OECD Education Working Papers No. 97

Balancing Trust and Accountability? The Assessment for Learning Programme in Norway: A Governing Complex Education Systems Case Study

Therese Hopfenbeck,

Astrid Tolo,

Teresa Florez,

Yasmine El Masri

https://dx.doi.org/10.1787/5k3txnpqlsnn-en

OECD

Unclassified **EDU/WKP(2013)12**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **13-Dec-2013**

English - Or. English

DIRECTORATE FOR EDUCATION AND SKILLS

EDU/WKP(2013)12
Unclassified

BALANCING TRUST AND ACCOUNTABILITY? THE ASSESSMENT FOR LEARNING PROGRAMME IN NORWAY

A Governing Complex Education Systems Case Study

Working Paper No. 97

Therese Hopfenbeck, Astrid Tolo, Teresa Florez, Yasmine El Masri

Therese Hopfenbeck, Oxford University Centre for Educational Assessment,
therese.hopfenbeck@education.ox.ac.uk
Tracey Burns, CERI, OECD, Tracey.Burns@oecd.org

JT03350436

Complete document available on OLIS in its original format
This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

English - Or. English

OECD DIRECTORATE FOR EDUCATION

OECD EDUCATION WORKING PAPERS SERIES

This series is designed to make available to a wider readership selected studies drawing on the work of the OECD Directorate for Education. Authorship is usually collective, but principal writers are named. The papers are generally available only in their original language (English or French) with a short summary available in the other

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The opinions expressed in these papers are the sole responsibility of the author(s) and do not necessarily reflect those of the OECD or of the governments of its member countries.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org.

Comments on the series are welcome, and may be sent to edu.contact@oecd.org.

---

www.oecd.org/edu/workingpapers

Copyright © OECD 2013

ABSTRACT

This report explores the development of implementation strategies used to enhance the programme "Assessment for Learning – 2010-2014" in Norwegian schools. Norway's educational governance is highly decentralised, with 428 municipalities and 19 counties responsible for implementing education activities, organising and operating school services, allocating resources and ensuring quality improvement and development of their schools. This case study is based on56 interviews with 98 key actors and stakeholders in the Norwegian education system, as well as analysis of key policy and legal documents and a range of media articles. Key findings include the importance of clear communication between governance levels and a high degree of trust between stakeholders; the need for a clear understanding of programme goals, the role of learning networks between schools to aid the exchange of knowledge and provide peer support during the implementation process. Innovative forms of capacity building were of particular importance for the smaller municipalities, who reported being overextended by the continual stream of policy changes and struggling with prioritising activities. The case study also provides a series of recommendations for improvement.

RÉSUMÉ

Ce rapport examine le développement de stratégies de mise en œuvre utilisées pour l'amélioration du programme « Assessment for Learning – 2010-2014 » dans les établissements scolaires norvégiens. La gouvernance éducative de la Norvège est fortement décentralisée, comptant 428 municipalités et 19 provinces en charge de mettre en œuvre les activités éducatives, d'organiser et de faire fonctionner les services scolaires, d'allouer les ressources, d'améliorer la qualité des établissements et d'assurer leur développement. Cette étude de cas s'appuie sur 56 entretiens avec 98 acteurs-clé et parties prenantes du système éducatif norvégien, ainsi que sur l'analyse de documents-clé politiques et légaux, et d'un ensemble d'articles médias. Les principaux résultats soulignent : l'importance d'une communication claire entre les différents niveaux de gouvernance et un haut niveau de confiance entre les différentes parties ; la nécessité d'une compréhension nette des objectifs du programme, du rôle des liens d'apprentissage entre les établissements pour faciliter l'échange de savoir et fournir un soutien par les pairs pendant le processus de mise en œuvre. Des formes innovantes de renforcement des capacités se sont montrées d'importance majeure dans les municipalités de petite taille, lesquelles ont déclaré avoir été dépassées par les fluctuations incessantes dans les politiques à mettre en œuvre, et avoir eu du mal à décider des activités auxquelles donner priorité. Cette étude de cas conclut par une série de recommandations pour l'amélioration du programme.

FOREWORD

In 2011, the OECD launched the project "Governing Complex Education Systems" (GCES) to better understand the increasing challenges in steering the implementation of education policies. GCES aims to establish the state of research and the evidence base in the areas of governance, complexity and knowledge, and to explore current practices in OECD member countries. The country case study of Norway is part of this work. It describes and evaluates the implementation of a specific national programme, *Vurdering for Læring* or "Assessment for Learning" which was initiated by the Norwegian government in 2010.

The purpose of Assessment for Learning is to enhance assessment practices in Norwegian schools (years 1-12). The country case study analyses the views and experiences of the Ministry of Education and Research, the Directorate for Education and Training, supervisors in the project, heads of municipalities, policy makers, teacher unions, teachers, parents, and students in the implementation of Assessment for Learning.

The findings of case study underline the importance of good communication and trust between the various stakeholders and government levels to succeed in the implementation of the programme. The case study also points to the importance of municipalities and school owners fully understanding the goals of the programme and ensuring that capacity building and leadership are given the necessary attention. Moreover, the report shows the various ways the programme has been implemented on the local level and the importance of providing sufficient time to fully implement new practices. I am convinced that these findings will be useful for the implementation of other programmes, too.

I find it very valuable to have the voices of so many different stakeholders in one single report and am grateful to all for taking the time to share their experiences with the researchers. We will learn from it. The Assessment for Learning programme is at the core of our on-going efforts to improve our education system.

I also recognize the valuable work done by the OECD to bring together stakeholders and present evidence from various countries so that we can learn from each other. Context, systems, history and culture vary but we share the common goal of providing our students with the best education possible.

Oslo, September 2013

Kristin Halvorsen
Minister of Education and Research

ACKNOWLEDGEMENTS

This case study aims to look at governance of the implementation of the new Assessment for Learning programme in Norway. In order to do this the research team conducted 56 interviews with almost 100 participants. A wide range of national, regional and local authorities was consulted, as well as members of the former Assessment Reform Group in the UK. The authors would like to express grateful appreciation to the many stakeholders who kindly spared their time to work with the interview team from May to October 2012. School leaders, teachers and researchers who made themselves available again after a strike in Norway in May 2012 are particularly thanked, as well as all the students who spent their time explaining how they felt about being assessed in schools. After all, students are the ones who are most affected by the policy implementations in schools.

The authors acknowledge the commitment stakeholders have given to this programme and also appreciate their willingness to share their experiences. The report is the responsibility of the research team, and although it has been informed by a number of important stakeholders, the authors assume full responsibility for any errors or misinterpretations.

The research team would like to extend special thanks to Professor Sølvi Lillejord, Director of The Norwegian Knowledge Centre for Education, for valuable feedback on drafts of this report, particularly in relation to the Norwegian context. Special thanks to Professor Richard Daugherty, Professor John Gardner, Professor Jenny Ozga and Professor Jo-Anne Baird for their valuable feedback, discussions and comments on this report. The Directorate of Teaching and Training in Norway has been very helpful in providing data for national tests for the 2010 – 2012 period, and contributing documents for analysis. Thanks to Daniel Caro for help and guidance with the Norwegian national test data, Jo Hazell for proofreading and Natalie Usher for support with data cleaning, proofreading and editing. The research team in Oxford are grateful for the support from the OECD team Tracey Burns and Harald Wilkoszewski, and Eli Sundby and Annette Skalde from the Norwegian Ministry of Education and Research.

This report was authored by Therese Hopfenbeck, Astrid Tolo, Teresa Florez and Yasmine El Masri of Oxford University and the University of Bergen. Within the OECD Centre for Research and Educational Innovation (CERI) the preparation of this report was coordinated by Tracey Burns and Harald Wilkoszewski, with editorial support from Anna Barnet.

TABLE OF CONTENTS

OECD DIRECTORATE FOR EDUCATION 2

ABSTRACT ... 3
FOREWORD ... 4
ACKNOWLEDGEMENTS ... 5
ACRONYMS AND ABBREVIATIONS ... 8
EXECUTIVE SUMMARY ... 9
Introduction ... 9
Key findings ... 9
Key recommendations ... 10
CHAPTER 1. GOVERNING COMPLEX EDUCATION SYSTEMS ... 11
Introduction ... 11
Research questions for this report ... 11
Methodology ... 12
Some key terms ... 14
CHAPTER 2. EDUCATION IN NORWAY ... 17
Demographic and social context ... 17
Economic context ... 17
Governance actors and different stakeholders ... 18
Cultural, political and historical factors in Norway ... 20
Assessment practices before and after PISA ... 22
The Norwegian quality system ... 23
CHAPTER 3. THE CASE STUDY: ASSESSMENT FOR LEARNING (AFL) ... 25
Description of the policy programme and implementation strategies ... 25
Assessment for learning as policy: International evidence ... 29
Impact of the Assessment for Learning (AfL) programme in Norway ... 32
CHAPTER 4. THE AFL PROGRAMME IN NORWAY – A SUCCESS STORY WITH MAJOR CHALLENGES? ... 35
Drivers of change: Governing from the central to the local level ... 35
Implementing AfL in schools: How successful has the programme been? ... 40
Perceptions of key stakeholders on the AFL programme ... 50
Impact of the AfL programme: Students' learning outcomes ... 57
CHAPTER 5. CONCLUSION AND RECOMMENDATIONS ... 60
Expectations of the programme – What next? ... 60
Balancing accountability and trust ... 61
Dealing with "reform fatigue" ... 62
The gap from theory to practice ... 62
Change takes time ... 64
NOTES ... 65
REFERENCES ... 67
ANNEX A. METHODOLOGICAL NOTES ... 77
Coding of interviews ... 78
National test scores ... 79
ANNEX B. LIST OF INTERVIEWEES ... 80

ANNEX C. INTERVIEW GUIDE 81
Interview schedule 81
Interview guide: Policy makers and researchers 81
Interview guide: School leaders 82
Interview guide: Teachers 83
Interview guide: Students (approximately age 14) 84
ANNEX D. WEBSITES FOR THE MEDIA ANALYSIS: EXAMPLES FROM DIFFERENT CATEGORIES 86
1) Information about AfL 86
2) Schools working on AfL 86
3) Problems in relation to assessment 86
4) Think tanks 88
RECENT OECD PUBLICATIONS OF RELEVANCE TO THIS WORKING PAPER 89
THE OECD EDUCATION WORKING PAPERS SERIES ON LINE 90

Tables

Table 1.1. Participating schools and their status in this programme 12
Table 1.2. Participants, different stakeholders 13
Table 3.1. Participating groups and timeline 26
Table 4.1. Ministers of education from 2000 in Norway 36
Table 4.2. Implementation of the AfL programme in nine schools 41
Table 4.3. Results from national tests: Reading Grade 5 57
Table 4.4. Results from national tests: Reading Grade 9 58
Table 4.5. Results from national tests: Mathematics Grade 5 58
Table 4.6. Results from national tests: Mathematics Grade 9 58
Table 4.7. Results from national tests: English Grade 5 58
Table 4.8. Results from national tests: English Grade 8 58

Figures

Figure 2.1. Levels of governance and different stakeholders in Norway 18

Boxes

Box 1.1. Questions regarding knowledge and capacity building in this case study 16
Box 2.1. Student organisation in Norway 19
Box 3.1. Smoothing the implementation of the AfL programme 25
Box 3.2. Norwegian research programme on assessment 34
Box 4.1. Evaluation of Norwegian municipalities 39
Box 4.2. The changing role of teachers 40
Box 4.3. The website Vurdering for læring: A tool for implementing new assessment practices 42

ACRONYMS AND ABBREVIATIONS

AfL	Assessment for Learning
ARG	Assessment Reform Group
CERI	Centre for Educational Research and Innovation
GCES	Governing Complex Education Systems
EO	Elevorganisasjonen (Norwegian Student Organization)
FUG	Foreldreutvalget for grunnopplæringen i Norge. (National Parents' Committee for Primary and Secondary Education).
DET	Norwegian Directorate of Education and Training (Utdanningsdirektoratet)
OECD	Organisation for Economic Co-Operation and Development
VfL	*Vurdering for Læring* – the Norwegian translation of Assessment for Learning (AfL)
LK06	The National Curriculum for Knowledge Promotion 2006 (Læreplanen Kunnskapsløftet)
PISA	Programme for International Student Assessment

EXECUTIVE SUMMARY

Introduction

With the increasing complexity of educational systems, countries face challenges in governing the implementation of educational reforms and programmes. The present country case study explores the development of implementation strategies used to enhance the programme "Assessment for Learning – 2010-14" in Norwegian schools as a part of the OECD's Governing Complex Education Systems (GCES).

Norway's educational governance is highly decentralised, with 428 municipalities and 19 counties appointed as "school owners." The municipalities and counties vary greatly in size, number of schools and competence at the municipal level. The Ministry of Education and Research formulates national education policy including acts, regulations and curricula. Within this framework, the school owners (counties, municipalities and private providers) are responsible for implementing education activities, organising and operating school services, allocating resources, ensuring quality improvement and development of their schools.

The case study is organized in four parts. Chapter 1 introduces the GCES framework and presents the background for the study while Chapter 2 describes the Norwegian school system within the context of cultural, historical, and socio-economic factors. Chapter 3 examines the Norwegian programme *Vurdering for Læring* (Assessment for Learning) that is the focus of this case study. Chapter 4 presents the main results including policy recommendations. Details about the methodology, participants and analytical tools used in the present study can be found in the Appendices.

Key findings

A substantive body of empirical data was collected for this report. In total, 56 interviews with 98 key actors and stakeholders in the Norwegian education system were conducted. In addition, key policy and legal documents as well as a range of media articles were analysed. From the results of these analyses the following key findings can be drawn:

- The municipalities that successfully implemented the programme demonstrated clear communication between governance levels and a high degree of trust between stakeholders.

- Municipal and school leaders who based their implementation strategies on a clear understanding of the programme goals and who could integrate these goals within the broader aims of educational policy and school practice were more successful in their implementation.

- The establishment of learning networks between schools aided the exchange of knowledge and provided peer support in the implementation process.

- Innovative forms of capacity building were of particular importance for the smaller municipalities, who reported being overextended by the continual stream of policy changes and struggling with prioritising activities. Within the framework of the Assessment for Learning

programme, a set of tools was developed that included an online platform where teachers and school leaders could access information on best practices. Smaller municipalities in particular reported that this tool helped them to implement the programme goals.

Key recommendations

Facilitate communication and build trust between the various levels in the system

- Facilitate a continuous exchange between the various actors to create a forum for feedback and learning, which will contribute to greater trust in the policy programme
- Focus particularly on the learner: students' feedback on the implementation processes can foster change and be a valuable classroom reality check
- Communicate clear goals and expectations for the policy programme

Assess programme priorities in the context of other policy reforms

- Provide guidance on priority setting to smaller municipalities and others that may be overwhelmed by political reforms
- Establish networks between outstanding teachers, school leaders, schools and municipalities and those that struggle with change, in order to overcome implementation problems

Build capacity on a large scale

- In order to facilitate the scaling-up of a policy programme, enlarge the capacity building element of the reform to facilitate its implementation
- Following the example of the Norwegian national centres for reading, writing, science, language, and mathematics, establish a national centre for assessment to support teachers, school leaders, and schools with expertise on the design and use of assessment models.

