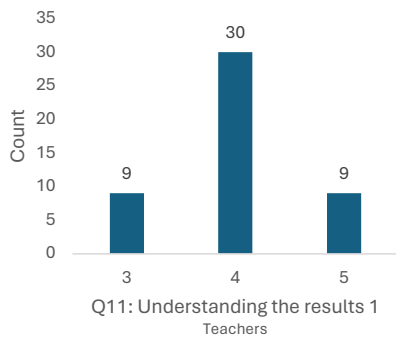
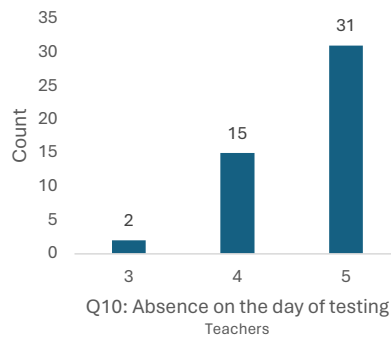
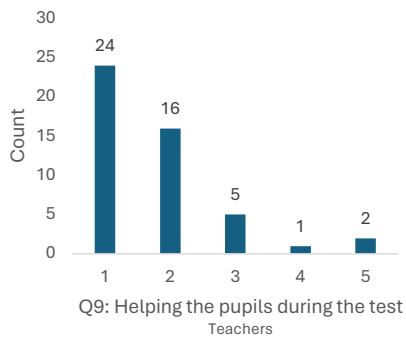
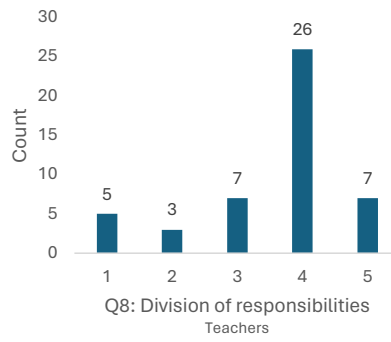
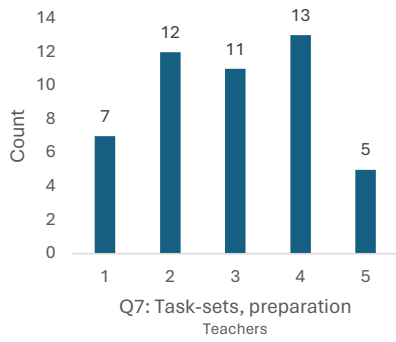
Svært lite  
Lite  
Verken lite eller godt  
Godt  
Svært godt

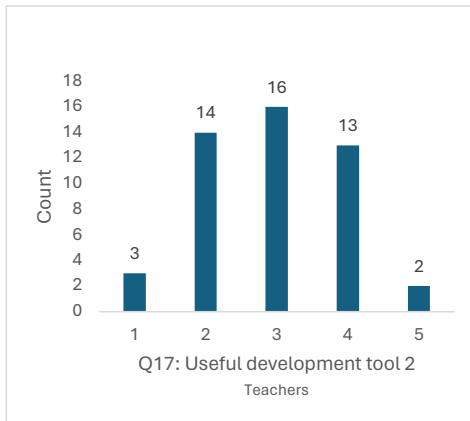
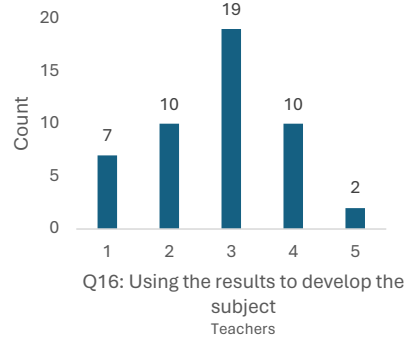
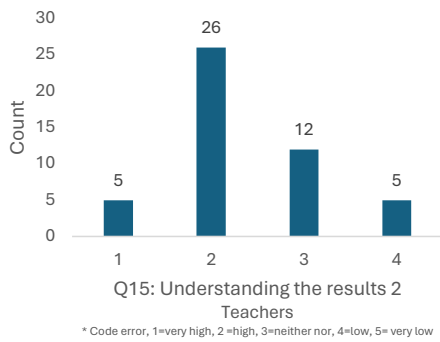
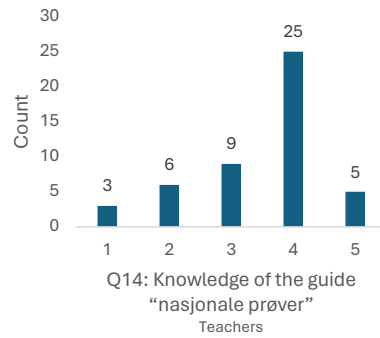
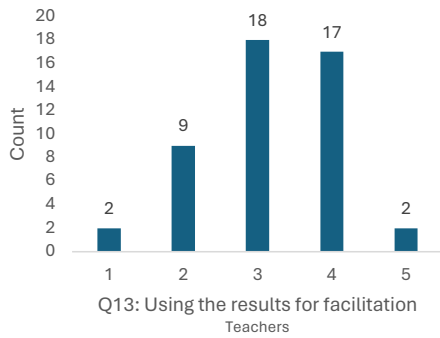


## Appendix 5, individual graphs, teacher survey









# Appendix 6, individual graphs, school leader survey