Develop a sustainable implementation strategy

- Ensure that the necessary sources of capacity building are provided beyond the nominal time frame of the policy programme
- Empower school leaders to facilitate sustainable change and provide them with the necessary knowledge on the change process, underlying assumptions, and research evidence, as well as potential pitfalls.

CHAPTER 1.

GOVERNING COMPLEX EDUCATION SYSTEMS

Introduction

Over the last decade a number of reports and articles describing the challenges of implementing educational reforms and programmes and steering the education system in Norway have been published (Engelund and Langfeldt, 2009; Allerup et al., 2009; Sandberg and Aasen, 2008; OECD, 2011; Aasen et al., 2012; Sivesind, 2012; Olsen et al., 2012). A recent report from the OECD (2011) argued that Norway lacks "a clearly defined implementation strategy for education reforms throughout different levels, including municipalities and schools" (ibid., p. 12). The report made the following recommendations (OECD, 2011):

1. define and develop clear implementation strategies,
2. reinforce the role and capacity of policy makers at the different levels of the system and;
3. build a culture of evidence using data strategically.

Norway has taken a number of steps to address these recommendations. This case study is one such example. Conducted as part of the OECD Governing Complex Education Systems (GCES) project, this case study takes an in-depth look at the policy programme "Assessment for Learning – 2010-2014" and its implementation in Norwegian schools. The programme was developed by the Directorate of Education and Training (DET) with a heavy emphasis on participation and dialogue, a strategy which can be seen as part of the Norwegian philosophy where all participants need to feel a sense of ownership of the approach to implementation for this to work in practice.

Education systems have never been easy to manage. Nonetheless, there seems to be wide consensus in the literature and among policy makers that complexity is increasing as systems become more decentralised. In this context, processes of reform cannot be understood as top-down unilateral delivery chains and treated as systems engineering processes, but require new structures of collaboration, participation and networking. This has made governance a central question for education and education systems more generally.

Research questions for this report

The Assessment for Learning (AfL) programme is a national programme initiated by the Norwegian Directorate of Education and Training on behalf of the Ministry of Education that aims to improve formative assessment practices in the classroom. As set forth by the Framework for Case Studies of the GCES project, the research questions that were the basis for this report are as follows:

- What were the expected **results** of the AfL programme? How do they compare to those actually achieved? How is the programme designed and organised? Are the expectations clearly stated in the design?
- To whom were the **results communicated**, how and when? Were the underlying concepts clearly stated and easy to understand? How many actors communicated at how many levels? Was the communication distorted along the way? Why and how?
- What were the **perceptions** of different stakeholders with regard to the process and the final setup of the policy programme, its implementation and its outcomes? How were the programme ideas acted upon? Were local activities consistent with the intended design and organisation? Were there "undesired" outcomes, and if so, for whom?
- Was there a **redefinition of goals**? If so, to what extent was it based on undesired outcomes of the policy programme?
- Did the results of the programme have an **impact** on central or local education policy?
- Was there a **follow-up** to the policy programme? If so, to what extent did it differ from the original policy?

Methodology

To inform the analysis, research studies on evaluation, assessment, reforms, implementation and policy change in Norway from 1980 to 2012 were included. For the Assessment for Learning Program, official and unofficial notes and letters from the Norwegian Ministry of Education and Research were given to the research team in June 2012. These documents have mainly served to inform the interviews that were carried out, and to clarify how the implementation process was conducted from the central level. As some of the documents are not official documents, no references have been made to these in the present report. In addition, articles from the AfL literature that were suggested by participants were included – articles that they relied upon when implementing AfL. Finally, empirical research articles from Norway, describing AfL in a Norwegian context, are included. Students' test scores on national tests in reading, mathematics, and English have been compared from year 2010 to 2012, for both schools in the AfL programme and schools outside the intervention program. For a full description of the methodology used, including participants, coding and interview guide, please see Annex A.

Semi-structured interviews and participants

Interviews were conducted in Norway from May through September 2012, and the participants were sampled to give views from (1) schools in municipalities that were not part of the programme (2) schools in municipalities that had been part of the implementation since 2010 or had been a part of the Better Assessment for Learning program earlier, and (3) schools in municipalities that would join the implementation programme in the fall of 2012. As one of the main purposes of the interview study was to investigate how the programme had been implemented, some of the participating schools are compared with schools that had just started the programme. Also investigated were schools that had been selected by the authorities at the start because they were known to be highly valued by stakeholders.

Table 1.1. Participating schools and their status in this programme

ID	Schools	Number of students	Location	Relation to AfL	Municipality
1	Primary school	350	Urban	Not official member	A
2	Primary and Secondary	650	Urban	Not official member	A
3	Primary	250	Rural	Participated from 2012	B
4	Secondary	250	Rural	Participated from 2008	B

5	Primary	200	Rural	Participate from 2011	C
6	Secondary	350	Urban	Participated from 2010	C
7	Primary	230	Rural	Participated from 2012	D
8	Secondary	150	Rural	Participated from 2012	D
9	Primary	500	Urban	Participated from 2012	D

In addition, researchers and policy makers who had been active in advocating the implementation were selected for interviews. Due to a strike involving teachers, researchers and leaders on the municipality level in Norway in May 2012, some of the original interviews were either postponed or cancelled. During the first interviews, new names of possible interview candidates were identified in a snowball sampling approach. To protect the anonymity of participants, little information was given about their specific duties. Anonymity fostered a more open discussion regarding problems. Only one interview candidate did not respond to our invitation to participate, and a second candidate could not participate due to annual leave. Overall, there was an extremely high positive response rate in this project (56 out of 58 interviews scheduled).

Table 1.2. Participants, different stakeholders

Type of organisation	Information about participants	N=people
Ministry of Education and Research	The Minister of Education and Research	3
Directorate of Education And Training	Leading	4
Deputy-Director	OECD	1
University and University Colleges	Professors, different Universities in Norway	5
Teacher Unions	One national leader	1
Parents	Two couples	4
Municipality leaders	Four different counties	5
School leaders	Nine different schools	9
Teachers	Primary and Secondary teachers	21
Pupils	Grade 5, 6, 8 and 9	49
UK researchers	Members from the Assessment Reform Group	4
Journalist	Writer and Think Tank member	1
Total number of interviewees		98

Interviewees from the Directorate were all working on the implementation of the AfL programme in different positions and roles. The school researchers in higher education institutions were chosen since they had a particular focus on how to support schools in the implementation of AfL, and some of them had been actively involved in the supervision of some of the participating schools. The participants came from three counties and from four different municipalities in Norway. Municipality leaders delegated most of the responsibility to a project leader, and did not personally work on the implementation. For this reason we have not included the municipality leaders in our analysis in the result chapter. The municipalities represent different school systems; some have a school office that leads the schools in the county, while other schools have been delegated all the power, and are in charge of budget, steering and school development. This is currently known as the difference between three-level steering and two-level steering in Norway.

Analysis of national test scores from participating schools in the AfL programme

According to the AfL literature, AfL can increase students' learning outcomes (Black and Wiliam, 1998, Hattie 2009). In the present study, national test scores from schools participating in the AfL 2010 – 2012 were analysed to see whether there were any changes in the scores (English reading, Reading, and Mathematics from the years 2010, 2011, and 2012). Students in Norway are tested in year 5, 8 and 9, which matches the primary and secondary schools participating in this case study. A list of approximately 145 primary schools and 80 secondary schools participating in the AfL programme in 2010-2012 was received from the Norwegian Directorate for Education and Training in January 2013, together with the

data files of national test scores. The test scores from these schools were compared to the rest of the schools in Norway, consisting of approximately 1450 primary schools and 820 secondary schools.

Media analysis

For this case study a media search was conducted in some online newspapers using the search term *Vurdering for læring*. (See Annex D for overview). Also included were articles from Norwegian think tank websites, such as Civita and Manifest Analyse, which were selected because they represent two of the most active think tanks in Norway and two different political and philosophical perspectives. The Managing Director at Civita is the previous Minister of Education, Kristin Clemet, from the Conservative Party. Civita promotes "the meaning of freedom, personal responsibility and a free market economy". Manifest Analyse is led by a previous leader from the Socialist Youth Party, Julie Lødrup, and is supported by organisations associated more with the left wing compared to Civita. Further searches were conducted in journals that specifically targeted teachers and school leaders, such as Utdanning, Skolelederen and Bedre Skole.

Methodological limitations

A sample of nine schools from different areas in Norway was used to gather data for this case study. As such, it is not a representative sample and the findings herein cannot be simply generalized to all schools in Norway. However, the present analysis offers in-depth findings that can shed light upon implementation strategies that might be of broad interest across the country and indeed, outside of Norway. Though based upon self-reports (interviews) and document analysis (selected by the government), this study nonetheless involved 98 participants from 56 interviews. It is a therefore a substantial body of data that can be used for future secondary analysis. In addition, the participants interviewed are leaders, at all levels from the central to the local, within Norway's political system. Further, the scores from the national tests give some information about the participating schools, even though there are some limitations to the sample we have used, which is further explained in Appendix A.

Another limitation of the study is the lack of observation data from inside the classroom. The analysis of the interviews thus relies heavily on self-reports of school leaders, teachers and students. Despite these limitations, school leaders, teachers and students give the same message independently, which suggests that the findings reported in this work are robust.

Some key terms

Governance

It is important to stress that not all authors who are working on issues of complexity in modern educational systems agree about the roots of the concept of governance. For some authors, governance is part of a natural process to which countries are moving as part of a more globalised and decentralised world, where knowledge and values on education can be shared with the aim of improving education everywhere. For example: "All over the world, governments, universities, school systems and various other parties are looking at new ways to find, share, understand and apply the knowledge emerging from research, leading to increasing conceptual and empirical work to understand how this can be done" (Levin, 2011: 15).

This new structure of governance rather than government is seen as providing more autonomy to schools and local authorities and as allowing for more diversity between schools (Resh, 2009), and also as a means to "increase efficiency, encourage innovation, and combat social inequality and segregation in

education" (Hooge et al., 2012: 6). Again, all these changes are attributed to some naturally emerging trends in education management. In summary: "Governance refers to the process of governing societies in a situation where no single actor can claim absolute dominance" (Fazekas and Burns, 2012: 7).

However, several authors think it is important to highlight the origins of this new landscape, as they respond to very specific economic and political agendas and not simply to a natural process of decentralisation and globalisation. These authors refer to a historical change in developed and developing societies that were governed by a stronger and more directly responsible state where the limits between private and public are clear, which also involved more hierarchies and bureaucracy. After the introduction of neo-liberal agendas and market-centred models of society, this way of government changed to what has been called interchangeably soft governance, weak governance, network governance, the post-bureaucratic state, and real governance, among others (see for example Alexadiou et al., 2010; Al-Samarai, 2009; Ball and Junneman, 2011; Titeca, 2011). According to Ball and Junneman, quoting Peterson, there is "a shift from government to governance, (...) [that] involves the development of relations involving mutuality and interdependence as opposed to hierarchy and independence" (Peterson, 2003: 1).

To understand and analyse reforms and education programmes, researchers need to look at education as a profoundly national endeavour, which is increasingly met with new transnational governing structures. In other words, to review what is going on in educational programmes in specific countries, we need to investigate the national context where the programme is situated.

Knowledge

There are a number of types of knowledge generated and circulated in education systems. These types of knowledge are not mutually exclusive but rather co-exist in a system, and their position and circulation may vary in different educational contexts. They include:

- Knowledge from international and national accountability systems (e.g. indicators);
- Knowledge of local authorities (context);
- Knowledge from educational research;
- Teacher professional knowledge (experience, context, tacit);
- Student knowledge (experience, learning).

The present study looks at how these different types of knowledge are being used to govern the AfL programme from the central to the local level, and to what extent certain types of knowledge are more powerful than others.

Capacity Building

Capacity building in this report refers to the process of learning and knowledge production at various levels of the education system, which is intended to enhance the ability to work with 'Assessment for Learning' in a constructive way. Education systems are based on professional knowledge developed through the education of teachers and through development of knowledge based on sharing of practical experiences and reflection in schools. However, there is evidence of a lack of professional development based on theoretical and empirical knowledge in the teaching profession (Klette og Carlsen, 2012; Tolo,

2011). Confronted with the demands of a knowledge society, this is a challenge for the development of the education sector (Jensen, 2012).

Capacity building is not a linear process. For example, the government cannot simply order courses from consultants and have teachers or school leaders passively absorb this knowledge and change their practice in accordance with the governments' intention. As with education more generally, in professional development the learner is a creative individual who engages in the learning processes in ways we cannot foresee (Lillejord, 2003). New demands, new scientific knowledge, new and revitalized theories are embedded into the profession's knowledge base and lead to new collective thoughts about what constitutes good practice.

Professional development depends on the competence of each individual and the system as a whole, and individuals in the system have to know how to implement the intended change. This implies that school leaders also need to know how to facilitate change, they need knowledge about the content of the changing process in terms of what works and what are the theoretical assumptions underlying the new teaching paradigm and they need support to be able to maintain this focus over time.

Box 1.1. Questions regarding knowledge and capacity building in this case study

- To what extent has the Norwegian Directorate for Education been able to implement the AfL programme, and how have they done this?
- Has the programme succeeded in making an infrastructure for knowledge development on Assessment for Learning?
- How does "new" professional knowledge interact with the existing knowledge base of teachers and school leaders?
- Is it being integrated into the knowledge base of teachers and accepted as a core practice?

CHAPTER 2.

EDUCATION IN NORWAY

Demographic and social context

With a population of 5 million people, Norway is a small country, which is also geographically large, with long distances and scattered population settlements. In contrast to neighbouring countries such as Sweden (9.5 million), Finland (5.4 million) and Denmark (5.6 million), Norway is not part of the European Union. Norway enjoys a stable economy, partly due to its wealth in oil and gas resources in the North Sea. Democratic ideals are strong and there is an expectation of equal access for all to Norwegian schools and universities. The International Civic and Citizenship Education Study shows that knowledge of democratic ideas and citizenship is high among Norwegian students (Mikkelsen et al., 2011). Norway, together with countries such as Iceland and Sweden, has been at the top of the Quality of Life Index survey for the past few years. Education is seen as an important part of the welfare system, which offers free schooling for all, from primary to upper secondary school, as well as free access to universities and colleges.

Approximately 614 000 students attended Norwegian schools in 2011 and this number is expected to increase to 668 000 by the end of 2020. In 2011, there were 3 000 schools in Norway, 344 less than in 2001. Although many schools with very few students have closed, 33% of the schools in Norway still have less than 100 students. Forty percent of the schools have between 100 and 299 students, while only 27% of the schools have more than 300 students. In other words, 8% of the student body attended schools with less than 100 students, while the majority of students (55%) attend schools with more than 300 students. Most of these schools are public schools, but there has been an increase in the number of private schools (from 2% in 2001 to 3% in 2012). Of the private schools, two-thirds have less than 100 students, and they are mostly located in the main cities in Norway (The Directorate for Education and Training, 2012).

There are 19 counties and 428 municipalities in Norway. The capital, Oslo, is both a county and a municipality.

Economic context

Norway has for years built and extended a welfare state known as one of the Nordic models. It contains universal welfare schemes, a close cooperation between government, unions and private business organizations and a private sector that functions well. One of the overarching goals is to keep a low unemployment rate and to have a well-qualified work force. The unemployment rate in Norway is overall less than 4%, and youth unemployment is low compared to other countries, at 7.6 % in March 2012 (in comparison, youth employment in The European Union was on average 22.6 % in 2012).[1] In such a positive economic climate, Norwegian students grow up feeling more secure about the future than many other young students in Europe, and they do not face the same degree of pressure to succeed in school in order to later find a job as do many other young students around the world. The growth of the mainland

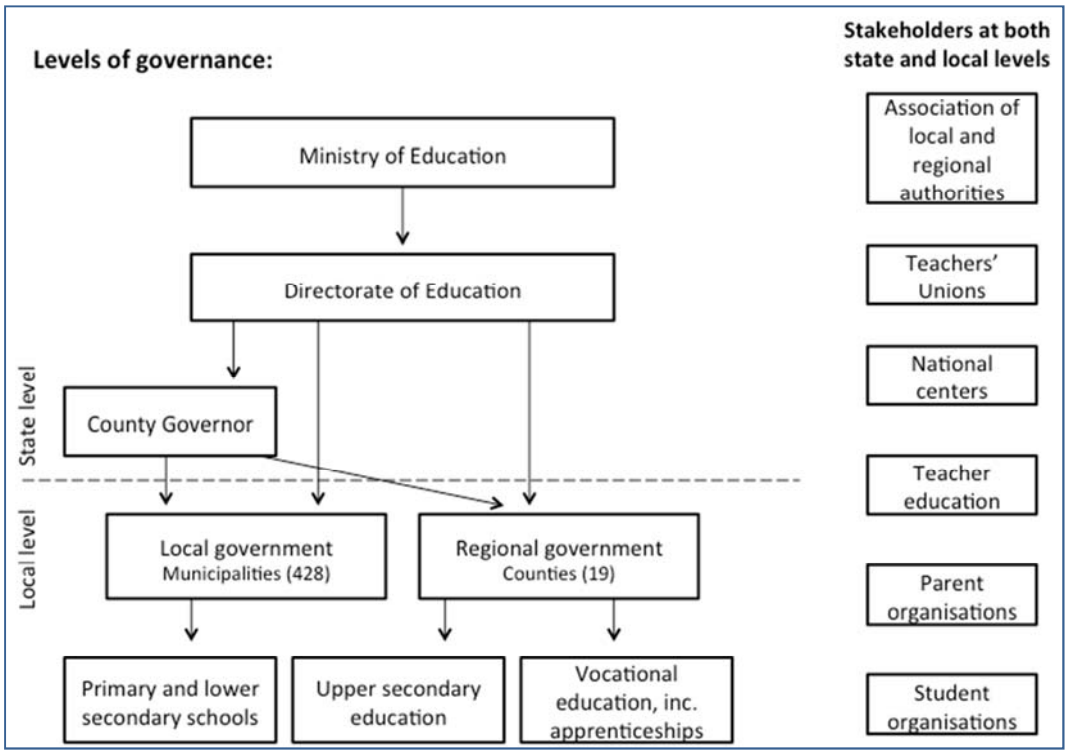
economy is essential for the development of Norway's welfare system. The revenues from the oil and gas sector put Norway in a more favourable situation than many other countries. In 2012, Norway was classified as one of the major oil exporters in the world.

The Norwegian Parliament has put restrictions on the annual spending of oil and gas revenues into the economy (so-called "*Handlingsregelen*"), and the Government has long agreed to save money for future generations and has passed legislation ensuring that oil revenue goes to a sovereign wealth fund used for investments around the world.

Governance actors and different stakeholders

The public school system in Norway is governed through a decentralised model (see Figure 2.1). The *Storting* (Norwegian Parliament) and the Government formulate the objectives for education, adopt legal frameworks (the Education Act and its regulations), and evaluate the status and condition of the day-care and education sectors. The local (428 municipalities) and 19 county authorities have the overriding responsibility for financing primary and lower secondary education.

Figure 2.1. Levels of governance and different stakeholders in Norway



Source: Adapted version of Theisens, H. (2012), "Effective governance from the centre", paper presented at the First Thematic GCES Conference, 21-22 November 2012, The Hague, OECD/CERI.

Costs are financed in part through block appropriations from the state and in part through municipal and county revenues. The education budget is adopted annually by the Norwegian Parliament and then transferred to the municipalities without being earmarked for education.

The Ministry of Education and Research is responsible for formulating national education policy. National guidelines are ensured through acts, regulations, the curricula and framework plans. The

Directorate for Education and Training is an administrative agency under the Ministry of Education and Research. The regional level (the offices of the County Governors) is the link between the Ministry of Education and Research and the Directorate for Education and Training on the one side, and the education sector in the municipalities and counties on the other. The school owners of primary and lower secondary schools are the local authorities, and the owners of upper secondary schools are the counties. The school owners of private schools are the school boards.

Education is compulsory until age 16. It comprises Primary (Year 1-7) and Lower Secondary Schools (Year 8-10). These schools are governed by local school owners that in 2012 represented 428 municipalities in Norway. The municipalities are generally small: 39 municipalities have between 20 000 and 49 999 inhabitants, while 95 municipalities have less than 2 000 inhabitants, and 134 have 2 000–4 999[2] inhabitants. There are only 13 municipalities with more than 50 000 inhabitants. All municipalities are expected to offer their students the same quality of education, but municipalities differ in the way in which they govern their schools, both with regard to the number of staff available to support the local schools and when it comes to the amount available in their budgets.

When students are 16, they start Upper Secondary School. In Norway, students have a legal right to attend Upper Secondary schools. Their grades determine whether they are able to get into their preferred school. Despite the free school system, less than 75% of the students finish these three years on time and receive their Upper Secondary School Diploma. Several policy initiatives have been designed to improve these numbers and the current AfL programme can also be seen as a part of this drive. The Upper Secondary schools are governed by one of the 19 regional government counties. At the State level, the County Governor is responsible for implementation and administration of the state's education policy in their region, and is also responsible for inspecting public schools. The County Governor reports to the Directorate of Education and Training, which in turn reports back to the Ministry of Education. Norway's capital, Oslo, is the biggest of the regional government counties, with more than 600 000 inhabitants, while Finnmark is the smallest county with less than 75 000.

Box 2.1. Student organisation in Norway

Students in Norway have their own Organisation, *Elevorganisasjonen* (EO), which is an active stakeholder working for students' rights in education. According to its own web pages, EO is politically independent. Since 1999, the organisation has been campaigning for students' rights in upper secondary school. From 2007, the organisation also started to include pupils from the lower secondary school and recently became one of the important stakeholders in the development of Norwegian schools. This is reflected in their annual conferences, where the Minister of Education was invited as a guest speaker.

Among the Organisation's achievements are the laws that give students in Norway free school books and the right to have student councils in all schools. EO launched their new campaign in November 2012 with a particular focus upon pupil participation. The campaign comprised material, which was developed in co-operation with the Teacher Union in Norway and *Kommunesektorens interesse- og arbeidsgiverorganisasjon* (KS), the Norwegian Association of Local and Regional Authorities. The material includes examples of questions that can be used in dialogues between pupils and teachers to improve learning in Norwegian schools.

Parents in Norway have their own Organisation, *Foreldreutvalget i skolen* (FUG), the National Parents' Committee for Primary and Secondary Education. FUG works on different topics such as how to improve dialogue between parents and schools, how to prevent bullying and how to improve subject skills for all pupils. Members from EO and FUG regularly participate in discourse on Norwegian schools in the media, in TV debates and in meetings with school leaders and policy makers. In addition to these groups, other stakeholders include different Teacher Education Universities and Colleges in Norway as well as Teacher and School Leader Unions, which all act as important stakeholders in the education system. The

Leaders of the Teacher Unions meet regularly with members from the government and participate in dialogues about education policy. For all stakeholders in education, there are high expectations of active participation.

Cultural, political and historical factors in Norway

Norway has a long tradition of equity, which has included developing a school system that is open and accessible to everyone. To understand the development of the Norwegian school system, it is necessary to consider the changes of political rhetoric over the last 40 years. In 1936, the unified school system was established under a law declaring that all children should attend the same school regardless of their place of residence, their abilities or socio-economic background. The underlying argument was that the whole society would benefit from a school system where students from all strata of society had the same opportunities for educational success and received an appropriate education (Dokka, 1975). This fits with the Scandinavian social democratic ideal of equality, where the state also plays a role in reducing inequalities in society.

This commitment to equity explains why the majority of Norwegian schools are public, and very few (3%) are private. Most students in Norway attend a comprehensive school (*Enhetsskolen*), which is a free school open to all. It has been a political goal for decades to offer free education for all, and include children from across all religions, ethnic origins, socio-economic statuses and genders. However, today the ideals of the comprehensive school are contested. Despite these long-standing goals, Norwegian schools have not been able to decrease the performance gap between the majority of students and those from lower socio-economic backgrounds (Hernes and Knudsen, 1976; Bakken and Elstad, 2012). However, compared to most OECD countries, difference in performance across all students is still relatively small (OECD, 2010).

Despite the system's emphasis on equity, there is evidence of comprehensive schools producing inequalities in learning conditions, achievements and opportunities. There has therefore been a shift away from describing the Norwegian school as "comprehensive schools" towards the term "quality schools" (Welle-Strand and Tjeldvold, 2002). This change in attitude may be due to the "reality orientation" stemming from the PISA tests results in 2001 (Lie et al., 2004). Prior to this, Norwegians believed that overall they had a good school system, but the results from PISA 2000 revealed that Norwegian student achievement was average when compared to other countries' educational systems. To understand the bigger picture, it is important to consider why tests like PISA were introduced and why the Norwegian educational system turned its attention towards international rankings.

In 1975, the Labour Party was in power in Norway. In an interview in Aftenposten, a leading Norwegian newspaper, the Minister of Education and Church Affairs, Bjartmar Gjerde, was cited as saying that Norwegian schools were the best in the world.[3] Until the 1980s, no data were produced on the status of standards in Norwegian schools, so there was no real reason for the public to question the appearance of high standards (Telhaug, 2006, 2007). Norwegians based their conclusions on their strong beliefs rather than on any real evidence. It therefore came as a shock when the OECD Review of National Policies for Education in Norway claimed that policy makers lacked evidence on the quality of the Norwegian school system, as well as knowledge of how funds were spent and invested (Kogan, Lundgren and O'Donoghue, 1989). According to the expert group from the OECD, a major problem was the lack of documentation of knowledge which the Norwegian government could use for steering purposes (Kogan, Lundgren and O'Donoghue, 1989, 1990; The Ministry of Education and Research, 1988: 28, p. 25). It was further argued that management by objectives rather than management by rules would produce better utilisation of resources in the education system (Lillejord, 1997). The OECD report criticised the lack of valid statistics from Norway and argued that the political governance of Norway was weakened by the paucity of information on whether national goals were being reached (Lundgren, 2003).

The OECD report was followed by a project called EMIL[4] initiated to improve school evaluation at the national level. According to Lundgren (2003), the project was considered a threat to local autonomy, and faced strong resistance from teachers. Public discourse in Norway in the early 1990s was focused on school-based evaluation (Monsen, 2003). A national evaluation system was thus seen as opposing the philosophy behind school-based evaluation, a philosophy based upon school development and less on control and accountability. These attitudes changed during the 1990s, when the concept of quality of education came to the foreground. Lundgren writes: "It was as if (the) prefix *e* in equality was dropped" (ibid., p. 106). The OECD report led to what has been called the "new political rhetoric in education" (Telhaug, 1990), as well as a shift of focus from the school level to the system level.

A few years later, a committee led by Astrid Søgnen on behalf of the government published a report that proposed a national quality assurance in Norway: First Class from First Grade (The Ministry of Education and Research, 2002). It was followed by a White Paper to the Storting[5] (2003-2004) in which all the politicians agreed on the new system and that the purpose of the quality system should be for school development, not for control (Lillejord and Hopfenbeck, 2013).

To understand the rhetoric of Norwegian education, it is important to consider the discourse from the 1980s. Gudmund Hernes, later Minister of Education and Church Affairs (1990-1995), led the work on a public report (Med viten og vilje, The Ministry of Education and Research, 1988, p. 28). This report has been referred to as one of the most important background documents to the policy events of the following years (Telhaug, 1997; Tolo, 2011; Trippestad, 1999). The introduction stated, "The challenge for the Norwegian knowledge policy is that the country does not get enough competence out of the talent of the people" (The Ministry of Education and Research, 1988, p. 7).

The timing here was not accidental. Education policies had become influenced by a new kind of economically motivated seriousness. In times of economic recession, as was the case at that time, education becomes crucial and central to the political debate. This was not a new thought; this trend can be traced to a time 150 years ago when there were discussions about "the appropriateness of the knowledge taught" (Dale, 1985, p. 2) in the British education system. There were two aspects to this discussion: one arguing against 'academic drift,' and one arguing for the education system's responsibility to meet the needs of industry and businesses (ibid. 1985). During the 1980s, however, the global education trend was woven into a picture where the State and the market were seen as closely related. The same arguments about the education system's legitimacy and tasks won out in country after country, eventually winning over Norway as well (Apple, 2005; Ball, 1990; Biesta, 2004; Hyslop-Margison and Sears, 2006; Rizvi and Lingard, 2009).

It may seem strange that Norway, a rich oil-producing country, would prepare for recession, but this has indeed been an issue since the 1980s. At that time, the sociologist Hernes was formulating a central argument for developing concern for the long-term state of the Norwegian economy. In 1987, he wrote in a leading newspaper that the "debate has been running for three months now. And this is about more than the University of Oslo. It is about the Norwegian welfare state's future. In a few decades, the oil will be drained and the gas will be burned. But even today, we first and foremost earn our money by exporting know-how ..." (Hernes Dagbladet 19 March 1987, quoted from Trippestad, 1999, p. 198, translated from Norwegian).

The "Knowledge Society" has been presented almost as an organism in need of expertise. As Hernes points out, even a country that exports raw materials must prepare for a future situation in which the nation's first and foremost aim is to produce knowledge. The main capital in a country is developed through skills and knowledge in the broadest sense. This thinking led to system-level curriculum reform in the 1990s (Reform 94), and to a further reform in 2006. The 2006 reform promoted elements deemed essential to a knowledge society: goals, competencies, basic skills, learning how to learn, and learning

strategies. Although it fine-tuned Reform 94, the 2006 reform safeguarded the democratic spirit of Norwegian values such as participation for all.

Assessment practices before and after PISA

Before 2000, Norwegian schools had been experimenting with a variety of approaches to assessment and evaluation, and many case study reports described local cases where teachers successfully collaborated with researchers (Lund, 2000). The municipalities were responsible for developing quality assurance for their schools, but since only half of the counties managed to do this, a stronger national approach was needed to guarantee that all children in Norway had a proper education (The Ministry of Education and Research, 2002a).

In 1999, Norway participated in the piloting of PISA, partly building upon work from the Indicators of Education Systems programme (INES). When the results from the first PISA cycle were published, Norway had its PISA shock: the country placed 13th of 30 countries in achievement on reading, a full 30 points behind the leader (and neighbour) Finland. According to Petter Skarheim, Director of the Department of Education and Training, PISA marked the beginning of one era in education and the end of another. In an international conference in Oslo in 2010, he stated, "Before PISA there was nothing", making the point about there being little statistical knowledge available to Norwegian schools (Skarheim, 2010).

As in Germany (Ertl, 2006), the PISA results came as a shock to the Norwegian public who believed the Norwegian school system had been among the best in the world. At the same time as the first PISA results were released in 2001, other research studies confirmed the impression of a school system with several challenges: not enough focus upon learning goals, teachers who did not have high enough expectations for all students, and lack of leadership from teachers in the classroom (Haug 2003, 2004; Hertzberg, 2003; Klette, 2003). Although it was made clear that student-teacher relations in Norwegian classrooms were generally very positive, teachers often appeared to praise their students without explaining explicitly what was good about their answer or achievement. Several researchers therefore accepted PISA results at face value, since they corresponded with their observations of what they perceived to be ineffective classroom practices, with less focus upon learning goals than expected. Political parties also used PISA in their rhetoric when they wanted to influence the schools. The Secretary General of the Minister, Helge Ole Bergesen, later wrote a book after his term explaining how PISA was a "gift from heaven" (Bergesen, 2005) and that the timing was excellent for implementing new policy.

After the release of the PISA 2001 results, the Ministry of Education launched a report called *The School Knows Best*. In the introduction, Kristin Clemet, the Minister of Education stated:

If we are supposed to improve our achievements, we need to change the way we are steering our schools. We have focused too much on detailed regulations and steering of resources and processes from the top down. Our experiences and international research shows us that this model does not work well enough. We need to decentralize responsibility, improve quality control and give participants increased influence. Schools should be steered from the bottom up, and not from the top down, but according to national goals (Ministry of Education and Research, 2002b).

This report stated that international student tests indicated achievement levels in Norwegian schools were mediocre, despite the high level of economic support for Norwegian schools. Even more disturbing was the scarcity of knowledge available from national research about the quality of Norwegian schools. The lack of empirical research on learning outcomes in Norway was viewed as problematic. In addition, Norway did not have a national quality assurance system, like most other countries in Western Europe, and

therefore the Ministry proposed new strategies for introducing quality assurance as suggested in previous policy documents (Ministry of Education and Research, 1991, 1996, 1999)

As a part of the initiatives taken to improve Norwegian education, different centres were established to support teachers and researchers in their work in schools: *Matematikksenteret* in Trondheim (The Norwegian Centre for Mathematics Education 2002), *Naturfagsenteret* in Oslo (The Norwegian Centre for Science Education, 2003), *Lesesenteret* in Stavanger (The Reading Centre, 2004). *Fremmedspråksenteret* in Halden (The Norwegian Centre for Foreign languages in Education, 2005), Skrivesenteret in Trondheim (The National Centre for Writing Education and Research, 2009) and in Stavanger and Porsgrunn *Nasjonalt senter for læringmiljø og atferdsforskning* (Norwegian Centre for Learning Environment and Behavioral Research in Education, 2013). Major reforms were initiated, such as *Realfag naturligvis*, (Science of course!) and *Les!* (Read!). When evaluating the implementation of the current programme, these initiatives will be taken into account as well.

The Norwegian quality system

Norwegians have had a long history of being proud of their schools, but the development of a national quality assurance system first took off after the release of the PISA results in 2001. Before 2000, mapping tests had been used for diagnostic purposes in reading, but national tests in Norway were not introduced until 2004, when Reading (Norwegian) and Mathematics in Years 4 and 10, and English reading in Year 10 were introduced. In 2005, national tests were introduced for Year 7 and the first year of Upper Secondary School, and included writing tests in Norwegian and English. The Norwegian Parliament, the *Storting*, supported the introduction of these tests as part of a new quality assurance of assessment. However, in 2005 there was resistance from teacher organisations and students, and concern that the writing tests in English and Norwegian were not marked in a reliable way (Lie et al., 2004, 2005). The government thus postponed the tests in 2006 and used the time to improve them before introducing the new national tests in 2007 (Eklöf et al., 2012). As of 2012, the national quality assurance in Norway includes mapping tests, national tests, and end of school exams. According to the DET's Web site, the main purpose of the national tests is "to assess to which degree the school has successfully developed students skills in reading and mathematics, and English. The results are supposed to be used by the school and the school owner as a part of the quality development in schools" (http://www.udir.no/vurdering/nasjonale-prover (accessed 10 December 2012).

The Directorate also developed a new brochure in 2012 explaining how the national tests are supposed to be used for formative assessment. Under the heading "How to use tests in the work on assessment for learning" it states that it is important to make sure that students know why they are assessed and what is expected of them. Teachers are encouraged to tell students that the purpose of the tests is to know what students are able to do, so teachers can help them to learn more. Teachers are also encouraged to tell students how to take the tests, get them to know the tasks, use example tasks from the Web site, and help them feel comfortable with the test situation.

The document sets out the four principles of Assessment for Learning (see Chapter 3) and emphasises that the main purpose of the tests is a formative one. The folder also suggests how to talk to students and parents about the results and how to work with colleagues to understand the results and make improvements in the school. Since the introduction of the National Assessment System (NAS), many researchers agree that there has been an increased focus on students' performance and educational outcomes that demand new modes of school governing and the use of data (Skedsmo, 2011).

Recent research reports further indicate variation in the way school leaders use such results from the NAS, and that not all schools have the capacity to use and interpret the results in the new national system

(Aasen et al., 2012). Another recent study evaluating annual reports written by school owners (private, municipalities and counties), also points to the problems many school owners have in using information to improve schools. There is wide variation between school owners in how they conduct and use reports, and who is involved in this work. For example, this can vary according to the size of the municipality or rwhether it is two-, or three-level system. Only 33% of school owners had AfL as a written goal for improvement, and a qualitative analysis of the reports revealed vague, unclear, imprecise, unrealistic, standardized (similar to goals in state documents) and over-ambitious goals (Rambøll, 2013). One of the conclusions is that the evaluation capacity of school owners needs to be strengthened so that evaluation can be used as a tool for quality improvement (Rambøll, 2013). This raises issues concerning data management in more decentralised systems, particularly in the Norwegian context, and the way in which schools and school owners attempt to use these data but do not always succeed, perhaps because of the need for more capacity building (Campbell and Levin, 2009) or, on the other hand, because of the overload of external data that schools today must handle (Lawn and Ozga, 2009; Ozga, 2009).

The end of school exam in one of the subjects – Mathematics, English, or Norwegian – is the first high-stakes test students take in their life. It takes place the year students turn 16, and it is the only exam that is centrally set and marked by two external raters. Trust in teachers is, therefore, high in comparison to countries that rely more generally on externally marked tests. For most subjects in Norwegian schools, students are graded (including their final grades) by their subject teacher.

At the end of Year 10, pupils are marked in 13 subjects by the respective subject teachers, a system based on a numerical scale with marks from 1 (very low competence in the subject) to 6 (excellent competence in the subject). Based upon these grades, each student receives an overall score representing their results. In 2012, the average score was 40. These scores have been stable in recent years, but they also show that girls achieve on average 4 points higher than boys (Girls: 41.5 points, Boys 37.6 points).[6] Overall, subject marks given by teachers to their pupils are slightly higher than those given on final exams. Marks in Mathematics vary the most in the 13 biggest cities in Norway. Finally, only 2.4 % of the Norwegian pupils were enrolled in private schools in 2011–2012, and these students performed better than the rest of the students (by 3.7 points). The Directorate accounts for these results with the different backgrounds of pupils in private schools, which are not representative of Norway as a country.[7] In the present study, differences in the AfL programme with respect to private and public schools were not analysed, but it should be noted that private schools participated in the programme together with public schools.

CHAPTER 3.

THE CASE STUDY: ASSESSMENT FOR LEARNING (AFL)

Description of the policy programme and implementation strategies

The Assessment for Learning (AfL) programme involves 184 of the 428 municipalities in all 19 counties in Norway. The initiative is a continuation of an earlier two-year programme project, *Bedre Vurderingspraksis* ("Improved Assessment Practices"), which included 77 schools between 2007 and 2009. The overall goal of this project was to improve formative assessment practices in the classroom by developing distinct criteria to clarify how to reach curriculum goals (Hopfenbeck et al., 2012). Based on experiences from this project, the Norwegian Directorate for Education and Training recommended further systematic investment in assessment in Norway. The recommendation was endorsed by the Ministry of Education which then asked the Directorate to outline and lead the AfL programme. The programme was based upon four principles for quality formative assessment outlined in The Education Act[8], where it is stated that the main purpose of continuous assessment should be for learning. The four principles of the Education Act are also mandatory for Vocational Education and Training (VET) and are within the counties' responsibility. They state that students and apprentices learn better when they:

1. understand what to learn and what is expected of them (Regulation § 3-1);
2. obtain feedback that provides information on the quality of their work or performance (§ 3-11);
3. are given advice on how to improve (§ 3-11); and
4. are involved in their own learning process and in self-assessment (§ 3-12).

The main goal of the AfL programme has been to improve assessment practice in Norwegian schools by working on integrating the four AfL principles into their teaching practice (for general guidelines on the programme see Box 3.1). It is interesting to observe how the Norwegian version of AfL is centred on the aspects of criteria, feedback and self-assessment rather than aspects that have emerged as priorities in other contexts, such as the need to reconcile formative and summative assessment or improving the quality of questioning (Black et al., 2003). This illustrates how processes of policy borrowing involve adaptation of the policy to the local context and its values (see for example Phillips, 2004).

Box 3.1. Smoothing the implementation of the AfL programme

In order to help implement the AfL programme, a range of core documents are provided by the Directorate to the municipalities, which include:

- *The Base Document,* which describes the aims of the programme, common guidelines, roles and

responsibilities for all participants.

- *Invitation letters* to each group of school owners. These are based on the base document and elaborate the responsibilities, expectations and criteria for participation and financial support. Municipalities are required to participate with at least three schools (and counties with at least five upper secondary schools and three training establishments) to receive the 250 000 NOK (about 33 000 EUR) for municipalities or 800 000 NOK (about 106 000 EUR) for counties in financial support.

Box 3.1. Smoothing the implementation of the AfL programme (*cont.*)

- *Templates* for the school owner's plan for the local programme and mid-term and final reports from the school owner. For instance, school owners would have to describe:
  - the plan and outcome of competence building for schools and teachers;
  - how to involve school leaders, pupils, parents and other actors;
  - how they plan to roll out, share experiences and competence and continue the work after finishing the national programme.

In addition to these documents, further information was made available as part of the documentation on the AfL programme:

- *Reports and self-evaluation* from all participating school owners;
- *The Pupil Survey*, which contains questions concerning assessment practice. Results are available both at national, school owner and school level.

Source: Norwegian Directorate for Education (DET): www.udir.no/Vurdering-for-laring/VFL-skoler/ and www.udir.no/Laringsmiljo/Elevundersokelsen/ (both accessed May 2013).

The programme was launched in 2010 and led by a team from the Directorate. As a part of the implementation strategy it was decided that the Directorate would work closely with participants over a time span of 16 months. From the point of view of the Directorate, these 16 months are framed as the beginning of a longer process of developing and facilitating change in the assessment culture of schools.

Four waves of participants were envisaged for the period 2010-13 (see Table 3.1):

Table 3.1. Participating groups and timeline

Group	Participants	Period
1	50 municipalities and 287 primary and lower secondary schools from 6 different counties	Sept. 2010–Jan. 2012
2	All 19 county administrations, 97 public schools, 19 private schools and about 55 training establishment	Feb. 2011–June 2012
3	63 municipalities, 248 public primary and lower secondary schools and 15 private schools from 7 different counties	Feb. 2012–June 2013
4	71 municipalities, 200-300 public primary and lower secondary schools and 13 private schools from 6 different counties	Feb. 2013–June 2014

Each group attended introductory meetings with the Directorate, where the goals and expectations of the programme were outlined together with the implementation strategies. A letter from the Directorate[9]

explained that the overall goal of the project was "to ensure that the school authorities, schools and training establishments would develop an assessment culture and a practice with focus upon learning".

The municipalities, counties and private schools have the freedom to decide how the AfL programme should be designed and developed at the local level, but certain basic principles must be followed:

(1) The programme should include Assessment for Learning.

(2) All schools should include at least one of the four assessment principles from the *Education Act*.

(3) Schools should use assessment to adjust teaching and improve student learning.

(4) Schools should provide a safe learning environment where the focus resides upon student learning and goal achievement.

(5) The programmes should involve all participants at all levels, and use feedback between all participants and levels in the system.

The school owners are also obliged to give financial support to the local implementation of the programme, in addition to support from the Directorate.

The Directorate further suggested how the school owner could develop a plan for checking whether and how schools implemented the AfL programme. The plan was to have measurable, specific, acceptable, realistic and time-bound goals. The plan would also include a description of which of the data-sources had been used for the evaluation of the programme. The school owners were then expected to write two reports to the Directorate (one mid-term and one at the end of their participation period), in order to evaluate the work in their county or municipality, describe their plans for future work on assessment practice, and advise on further implementation strategies needed from the Directorate and the County Governor (*Fylkesmannen*). These reports were all self-reports explaining how schools in their municipality had implemented AfL.

Each school owner had to appoint one person to be the school owner's "resource person". The role of this person was to support the school owner in running the programme locally, for example by developing an implementation plan and "learning networks" for schools. According to the AfL guidelines, these networks should be based on the processes to improve assessment practice in each school or learning establishment. The local and regional networks can be set up in various ways, but meeting places must be provided between schools. The networks are seen as arenas for professional development, sharing experiences and reflection and should be based on concrete examples from practice supported by theory and research. The use of "learning networks" and "resource persons"[10] for the school owner are two main aspects of the implementation model. Part of the reason for choosing these elements was that the school owner can use established structures and persons' enhanced competence (resource persons and leaders and teachers at participating schools) in their future work on assessment practice.

In addition, the Directorate organised a series of conferences to support the resource persons' work locally. These conferences also provided the opportunity for all actors involved to learn more about the programme and to work and learn from networking with other school owners, school leaders and teachers. Researchers from the Assessment Reform Group (ARG) in the UK were also centrally involved in the meetings (including, at various different times, world-renowned experts as Louise Hayward and Gordon Stobart). The programme focused on classroom practices that support AfL and elements concerning the implementation of AfL, in particular the "learning networks" as a tool for school development. While the

explicit goal was to increase student motivation and knowledge of their own learning, reports from school owners indicate that a positive (and unintended) side effect was that this systematic work with assessment decreased the number of complaints from students and parents on grades.

The Directorate was responsible for (1) the outline of the programme, (2) co-operating with the school owners, (3) preparing conferences for the participants, (4) developing the web-based tool *Vurdering for Læring*, (5) providing financial support to all participating school owners, and (6) contributing in summarizing and sharing the knowledge and experiences of the programme. The county governors were responsible for selecting schools for the programme, financially supporting and monitoring these participants, while the municipal and county school owners had responsibility for planning, conducting and following up the different schools in the programme. These different roles and responsibilities will be explored in more detail when discussing the findings of the case study in Chapters 4 and 5.

The main goal of this programme can therefore be seen as supporting school owners and schools in developing a better assessment and learning culture. The initial programme did not have any expectations of raising achievement, standards or national test scores. Rather, the focus has been on how Assessment for Learning can enhance students' learning. The AfL programme should not be considered a mere assessment tool, but more as a method for quality teaching (Stobart, 2008; Gardner, 2011).

After the first participating group of 50 municipalities had completed their 16-month programme, the participants wrote reports for the Directorate summarizing their experience and implementation of AfL, which were published on the web.[11] Among the first reported results were differences among schools, as well as an understanding of the need to work for a shared assessment practice in and between schools in the different municipalities. One reported success was the municipalities' use of networks as an implementation strategy tool among their schools. Challenges reported included the need for a school leader that endorsed the implementation in the school, a shared language of assessment, a shared practice around documentation, and having all teachers in a school involved. The final report was used to adjust the implementation and inform the process for groups 2, 3 and 4.

The website *Vurdering for Læring* (http://www.udir.no/Vurdering/Vurdering-for-laring/) contains resources for teachers and school leaders. New material is posted on the website as the project progresses. Research literature, books in Norwegian about AfL and video footage from classrooms in Norway showing teachers and students implementing some of the principles are available, as are guidelines on how school leaders can work and reflect with their teachers using these films, along with discussion questions. Keynote presentations and some video recordings from the conferences are also posted on this website as references for teachers and school leaders who were not able to attend the conferences.

The many documents available online on the official website show how this programme has been centrally initiated both in content and expectation of how AfL should be interpreted. Although several documents insist that no one recipe for AfL exists, the government expects all schools to implement the four principles of AfL, and even made it mandatory through the Education Act. In this sense, their approach is similar to that proposed by Thompson and Wiliam (2008) under the term "tight but loose", that is, keeping in mind core central principles (tight) whilst leaving the way to implement these principles according to the teacher's initiative (loose).

In this respect, it would be fair to say that the Assessment for Learning programme has been specifically targeted at about 880 schools and 55 training establishments, but that the ideas behind it are now embodied in law, so students can see these principles as "a right" when they interact with teachers in all Norwegian schools. In the next section, an overview of empirical studies on AfL is outlined, which will serve as foundational evidence for the analysis of the results in Chapter 4.

Assessment for learning as policy: International evidence

There seems to be a research consensus regarding the positive effects of Assessment for Learning as perceived by participants (see for example Ofsted, 2008; DfES, 2007; Condie et al., 2005; Hayward and Spencer, 2010; Kirton et al., 2007; Webb and Jones, 2009; Kellard et al., 2008). However, although the programme has been researched, piloted and/or implemented in a wide variety of contexts (Australia, New Zealand, the US, Canada, Hong Kong, Chile, Singapore, Rwanda, Cameroon, The Netherlands, to mention just a few), a series of caveats and difficulties have been found in relation to the feasibility of the implementation of Assessment for Learning system-wide. These are explained below.

Teacher resistance to peer and self-assessment

Various authors refer to teacher distress generated by the introduction of peer and self-assessment (MacPhail and Halbert, 2010; DfES, 2007; Willis, 2008; Kirton et al., 2007; Stiggins and Arter, 2002). This is attributed to the changes that such practices entail in terms of the traditional classroom power relations. Teachers who are accustomed to leading their classroom in conventional ways, such as teaching students from the blackboard and doing all the talking, can feel uneasy about a potential loss of control once more power is given to students. Dysthe (2008) argues that a change of teaching style is challenging, and very often researchers have understated just how challenging such a change is for teachers.

Teacher resistance to change in teacher and student roles

The change entailed in teacher and student roles is rather crucial to peer and self-assessment: "It is essential for teachers to embrace the change in teacher-student relationships that is involved in implementing formative assessment. Without this change, students will not use feedback for assessment to be used for learning" (Harlen, 2007: 20; James et al., 2007) described the challenge for teachers when they need to think about ways of shifting the responsibility of AfL practices to the students. They have argued for the need to teach students how to "step back" from the learning process and reflect upon the task in dialogues with other learners, and then to "step back in" to restructure or transform the learning process. Key to this is promoting students' autonomy and self-regulatory skills, so that they are able to use self and peer assessment. These processes are challenging, but teachers can foster self-regulated learning when using AfL practices (Hopfenbeck, 2011).

Lack of commitment from senior staff

Black et al. (2003) refer to the importance of creating an assessment culture and generating a change at a whole-school level. Without this shift, teachers struggle against a different culture with different beliefs on assessment, leaving them alone and potentially demotivated. However, Ofsted reports[12] that in England detachment from senior staff in relation to Assessment for Learning in practice has been observed, although research considers engagement to be a crucial factor if Assessment for Learning is to be implemented in the long-term (Ofsted, 2008).

Shortcomings in teachers' disciplinary knowledge and assessment skills

The level of subject knowledge required from teachers remains an open question: What do teachers need in order to adequately understand and use Assessment for Learning in their specific subject? Which problems might occur if these needs are not being met? An additional impediment to deep reflection on learning expectations and assessment criteria, highlighted by a range of studies, is a lack of assessment literacy both on the part of in-service teachers and initial teacher education programmes. Concerns about subject knowledge and/or assessment skills in teachers in England are indicated by DfES[13] (2007),

Kellaghan (2004), Carless (2005), Gioka (2006), Stiggins and Arter (2002), Ofsted (2008), Thompson and Wiliam (2008), Azúa and Bick (2009), and Gardner et al. (2011).

Superficial understanding of the approach

In the context of an international conference that brought together Asian Pacific countries with experience in Assessment for Learning, Klenowski states that some misunderstandings "… have stemmed from a desire to be seen to be embracing the concept – but in reality implementing a set of practices that are mechanical or superficial and without the teacher's, and most importantly, the students', active engagement with learning as the focal point. While observing the letter of AfL, this does violence to its spirit" (2009, p. 263).

This concern about a superficial understanding of the approach leading to a mechanical or equivocal use of AfL in practice is supported by other authors like Hayward and Spencer (2010), Tapan (2001), Webb and Jones (2009), Ofsted (2008) and Azúa and Bick (2009). It is also found in the work of Harlen (2007), who describes how formative assessment can have both positive and negative impacts: On the one hand, formative assessment can have a positive impact on teaching and learning, on the other hand, a negative effect of formative assessment can be seen if teachers are following procedures mechanically without understanding the purpose of the assessment (ibid., p.20).

Busy classrooms and lack of knowledge on how to put AfL into practice

In their seminal and widely cited article from 1998, Black and Wiliam argue that teachers will not take up attractive sounding ideas, albeit based on extensive research, if these principles are presented as general principles that leave the task of translating them into everyday practice entirely to the teachers. The everyday classroom is too busy and Black and Wiliam suggested that transforming general principles into AfL practices will only be possible for the outstanding few. Black and Wiliam therefore suggest that teachers need "a variety of living examples of implementation, by teachers with whom they can identify and from whom they can both derive conviction and confidence that they can do better, and see concrete examples of what doing better means in practice"(Black and Wiliam, 1998, pp. 15-16).

They clarify their point by arguing that there are no "quick fixes" in formative assessment. Instead, they argue that each teacher should find his or her own ways of incorporating the lessons and ideas that are set out in the literature into their own patterns of classroom work. This is necessarily a slow process, and should include sustained programmes of professional development and support. If teachers find themselves too busy or overwhelmed, they will not be able to find the time and energy needed to change their teaching into AfL practices.

Scaling up

Some authors raise concerns when it comes to the replication of the practices at a wider level. The majority of the studies on Assessment for Learning have worked with groups of volunteer teachers who are willing to participate and are given special conditions for their involvement, namely, time for peer discussion and sharing experiences. This is seen as a crucial factor for the understanding of the underlying concepts. The question raised here is the extent to which ideal support conditions of small pilot programmes can be replicated when the approach is to be scaled up to, for example, an education system at the national level (Torrance and Pryor, 2001; DfES, 2007; Dori, 2003; Black et al., 2004; Black et al., 2003).

High-stakes testing systems and administrative requirements as an obstacle

In Black et al.'s (2003) study in England, teachers showed some level of anxiety towards the use of comment-only marking as they did not view this strategy as consistent with the school's administrative requirements in terms of grading. However, research findings refer to the teachers' surprise when parents, school senior staff and Ofsted did not react negatively to this. To solve this tension, teachers also attempted to reconcile formative and summative assessment in different ways (Black et al., 2003; Black et al., 2004). Despite these positive outcomes, the authors also recognized this reconciliation was only feasible in relation to summative assessment when teachers have control, which is not the case with external high-stakes assessment systems.

These concerns have been confirmed by other studies on the topic, which see testing as one of the main threats to the feasibility of the approach (Condie et al., 2005; Dori, 2003; Hayward and Spencer, 2010; Kirton et al., 2007; Gipps et al., 2005; Kellaghan, 2004). The pressures that such systems pose on schools motivate teaching to the test and a rhythm of work that is mostly centred on check-listing of test content. Teachers feel compelled to opt for this logic of work because of its connection with accountability mechanisms and associated public pressure despite its inconsistency with their teaching principles and values (Shepard, 1992). Teachers thus become subjected to role conflict (Berryhill et al., 2009) and feelings of deprofessionalisation. According to Shepard, the "kind of drill-and-practice instruction that tests reinforce is based on outmoded learning theory. Rather than improve learning, it actually denies students opportunities to develop thinking and problem-solving skills" (1992, p. 56).

Therefore, the underlying philosophy of Assessment for Learning, which requires deep learning, attention to difference in progress, detailed feedback, a variety of tasks to demonstrate learning, and a considerable amount of time, seems to be at odds with the unintended effects that testing systems have generated in school assessment cultures. To overcome this obstacle, the Assessment Reform Group (ARG) has suggested that "it is important that summative assessment procedures are in harmony with the procedures of formative assessment and that they are transparent, with judgments supported by evidence so that all involved can have trust in the results" (ARG, 2006, p. 3). Many schools give the impression of having implemented AfL when in reality the required change in pedagogy has not taken place. This happens, for example, when teachers feel constrained by external tests over which they have no control. "As a result they are unlikely to give pupils a greater role in directing their learning, as is required in AfL, in order to develop the capacity to continue learning throughout life" (ARG, 2006, p. 10).

Beliefs about assessment

Several authors refer to the relevance of considering teachers' and students' beliefs on assessment and learning as a relevant part of the implementation process of formative assessment strategies (see for example Torrance and Pryor, 2001; Brown et al., 2009; Brookhart, 2001; Willis, 2008; Tapan, 2001; Carless, 2005 , Marshall and Drummond, 2006. Initial resistance from teachers and students is likely found both because of the assessment culture of their school context and the misalignment between their beliefs and/or knowledge about assessment and those proposed by Assessment for Learning. However, researchers recognise that, after a process of dialogue and use of the approach, teachers finally find "congruence of AfL with their own values and beliefs" (Carless, 2005, p. 51).

This seems consistent with Ball et al.'s (2012) policy approach to Assessment for Learning. Assessment for Learning was perceived by case study teachers to be an exhortative or developmental policy, that is, a type of policy that motivates an "active policy subject, perhaps a more 'authentic' professional who is required to bring judgement, originality and 'passion' […] to bear upon the policy process" (2012, p. 94). However, in the process of enactment, schools struggle to reconcile this policy with

standards and testing policies, which Ball understands as an imperative or disciplinary policy that produces "a primarily passive policy subject, a 'technical professional' whose practice is very determined by the requirements of performance and deliverology, particularly in the high-stakes disciplines […]. Little reflexive judgement is required of this teacher; indeed, it could be argued that it is 'required' that judgement be suspended and ethical discomforts set aside" (2012, p. 92). These two policies bear different beliefs on assessment. Assessment for Learning, it seems, is closer to teachers' values and beliefs; thus, after some discussion it becomes accepted. Nonetheless, its enactment is in permanent conflict with the values of the standards and testing agenda, which are not consistent with those of teachers, who, due to the way the programme is implemented, do not feel free enough to disagree.

Alongside this, beliefs at the policy authority level on assessment seem to be relevant in the process of enactment, although this aspect has been studied little. James (2007) and Ecclestone and Daugherty (2006) have highlighted how ideological disputes are embedded in the discussion on Assessment for Learning as a policy. Ecclestone and Daugherty (2006) go further suggesting that the current enthusiasm of policy authorities towards Assessment for Learning is due to a distorted version of the original approach which is consistent with the performance agenda. According to them, AfL is treated in policy texts as frequent summative assessment, as part of personalised learning, and as other ideas that contradict the original sense of the Assessment Reform Group perspective.

All the aforementioned difficulties are exacerbated by the fact that AfL can rarely be observed in schools' daily practices, even when initiated as a national policy (Black et al., 2003; Thompson and Wiliam, 2008; Ofsted, 2008; Weeden and Winter, 1999; Tapan, 2011; Gioka, 2006; Kellaghan, 2004). As a result, it is difficult to study the programme and its effects empirically (Tierney and Charland, 2007). Due to this, Hayward and Spencer, in a qualitative study about formative assessment in Scotland, concluded that complexity should be incorporated as an aspect of research and policy making with consideration of all the factors and actors involved in the policy process as well as their interactions. They stress the need to "explore the world as seen from a range of perspectives" in order to "begin to understand a little more deeply what really matters in improving learning." (2010, p. 175).

Similarly, Bennett (2011) emphasises "the system issue" as the most challenging one to formative assessment reforms. He argues that coherence is needed throughout the entire education system in order to avoid discourses on assessments working against one another, as he sees is the case with accountability tests and formative assessment. He concludes: "…ultimately, we have to change the system, not just the approach we take to formative assessment, if we want to have maximum impact on learning and instruction." (2011, p. 19).

The effectiveness of AfL is thus very much dependent on a number of rather practical implementation factors. One key element is the national context in which it plays out. Despite the increased focus on educational outcomes in terms of student achievement and accountability, in the Norwegian context there is still relatively little pressure put on key actors compared to countries such as the United States and England (Elstad, Nortvedt and Turmo, 2009).

Impact of the Assessment for Learning (AfL) programme in Norway

Norway conducted an extensive four-year evaluation (2006-2010) of the Knowledge Promotion reform in general, which included Assessment for Learning as one of a larger parcel of assessment reforms in the Norwegian system. One study, acknowledging that assessment was one of the areas in which the largest changes on the school level could be observed, found that the AfL programme contributed to the "establishment of learning networks between schools, school owners, experts and national authorities" (Aasen et al., 2012, p. 15). Schools included in the study also reported a more systematic approach to

assessment than before. A second study concluded that during the reform period teachers communicated learning goals to students in a clearer way than before the reform (Hodgson et al., 2012). In addition, a national survey in 2012 interviewed school owners and leaders and compared schools participating in the AfL programme to those not participating. It found a "… small, but systematic difference […] when it comes to […] assessment practices in school" (Vibe, 2012, pp. 59-60), especially in relation to the link between assessment and learning and open discussion at the school of these issues.

Box 3.2. Norwegian research programme on assessment

In 2010, the Directorate commissioned a new programme with the goal of evaluating assessment practices in Norway after previous reforms, and with a focus upon assessment for learning. The first study, published in September 2012, reported the views of teachers, school leaders and school owners on assessment. The researchers claimed that teachers were positive about the AfL programme overall but found differences among primary, secondary and upper secondary teachers. Primary and secondary teachers reported having changed their assessment practice to a greater extent than upper secondary teachers (Sandvik et al., 2012). Even though the report had positive results, it relied mainly on self-reported data from interviews and questionnaires in which teachers reported that they had changed their practice. However, the next report, which will be published in summer 2013, will include classroom observations that will provide information about what is actually happening with AfL practices in the classroom. It is therefore eagerly awaited.

The overall goal of the AfL programme, as described in official documents, has been to enhance and develop a better assessment culture to support students' learning through a shared understanding of Assessment for Learning as described in the four programme principles. These principles are based on the premise that students should know the learning goals, should know about assessment criteria and how to be actively involved in the learning and assessment process. This allows for tracing the implementation and impact of the AfL programme by interviewing students about their assessment literacy and how much they are involved in the process. The impact of the AfL programme in this case study has therefore been investigated by interviewing students, teachers, school leaders, and municipality leaders about their assessment practice in schools. This data will be presented in Chapter 4.

As the programme did not specifically focus on whether it would be possible to measure the impact of the programme on student learning outcomes such as national test scores, such analysis must, therefore, be conducted with caution. Yet, to address research questions regarding expected results of the AfL programme, national tests scores from Mathematics, English and Reading are analysed and also discussed in Chapter 4.

CHAPTER 4.

THE AFL PROGRAMME IN NORWAY – A SUCCESS STORY WITH MAJOR CHALLENGES?

This chapter presents and discusses the results from the case study and aims to answer the research questions posed at the beginning of this report. In order to do so it concentrates on three central questions:

(1) How was the implementation of the AfL programme designed and organised? How were the expected results of the reform communicated and to whom?

(2) What were the perceptions of different stakeholders with regard to the implementation process and its outcomes?

(3) What was the impact of the reform and did it differ from the expected results of the AfL programme?

This chapter is divided into four main sections. The first presents the views of the drivers of change: the three Norwegian Ministers for Education who were in charge over the period of the programme, as well as an external perspective from the international community. The second presents an overview of implementation activities in nine schools across four municipalities, with a specific look at the involvement of key actors in the implementation process: teachers, students and the Directorate for Education and Training. The third looks at the perceptions of the implementation and outcomes of the policy from school leaders, unions, researchers and the media. The final section assesses the impact of the AfL programme by examining students' learning outcomes before and after the implementation of the first phase of AfL.

Drivers of change: Governing from the central to the local level

The view of the Ministers

Three Norwegian Ministers of Education were interviewed for this report: Kristin Clemet (H), Øystein Djupedal (SV) and Kristin Halvorsen (SV).[14] These three served under different governments and were responsible for different phases of the AfL programme and its preceding reform on Knowledge Promotion respectively. These interviews are, naturally, more political and ideological than the other interviews carried out for this report, but they share some common ideas about Norway.

In order to provide a better understanding of Norwegian reforms since 1990, when steering by objectives was introduced into the education system, key policy documents were taken into account. These include commentaries by, and interviews with, previous Ministers of Education, for instance Gudmund Hernes from the Labour Party who served from 1990–95. In addition, several documents written by further Ministers of Education who have served during the past 25 years were included in the analysis.

Table 4.1. Ministers of Education from 2000 in Norway

Year	Minister of Education	Government	Coalition[2]
2001-05	Kristin Clemet[1]	Bondevik II	H + KrF + V
2005-07	Øystein Djupedal[1]	Stoltenberg II	AP + SV + SP
2007-09	Bård Vegar Solhjell	Stoltenberg II	AP + SV + SP
2009-13	Kristin Halvorsen*	Stoltenberg II	AP + SV + SP

Notes: 1). Interviewed for this report. 2). H = Conservative Party, KrF = Christian Democrats, V +Social Liberal Party, AP= Labour Party, SV=Socialist Left Party, SP=Centre Party.

The current AfL programme was initiated under the three-party coalition government led by Prime Minister Jens Stoltenberg (Labour). In 2010, when the AfL programme was launched, Kristin Halvorsen (Socialist Left Party) had just come into office as Minister of Education after serving four years as Minister of Finance. She succeeded two other Ministers from the Socialist Left Party, Baard Vegar Solhjell (2007–2009) and Øistein Djupedal (2005–2007). The AfL programme was thus implemented under three different ministers who were all from the same party. In the interviews conducted for this case study, the Ministers highlighted continuity as a key factor in the facilitation and implementation of new policy programmes spanning such long timeframes. The Ministers also explained that there was broad consensus across the different political parties in the Norwegian National Parliament (*Storting*) on most issues that related to the implementation of the Knowledge Promotion reform.[15] This may not be the case in other democracies.

With regard to the question of how implementation strategies for educational reform in Norway (should) look, the three Ministers varied in their answers and cited a range of factors:

- The need for more substantial knowledge on assessment literacy and implementation strategies;
- The need for data to support governance and steering of educational programmes;
- The need to improve dialogue and information between central and local governments, particularly the capacity to reach out to small municipalities;
- Trust and dialogue in the Norwegian context as a main criterion of success for implementing educational programmes and reforms.

The interviewees also stressed the challenge of governing from the centre without the necessary power for leadership and control. In the highly decentralised Norwegian context, Ministers cannot really steer counties or schools, but rather can only offer help and support through the Directorate of Education and Training. If some educationalists refuse to follow centralized regulations, there are few consequences.

The three Ministers stressed the challenges in governing and implementation differently. When asked to reflect upon challenges and success criteria of implementation, the minsters offered the following reflections, presented below in chronological order of their time in office.

2001-05: Kristin Clemet – Prior to the AfL programme

Kristin Clemet, from the Conservative Party, served as Minister of Education and Research from 2001 to 2005 under a coalition government led by Prime Minister Kjell Magne Bondevik (of the Norwegian Christian Democratic Party). She initiated the Knowledge Promotion reform. Recently evaluated, results indicate that Norwegian students improved their learning outcomes in international tests after the reform, particularly low-achieving students, and that there are fewer variations between student outcomes (Olsen et al., 2012). It is important to stress that all political parties supported the main parts of

the reform, making it fair to say that Norway's tradition of governing through consensus also applies to education.

Kristin Clemet reported that when she came to power, she was struck by the fact that there was a lot of knowledge in the Ministry about structures, resources and economy, but when she asked questions such as "How do we know that students learn what they are supposed to learn in schools?" there was less knowledge and documentation available. She described the years from 2001 to 2005 as a time of real change, when the need for documentation of learning outcomes increased, while a greater focus on core knowledge within the school seemed to be increasingly accepted. Also, according to Clemet, it became clear during that time that teachers lacked assessment literacy. When she asked for documentation on how teachers assessed students in schools, she was given a variety of procedures. Some teachers used the normal distribution when marking students' work, some used specific standards and assessment criteria and some teachers generally praised their students and wrote comments such as "Well done, Per". According to Clemet, the lack of consistency in assessment was overwhelming.

Norway introduced national tests in 2004. In the early days, there was some resistance from teachers regarding the tests themselves. Clemet stressed that the education system at the central as well as the local level seemed to lack knowledge of assessment literacy. Despite the frustration, the government did receive positive feedback from teachers who attended courses held on the use of national tests. She attributes this positive feedback to the limited number of assessment courses available to teachers, which served to increase the trainings' acceptance and popularity.

Clemet also described the challenges involved in communicating across the entire education sector. She found that teachers tended to rely more on their own Teacher Journal (*Utdanning*) than on other sources of information. Clemet therefore drew on both national and international research, convinced that research evidence would work better than political rhetoric when communicating with teachers, politicians and the general public. She also suggested that the strong alliance between Teacher Unions, some researchers in Norway and the policy level had hindered change in schools – simply because this alliance had become too strong. But overall, she explained, Norway is a country where there is a tradition of good dialogue with the unions, something she had experienced herself.

Clemet further emphasized the need for strong top-down policy and leadership: "If there is no pressure, teachers will not necessarily change". Even though the Minister is unable to directly steer teachers, Clemet stressed the importance of being an active and visible leader, who would meet with teachers and school communities personally. As an example, she mentioned several meetings she had attended in Norway, gathering hundreds of students, teachers, school leaders, and municipality leaders in local schools, to discuss different issues in education. As a former Minister, Clemet fears that a lack of top-down steering can lead to more inequity across schools. She believes teachers and school leaders should be trusted to do their job, but also that they be held accountable for what they do as a part of broader implementation strategies.

2005–07: Øystein Djupedal – Pilot Study of AfL

The knowledge gap in assessment was also highlighted by Clemet's successor Øystein Djupedal. He suggested that more scientifically-based assessment literacy was needed, but believed that the real challenge of implementation was related to the following key issues: (1) teachers were tired of change after decades of school reforms under Gudmund Hernes (1990s), and then Kristin Clemet (early 2000s) and (2) the teachers' general resistance to change. The reforms had led to what Djupedal described as "teacher fatigue", that is, feelings such as "Another reform? It will pass, why should I care?" Djupedal gave the following example:

A researcher and teacher in higher education once explained to Djupedal that all the reforms he had experienced had had to be adjusted to his specific pedagogy, and not the other way around. Djupedal saw this as the main challenge for change in schools, the resistance from teachers to change – instead of trying to understand the aims of the policy reform, they would adjust it to their own thinking. In fact, Djupedal felt that teachers' resistance to any change was the most challenging part of implementation in Norway. He therefore suggested that in order to avoid teacher fatigue, the design of reforms had to be reconsidered and their number limited. Referring to another reform he had implemented at the nursery level, he believed it was possible to implement reforms in schools as well, but that it would be difficult.

2009–present: Kristin Halvorsen – The AfL programme is initiated

Teacher resistance to change was also one of the issues addressed by the current Minister of Education, Kristin Halvorsen. She promised that she did not have another reform to announce the first time she met teachers at an annual Teacher Union meeting in Lillehammer, in 2009:

> When I explained I would not reform school all over again, you could see that some of the teachers in the audience relaxed, they sat down in their chairs and were quite happy. Of course, then I was worried they had gotten the wrong message.

Halvorsen gave the following background information for her point of view:

> The Curriculum Reform Knowledge Promotion was initiated in 2006, three years before I came to office. It takes a long time to implement reforms and our country has too often implemented reforms, only to replace them with a new one. The challenge was that we had a situation where we needed to adjust during the implementation. So, we have adjusted the Knowledge Promotion reform with regard to several issues since 2006 until now.

Kristin Halvorsen further focused on how implementation could be carried out in concert with different stakeholders. She had various meetings with researchers, school leaders, teachers, students and other stakeholders to discuss education topics such as assessment, bullying and motivation. These seminars could be seen as part of an implementation strategy aimed at getting people involved in school development. Halvorsen also reported that a further priority of hers was the importance of learning together; she stressed the need to increase assessment literacy at all levels, from the central government and the Directorate, down to the local municipalities. She therefore described the experience of learning about the ideas in Assessment for Learning, and the use of feedback, as a positive strategy and motivating in itself, and something she believed in.

Kristin Halvorsen further explained that she believed that Norwegians had one advantage in that all stakeholders in Norway attend the same type of public school, and thus share the same experiences and background. Institutions such as "Oxbridge" equivalents or private schools from which stakeholders could be recruited, do not exist in Norway. Instead, stakeholders are all engaged in the same public school, which she believes is of importance. When people discuss education around the dinner table in their homes, they more or less talk about the same school experience.

As with her predecessors, she was worried that a few schools were being left behind. According to Halvorsen, there are a few municipalities that do not seem to follow policies formulated on the central level, such as the AfL programme. The government has offered these schools guidance by trained teams, and these initiatives focus more on support than control. In this context, Halvorsen stressed that "all pupils in Norway have a right to good teaching, no matter which municipality they live in".

Box 4.1. Evaluation of Norwegian municipalities

A recent evaluation of Norwegian municipalities (Rambøll, 2013) shows that there are several municipalities that do not follow up on centrally initiated regulations. The challenge the three Ministers have experienced in reaching all municipalities in Norway are therefore still *the* main challenge. These recent results echo an OECD (2011) report, which summarized the challenges for Norway and the decentralized school system, and the challenge of steering schools. The findings also reflect the challenge of communicating with the entire sector.

Although the Ministers interviewed are members of different political parties, they have addressed one common challenge for governing from the central level: The Minister of Education is one of the most high-profile Ministers in the government and very visible in the media. Despite this visibility, it is challenging to communicate directly and effectively with teachers and get the message of political reform through. Ministers have no power to influence or steer schools at the local level, or control whether central initiatives are followed. The different counties and municipalities effectively own the local schools, and they have the power to decide how their schools are run. So, if Norwegian schools are criticised, the Ministers are given the blame in the media, but they in fact have little power to influence what is going on in each county. As an example, the current Minister of Education did not want municipalities to rank their schools by test scores from the national tests, but this is now taking place in several municipalities governed by politicians from other parties.

An international perspective: Governance, PISA, and the changing role of teachers

This challenge was also commented on by Andreas Schleicher, Deputy Director for Education and Special Advisor on Education Policy to the OECD's Secretary-General. In an interview, he pointed out that Norway "is a country where the Ministry of Education has really very limited control over what happens in the classrooms". Schleicher further argued that when the PISA results were published in 2001, it was the first time there had been any kind of objective reflections on results in Norway.

> Yes, I think the high degree of autonomy of teachers and schools is not matched by transparency and the accountability culture, in the sense that nobody questions teachers in the classroom and in the schools, and the Ministry did not have very strong mechanisms to look at and improve performance. One of the positive aspects of the Norwegian system is the high level of trust, but the downside of this is that it is very hard for the teachers to understand where they stand and for the schools to know where they stand. You see that reflected […] in the eyes of students: students who rate their own performance as much better than the performance actually is, tells me that the teachers are telling them they are doing fine when their results are not really good.

It must be noted, however, that the question of whether PISA and the priorities of the Norwegian education system are consistent have not been analysed thus far, and as a result it is hard to know whether its results are nationally relevant or not. This is pointed out by Grek when she indicates that international testing systems tend to be centred on basic and functional skills for future workers rather than less concrete/measurable aspects of education such as "democratic participation, artistic talents, understanding of politics, history, etc." (2010, p. 399), which may be relevant goals for local governments. In addition, there may be reason to expect a displacement when shifting from the central level, with its focus on basic skills, to teachers, who tend to focus on issues related to personal relationships with students. Indeed research has indicated that teachers tend to experience personal relationships with students as the core of their professional work; while policy changes that aim at influencing teachers' work are regarded as less important (Jensen, 2008).

Getting teachers on board with new policies is key to successful implementation, but no change happens in a vacuum. Previous research has shown that teachers worry about change that involves students' and teachers' roles (see Chapter 3). In the interviews with teachers for this study, several described feeling exhausted and stressed out when facing new initiatives, even in the pilot school where teachers had already been working with AfL for the past six years. Some teachers indicated that they were too stressed to cope with their daily life, and therefore, changing their teacher role was not something they felt they were able to do.

Box 4.2. The changing role of teachers

A study in Norway exploring job satisfaction and burn-out of 2 000 teachers revealed that even if the majority of teachers are satisfied with their job, one out of three say they would not choose to be a teacher if they could do it all over again (Skaalvik and Skaalvik, 2007). When Ministers discuss how to reach out to the teachers and implement new programmes, then, they also need to pay close attention to the difficulties of teacher recruitment and retention

On a positive note, there are key characteristics of the Norwegian context that are generally conducive to good governance, Andreas Schleicher reflected upon the dialogue between teacher unions and the government as "very impressive" and "very constructive discussions". This perspective was echoed by one of the ministers, who claimed that overall there are good relations between the organisations and the partners in the working force. The Minister further qualified the Norwegian system as transparent, efficient with little corruption, and an open economy. One of the ministers explained that Norwegian society is characterized by the need to agree. "We have a consensus model, where we argue and have debates, but overall we pretty much agree on the big ideas". In this respect, good communication and high levels of trust make it easier to implement new programmes, and this can be seen as one of the key-findings from this work. This, and teachers' resistance to change, will be discussed in more depth in the following sections.

Implementing AfL in schools: How successful has the programme been?

The four principles of the AfL programme build on the central premise that students should know about the learning goals, assessment criteria and be actively involved in the learning and assessment process. How successful has the programme been in communicating these goals to the relevant actors (in particular students, teachers and school leaders) and engaging them in the implementation of the programme goals? The following section sheds light on these questions by presenting an analysis of qualitative interviews with teachers and students that looks at their assessment literacy and level of involvement. As these are self-reported data, the results should be regarded as reports on beliefs on assessment and classroom practice, rather than evidence of what is actually happening in the classroom. There are no classroom observation data to confirm these beliefs. Still, the analysis of data from four different groups of key actors – school owners, school leaders, teachers and students – allows for a triangulation of evidence from different sources and provides a comprehensive picture of how those who are most affected by the policy programme are involved in its implementation.

All schools are generally obliged to implement the four AfL principles in their practice as they are part of the regulations of the Education Act. However, the schools participating in the AfL programme work on these principles with the intention of making them an integrated part of their teaching practice. In order to focus the implementation phase and help the schools efficiently improve their assessment practice, they were allowed to choose to work with one or two specific principles to begin. In addition, the school owners were advised to do an analysis of their current assessment practice and development needs before they decided on the content of their local AfL programme. Table 4.2 demonstrates the extent of the involvement of each of the participating schools in the AfL programme as reported by the school leaders,

teachers and students. A coding system indicates the degree to which the participants interviewed had knowledge of and were actively involved in AfL strategies in each of the nine schools, which is taken as a proxy for the extent to which AfL was implemented in the particular school. Examples of evidence for implementation include:

(1) students explaining how they use peer or self-assessment;

(2) descriptions of how teachers share assessment criteria with the students;

(3) descriptions of how teachers adjust their teaching as a result of students' feedback;

(4) school leaders explaining AfL principles with specific examples from their schools such as internal courses, teamwork or conferences.

School leaders, teachers and students were asked to share their knowledge and experience of AfL practices in the classroom concerning the four principles laid out by the Education Act (see Interview Guide in Annex C for a full methodology and list of questions). Answers were grouped according to the number of principles respondents could explain or relate to, as follows: ++ indicates that respondents were able to explain three or four of the AfL principles; + indicates that one or two principles were described; 0 indicates no reference or description of AfL practices.

Table 4.2. Implementation of the AfL programme in nine schools

	Municipality A Urban		Municipality B Rural		Municipality C Urban		Municipality D Rural		
School ID	1	2	3	4	5	6	7	8	9
School leaders	+	++	++	+	++	++	++	++	++
Teachers	0	++	0	+	++	++	+	+	++
Students	0	++	0	0	++	++	0	++	++

Notes: ++ strong evidence of AfL practice (three or four principles represented), + some evidence of AfL practice (one or two of the principles represented), 0 indicates no evidence of implementation of AfL.

As can be seen in Table 4.2, the extent to which AfL practice was perceived to be implemented differed widely across municipalities and also across respondents. In Municipality A, for example, School 1 showed weak evidence of AfL practices while School 2 demonstrated strong evidence of AfL, including at the student level. This is particularly interesting as Municipality A was the only municipality that was not officially taking part in the AfL programme at the time.[16]

Schools in Municipality B fared better, but notably the student respondents did not report any examples of AfL practices. In contrast, the respondents from schools in Municipality C uniformly reported evidence of AfL practice, while again the picture is more mixed in Municipality D. One key variable is the length of time in the programme. Several authors see the implementation of AfL as a slow and complex process that requires time (see for example MacPhail and Halbert, 2010; Hayward and Spencer, 2010; Willis, 2008). While this might explain some of the responses from students, it does not tell the whole story: at the time the interviews were conducted, schools 3, 7, 8, and 9 had all been participating in the AfL programme for less than a year (see Table 1.1 for the length of time in the programme), and yet their responses vary substantially.

These results demonstrate that even within the same municipality the extent to which the AfL programme was implemented differs. These varying implementation experiences will be discussed in the following sections, by looking at governing practices across the four municipalities as well as the perspective of various actors in the system.

Governing practices in four municipalities – Knowledge transfer is key

School leaders' experiences with the process and final set up of the AfL programme in the four municipalities varied considerably. Even though the school leader in School 1 struggled with the implementation of AfL, the support from Municipality A was acknowledged by the principal as being of value. However, in Municipality B, which is in a rural area, school leaders lacked municipality level support. They had to deal with budget cuts and reorganization processes, resulting not only in staff reductions at the municipality level, but also frequently changing leadership: in the past few years, the municipality had six different leaders. At the same time, some of the most qualified people obtained new positions elsewhere. School leaders underlined that the struggle of keeping qualified staff in important positions resulted in a clear loss of competence for the municipality, which also led to shortcomings in the communication between the various levels. For example, one of the school leaders had only had two meetings with his municipality leader over the last 10 years, despite these meetings being compulsory by law. The two school leaders in Municipality B emphasised that they needed more support and that in fact they felt left alone by the municipality with organising the implementation of the AfL programme.

These examples show that there is a significant difference in how municipalities organise the implementation of the AfL programme, especially when it comes to supporting the school leaders in tasks and processes necessary for the programme. Some municipalities support their school leaders with a consistent and constant dialogue (for example through conferences and other meetings), while other municipalities do not have the capacity to accompany the school leaders and steer them through the programme implementation.

For actors struggling with implementation, the AfL website developed by the Directorate (see Box 4.3) was perceived as a helpful tool. School leaders in Municipality B appreciated this platform and used it for example to communicate the programme in meetings with parents. They also planned to use the website for internal discussions with teachers as part of a knowledge development process.

Box 4.3. The website Vurdering for læring: A tool for implementing new assessment practices

The AfL website developed by the Directorate contains material from both international and national scholars, articles and book suggestions in Norwegian and English, and material specifically targeted at school leaders to use and stimulate reflection in their schools.

The website also contains keynote presentations from international and national scholars as well as video footage of students explaining how they experience the practice of AfL. It also hosts an online platform that documents the outcomes of a conference series for teachers organised by the Directorate.

Actors that have used these online tools considered them a valuable resource for knowledge sharing and connecting both practitioners and the public.

Source: www.udir.no/Vurdering-for-laring/

One of the school leaders in Municipality B participated in the assessment conferences arranged by the DET, and was very pleased with the support offered there, especially in light of the lack of assistance provided by the municipality. Both interviewed principals saw it as a challenge that no particular staff member at the municipality level oversaw the implementation of the programme, especially if AfL was to deliver sustainable outcomes. There was no steer for the principals as to what would happen after the schools had finished the implementation phase. Both school leaders suggested that they needed three to four years to change the practice in their schools, and therefore in their eyes the critical period for AfL would start only then. These results are in line with those of the study by Gardner et al. (2011) which

concluded that, along with impact, sustainable development was one of the aspects that was not properly considered during the design of the AfL programme.

A further pitfall in implementing the programme was that even though tools for knowledge generation and transfer had been provided, they did not always reach the targeted actors. For example, one of the school leaders interviewed was not aware of all the keynote presentations on the Directorate's online platform. In small municipalities, knowledge gaps like this can be particularly problematic as respondents from these communities expressed much more need for support from the DET than the ones in bigger communities. In Municipality A, for example, which had a large group of staff (approximately 50 members) working on quality improvement in schools and kindergartens, school leaders and teachers expressed almost no need for additional support from the Directorate.

In Municipalities C and D, all school leaders valued the support from the municipality level and qualified it as useful for the implementation of AfL. For example, county C had participated in work on assessment for more than a decade, and the schools used this experience when working on the AfL programme. Evidence of AfL was found in all interviews conducted in this county and there appeared to be a continuous exchange between the three different levels (municipality, school leaders and teachers) on assessment issues. In County D, three school leaders spoke about the value of having a municipal leader with whom they could talk about the challenges they faced and who also had a professional background in education. In fact, this municipal leader was very engaged in the implementation of the programme: regular meetings for school leaders were held to discuss the schools' progress, and bigger conferences facilitated exchange among all teachers and school leaders of the county.

In contrast, the schools in Municipality B lacked the support available in the other counties. The tools developed by the Directorate and their easily accessible material appeared to help fill in this gap and were therefore key to the successful implementation of the AfL programme across all municipalities.

From the empirical data collected for this report it becomes clear that not all participating municipalities and schools delivered the expected results when participating in the AfL programme. The research also suggests that these differences are due to a number of reasons, including governance, communication flows and the varying practice of capacity building for schools across municipalities.

Two different perspectives on teacher resistance to change

In addition to capacity issues, resistance to change can be a further obstacle to the implementation of policy reform, and the interviews conducted for this report indicate that this has been a factor for the AfL programme. The following section will present this finding in more detail.

Municipality A is not an official member of the AfL programme, but its schools participated in the study *Better Assessment Practice* (2007-09). As a follow-up, the municipality developed strategic plans for assessment in order to implement the four AfL principles. Situated in a big city, the two schools in Municipality A generally have the possibility of support from advisory staff at the municipal level. However, the two school leaders had quite different experiences of municipality level support for AfL.

In School 1, where there was little interview evidence for AfL implementation, the school leader was not quite sure whether the school should participate in the project. The municipality had been compiling strategy papers on assessment since 2009 and the school leader had actively tried to involve the teachers in this area. The main implementation problem occurred when teachers attended a large conference on assessment, after which disagreements emerged over whether or not they were supposed to share their

experience and knowledge gained with the rest of their colleagues. It became an open conflict, since the teachers refused to take on the role of change agents by sharing their experiences with their colleagues.

The school leader explained that it was hard to change a school culture against a resistance that was based on teachers' satisfaction with the quality of their own work and their focus on their own rights. Interestingly, according to the school leader, the municipality – which had organised the conference – did not expect the teachers to share what they had learnt at the meetings with their colleagues. This left the school leader with no real support when trying to facilitate a knowledge transfer within the school. When asked about support from the municipality level, this school leader did not feel supported in the implementation of AfL.

When planning further AfL implementation, School 1 and Municipality A took into account these experiences from the municipality's assessment conferences. This further shows that if a municipality is unclear about expectations for school leaders and teachers when offering conferences, it can lead to participation with no commitment to change as no-one feels responsible for following up the work.

In general, the school leader in School 1 also expressed a feeling that the municipality level was more concerned with assessment defined as results, or "what can be counted," and less about AfL practices *per se*. The main target, set at municipality level, was to improve students' reading skills, so the reading programmes were allocated more time than the AfL programmes. This is also one of the main findings from struggling municipalities that were included in this report: they do not seem to be able to co-ordinate competing programmes initiated by the Directorate, and so sometimes AfL is not prioritised due to lack of capacity. Both school leaders and teachers express concerns and a sense of guilt when they are not able to work on AfL, as they know it is a part of the Education Act. Despite the shortcomings listed above, however, the school leader from School 1 was satisfied with the municipality-level support in general.

School 2, which showed strong evidence of AfL practice, is run by a school leader team, which consists of the school leader, a head of department and two teachers. The school leader stressed the importance of acting as a team, which reflected also the way work was organised throughout the entire school, where a group of teachers would usually work on new developments proposed by the school leader. Even though this school has not formally been a part of the AfL programme, the school has been working on assessment for more than six years. It is also a common practice to send teachers to conferences or seminars and then let them train the rest of their colleagues. If resistance to change occurred within the community, a "trial-and-error"-strategy was applied, in which teachers would try out new practices and report back to the school's leader team on their experiences in order to inform a continuing dialogue on the change process. The interviews conducted for this report show that the school leader team in this school had a clear expectation of knowledge sharing between all teachers, which helped to create an atmosphere that was more open to possible change. The focus on peer learning and dialogue, the integration of all teachers in the change process (e.g. by organising pre-planned visits to classrooms), as well as the willingness to adjust the implementation strategy resulted in all teachers participating in AfL practices.

The school leader team reported that after six years the teachers were still working on how to improve their assessment practices, but were more confident in what they were doing. The teachers of School 2 had also visited other schools in Norway and shared their experiences with them. The school leader stressed the importance of working together with teachers, being patient, persistent, understanding and having regular communication. In addition, the school leader team was planning to use what they called "school walks", or pre-planned visits to classrooms, to see how teachers teach and use assessment. The teachers were to be involved in deciding what the leaders would observe and the focus of discussion of what the teachers achieved. Again, this demonstrates school strategies for involving teachers and a focus on dialogue.

Another factor that facilitated the implementation of AfL practice in this school was its very good relationship with the municipality, possibly due to the school leader's having worked at the municipality level for two years and still knowing staff there. Communication was easy and the school leader team stressed that the help of the municipality and the expertise of the staff there was instrumental to the success of AfL in the school. In addition, the school leader had been involved in work on assessment from the beginning of his career. The school leader acknowledged the work and support from the municipality level, and attributed part of the school's success to this. The school leader team in School 2 saw themselves as privileged to work in a municipality with a number of experienced people at the municipal level that could support them.

As illustrated by these examples, even within the same county, two school leaders can experience support from the municipality and teachers' resistance to change quite differently. This has a clear impact on practice, as there was little evidence from the teacher interviews and no evidence from the student interviews of AfL practices in the case of School 1, while in School 2 there were several examples from both students and teachers of AfL practices. School 2 is considered successful because of the organisation of its development programmes. They had been working with assessment for more than 6 years, and they had linked experience to new central initiatives. Even though the school was not part of the official AfL programme, it had been able to pick up centralized ideas, implement the new Education Act and develop its own practice. The school leader team characterises their approach as active engagement of teachers in the development programme, with changes made through dialogue as well as through trial and error.

Even though the school leader in School 1 struggled more with resistant teachers, the interviews with the municipality level and the school leader in School 1 both reinforce the idea of a long-term plan to implement AfL. In this longer-term view, local variations are expected and accepted, and perseverance and teachers' development of competences over time are considered to be crucial elements. The school leader in School 1 summarised this as follows, "In our school, we have to move with small steps – if not, we will not be able to move forward at all."

These results indeed support the call from many authors for more involvement of all actors in the policy processes (Barret, 2011; Bunar, 2011; Berryhill, 2009; Dobbins and Martens, 2012; Sookrajh and Joshua, 2012). The involvement of key actors is important to avoid possible resistance when it comes to implementation and a subsequent lack of sustainability of policy reform. The more hierarchical understanding of the policy cycle observed in School 1 seemed to hinder the enactment of new practices, while the more dialogic-oriented, interactive approach in School 2 seemed to lead to a more sustainable implementation. This is in line with Blackmore's (2011) argument about the "opportunities of governance," wherein more room for horizontal collaboration between actors will help to better develop and implement policy reforms in the field of education.

Student participation in AfL practices

In general, the implementation of AfL has been a complex endeavour because the programme needed to be adjusted to the particular needs of each school in order to have the best impact. An excellent example of this comes from the three schools in Municipality D. As these schools only recently joined the programme, pre-existing differences in teacher practice and assessment culture can be more easily observed, as well as the impact that has on their respective abilities to implement the AfL programme.

One of the expected results of AfL is that students should be more involved in their own learning and assessment. In the interviews, students were asked to describe and explain what they knew about assessment in general, and how they were involved in the learning and assessment process in their schools. In School 7, students (age 11) did not provide any examples of AfL practices but explained more generally

how their teachers assessed their work by talking to them in class and offering help and support. When asked to explain experiences with assessment, they provided examples of more general written comments from teachers such as, "Good, but you should have practised more." When asked specifically about whether they had any experience with assessing each other, one of the boys said yes, he knew about peer-assessment. He then went on to explain that he had tried it with his little brother at home, but it did not work since the brother was only two years old.

Students also spoke about a special reading programme where the teachers discussed their progress, and results from tests were shared with their parents in the annual parent and teacher conferences. Despite the fact that they did not report AfL practices, these students were really happy with their school, believed their teachers were very nice and kind and they loved all the activities they were able to do outside in the schoolyard. When the researcher asked this group of students if they had ever heard of Assessment for Learning, they said yes, but only in the briefing given by their teacher before the interview. They then told the researcher that the teacher had tried to explain what AfL practices were, so they would know when interviewed by the researcher. Apart from this student group, there was no indication that their teachers or school leaders had influenced students before their interviews.

In School 8, teachers interviewed for this report showed no clear understanding of assessment and AfL, even though the school leader did. Interestingly, despite this lack of knowledge from the teachers, students were able to share knowledge of AfL using practical examples both for peer-and self-assessment. It became clear that this knowledge originated with one of the language teachers who had not been interviewed. A single teacher thus made a difference for implementation at a school. However in order to avoid erratic outcomes, a whole-of-school commitment towards AfL implementation is necessary (see also Ofsted, 2008; DfES, 2007; Tapan, 2001; Kellard et al., 2008). As discussed in Chapter 3, creating a culture of assessment is necessary to generate change throughout an entire school. School leaders are seen as a crucial figure in AfL initiatives (Hayward and Spencer, 2010; Willis, 2008; Gipps et al., 2005; Webb and Jones, 2009; Kellard et al., 2008, among others), and for them the challenge seems to be how to get all teachers involved, and how to build collaboration within the school.

School 9 is an example of a school that was able to build such a whole-of-school approach. Students (age 12) reported that their teachers always endeavoured to share the goals of the session and what kind of criteria they would use for assessment. Students believed the criteria helped them to know what to include in their tasks and therefore found it easier to deliver them. They also explained that they used peer-assessment in written tasks, and how they were allowed to comment on each other's work based upon commonly set criteria. Students thought this was "really fun", but agreed it sometimes could be challenging to assess friends, since they did not want to be too harsh. On the other hand, they appreciated getting feedback from classmates, since it was sometimes easier to understand their language than that of their teachers.

When asked whether it was challenging to do peer-assessment, these students shared an example of how one of the teachers had developed special cards they could use when they worked in groups. These cards would guide the students and help them to ask relevant questions, support their different roles in group-work, and facilitate the process so that it would be easier to assess each other. One of the students would act as a group leader, telling the others what to do, based upon these cards, and students would then participate in the activity guided by the cards and the group leader. This was very much appreciated by the students.

It is important to note that these three schools had completely different needs when they started the AfL programme. School 9 already had a transparent assessment culture, where the school leader, teachers, and students shared knowledge of assessment practices. Their needs for further development were therefore naturally different to those of School 7, where students in the same age group did not know what

AfL was, or how they could use learning goals and knowledge of assessment criteria to improve their own learning. This can be seen as an indication of the importance of capacity building. In some schools, teachers and school leaders had the capacity to improve their practice, while other schools do not seem to have the knowledge or capacity to transform good practices supplied by the DET to the local level. It was expected that teachers would change their practice and work towards integrating the AfL principles into their teaching practice and to involve students in the assessment practices, but the interviews indicated that changing practice was still one of the most challenging elements for teachers.

The Directorate of Education and Training (DET) as an implementation agent

The Directorate of Education and Training, in close co-operation with municipality and central government leaders, has been a main stakeholder involved in steering and implementing the AfL programme. When working on its implementation, DET staff explained that they had taken the position that they needed to learn together with teachers and school leaders. They did not see themselves as experts brought in to tell teachers what needed to be done. Most teachers interviewed perceived this to be helpful. The interviews demonstrated that the more closely teachers and school leaders worked with the DET, the stronger this perception was. Some teachers and school leaders even said that due to the AfL programme they had a better impression of the Directorate than ever before.

School leaders appreciated the way the DET guided the schools throughout this programme. Some school leaders stated that they felt that, for the first time, someone from the central level was actually interested in what they were doing. One school leader was also actively involved in a DET resource group. This position gave him the opportunity to have regular dialogue with the DET, and he believed that they listened to what he had to say. It was not a solely top-down process, but a real dialogue.

The challenge was that the less contact teachers and school leaders had with the DET, the more negative attitudes they tended to express in interviews. Some school leaders questioned whether the DET knew enough about "the real world" and believed there was a mismatch between how they as school leaders assessed teachers' levels of competency and what the DET thought they were. Some suggested that there had been a lack of clear leadership from the DET throughout the programme, and found it problematic that participation was voluntary. Some municipality leaders asked for better advice from the DET regarding which experts they could use when seeking help for the AfL programme. Small municipalities did not always have the resources to find external experts for their project, and they would have preferred stronger leadership and advice in this matter.

It is clear from the data that participants in the programme had mixed feelings about schools working together with the Directorate. A DET staff member explained how they had been told that some teachers did not have a good relationship with the Directorate, and therefore, it was not always wise to inform the teachers that the AfL programme had been initiated by the Directorate. Indeed, the DET had been told by some school leaders specifically not to tell their teachers that the AfL programme was a Directorate initiative, since they wanted to avoid resistance from school staff based on this information. As a member of the Directorate said, their involvement could either hinder or help implementation in some of the schools, depending on the teachers' view of the DET.

One implementation strategy that was clearly successful was the invitation of foreign experts to Norway. The DET invited various researchers from the UK to Norway in order to facilitate an exchange on their experience of implementing AfL in Scotland and England. The experts were chosen for their status (all of them were professors of education) and their more than 20 years of expertise in assessment issues. In comparison, most Norwegian researchers have no more than a decade of experience in assessment research specifically.

The conferences with these foreign experts were a significant success and the researchers seem to have inspired Norwegian teachers to start using AfL in the classroom. Evaluations by teachers praised the experts for being able to share their knowledge in a way that they understood. In the present study, it seems that some of the international researchers in particular have been key reasons for the success of the implementation according to the DET. They were described as "being a great success", "our mentors", "helpers", and "advisers."

At the policy level, it was acknowledged that having an outsider talk to teachers was one of the most important factors for the success of AfL. The external experts were able to broach difficult or "unpopular" messages more easily than representatives of the central government. In an interview, one of the foreign experts, Professor Gordon Stobart, explained that one of the reasons he was so involved with Norway was that he found Norway to be a potentially better place for Assessment for Learning. Norway does not have a testing-based accountability system that tends to undermine some of the good AfL practice: "… in Norway there's a freedom to say: let's just get our teaching and our informed assessment better, which is a huge opportunity."

Despite this success, the use of foreign experts did mean that except for a few Norwegian scholars, most of the Norwegian researchers working within the field of assessment were not invited to be involved in the official AfL implementation phase. Several of the participants from the DET explained that they did not think the Higher Education sector was able to offer expertise in the field of assessment, despite a great improvement in recent years (i.e. from 2007 to 2012). Their view was that Higher Education needed to develop their assessment expertise together with the participating schools in the project. However, the DET has increasingly included Norwegian researchers in the programme. They also acknowledged that more Norwegian literature was being published which they were then trying to incorporate into the programme, but still, a concern was expressed that there was not enough knowledge of assessment among Norwegian researchers. The interviews for this report showed clearly that this led to tensions between researchers and practitioners as well as with the Directorate.

These tensions revolve around questions such as: Who has the mandate to guide schools? Who decides what is good assessment practice in the Norwegian context? Who will supervise the schools when the foreign expert has travelled home? The section on perceptions of key stakeholders will look at this issue more closely.

Knowledge production and use

As an implementation agent, the Directorate for Education and Training has focused on building the necessary capacity at the local level to implement the AfL programme. This section will look at how this has been done and if the DET has succeeded in developing the infrastructure needed for knowledge development.

As mentioned in Chapter 3, the website developed by the DET has been a key part of the capacity building efforts. The participants who used it found it very helpful, but not all participants were aware of the material that was available. In addition, the dialogue between the participating municipalities and schools worked better in some cases than in others, and it was not always clear why some of the municipalities were not able to use the material offered or to engage more actively in knowledge development. The data collected for this report shows that teachers' level of trust in research and their views on the learning abilities of students as well as schools' experiences with previous development programmes were crucial factors for successful knowledge production and use.

In Municipality A, school leaders were concerned with the negative attitude some of the teachers had towards research and a subsequent lack of research-based knowledge. One of the school leaders also explained that some teachers tried to explain students' problems with learning with general cognitive and non-cognitive inability, rather than looking for ways to help their students. This shows that the implementation of AfL is not only connected to new methods and strategies for teaching, but also to changing teachers' attitudes towards the adjustment of deeply rooted professional knowledge.

All schools that had previous expertise with assessment programmes, such as the study on "Better Assessment Practice" from 2007 to 2009, were also able to better use the knowledge provided to implement the AfL programme and adjust it to their needs. When interviewing teachers in these schools, some common elements were clear: all schools had been actively involved in different assessment projects before the start of the AfL project, they had a history of being active in school development, and new projects from the government were easily built into their strategic school plans. They were also actively looking for ways to take part and shape new initiatives. In School 2, for example, the school leader had the opportunity to influence the development of the project at the central as well as the school level. S/he was a member of a group asked by the DET to work out new strategies for the AfL project and another project around classroom instruction. This school leader did not experience the AfL policy as being implemented top-down; instead s/he was able to influence the process through a bottom-up approach.

The importance of having already worked to build a culture of assessment was important. Schools 2 and 6, for example, had undergone a process of trial and error in terms of implementing the Better Assessment Practice programme. In School 6 teachers reported that they had worked very instrumentally to get the process going, and then after a while they were able to identify what kind of professional competences they had to develop. The problem, back in 2007 and 2008, was that they did not know where to find knowledge about assessment. Even though there were a substantial number of international articles and books on assessment, and they had attended the DET programme, they were not able to access this international literature on their own. There was little work in the Higher Education sector in Norway in this area, and so very few Norwegian research articles had been published. Teachers believed the situation had much improved at the time when AfL was initiated, with more support available from Norwegian research.

Another school with prior assessment experience selected a different approach to develop the necessary knowledge for AfL implementation. School 5 first focused on theoretical ideas of how to implement new assessment practices, and then visited schools in Scotland to collect practical input on how to improve teaching and assessment. Several of the schools analysed for this study had also cooperated with other schools in the municipality, and were developing networks to share knowledge and practice from their experience in the project. These structures can be seen as signs of successful implementation strategies, and some of the schools were able to build a sustainable environment for AfL practice through these networks.

It is important to note, however, that the schools analysed for this study varied greatly, from being well organised and with extensive knowledge on AfL, to being less organised with teachers and school leaders who did not know much about the principles of AfL. Only a few schools were active and committed to engaging with all students. These schools seemed to have entered into a continuous and well-organized practice of developing professional knowledge where they made use of research knowledge, theories of learning, and shared their own concrete experiences of how to teach and assess. Most schools did not operate like this, and in those that did, the processes seemed to be slower and characterized by opposition from at least some of the teachers. Some teachers even reported that they struggled to understand the correct way of doing AfL, indicating that they believed there must be a "recipe" for good practice. The uncertainty about a "correct way" was found in all teacher interviews,

suggesting that more time was still needed to develop a deep understanding of the scale of change in practice that AfL requires.

A key question from these results is therefore, how does the central government steer hundreds of schools when the variation among these schools is so large? Xaba (2011) indicates the importance of understanding what kind of knowledge actors need before a policy reform is enacted in order to avoid ineffective implementation. Knowing where the schools were as a starting point was a key issue for effective implementation, especially important so that the capacity-building element of the programme could be adjusted accordingly.

Perceptions of key stakeholders on the AFL programme

School leaders and the role of networks

The nine school leaders interviewed for this report represent a diverse group. One school leader emphasised the importance of being trained specifically for the role of school leader, but until the early 2000s, Norway did not have any particular development programme for school leaders. Since then, national Masters level courses have been developed, and some of the school leaders interviewed had participated in these programmes.

In general, the school leaders stressed the importance of developing a spirit of "We-schools" (*Vi-skole* in Norwegian), emphasising teamwork within the school community. The OECD's Teaching and Learning International Survey (OECD, 2009) highlights the importance of team teaching in Norway. Teacher teams offer opportunities for teachers to reflect on their work and discuss their teaching practices. Team organisation in Norway usually involves a team of teachers sharing the responsibility for a group of students, generally from the same year group. School leaders talked about the strengths of leading teams of teachers, instead of individual teachers. One school leader in a primary school argued it would have been impossible to implement the AfL programme if the teachers were not organized in teams. The teacher teams also offered opportunities for teachers to reflect and talk about their teaching practice, even if some of the interviewed teachers admitted that much meeting time was used for administrative discussions.

This way of managing teaching staff seemed to be instrumental for the implementation of the AfL programme at some of the analysed schools, in that school leaders conferred special mandates to teams of teachers responsible for implementation. These teacher teams were often in charge of school development in general and supported the school leader in implementing policy changes.

When discussing relationships with teachers, all school leaders stressed the importance of including teachers in the school's strategy, and of spending time on involvement and participation. One school leader explained the importance of giving teachers good explanations of why the school needed to work on AfL, before explaining how he had used policy documents and research articles as evidence for why this was important. He emphasized the importance of treating teachers with respect since they are highly educated and will not do something just because the Minister has introduced a new programme. Overall, the school leaders described practices that emphasized participation, sharing responsibility and cooperation.

Several of the schools had also cooperated with other schools in the municipality, and schools were developing networks to share knowledge and practice from their experience in the project. These structures can be seen as signs of successful implementation strategies, and some of the schools visited were able to build a sustainable environment for AfL practice through these networks. On the other hand, school leaders demonstrated in the interviews that they needed to work with teachers and convince them of

the importance of school development such as AfL, as school leaders cannot offer any incentives to persuade teachers who would otherwise not want to participate. The "team leadership" and the relationship between the school leader and teachers can therefore be seen both as a strength and challenge for the implementation of programmes such as AfL. When teachers agree on new initiatives, their teams and networks can be an efficient structure for implementation. If teachers resist the programme, the same spirit can be a major challenge for the school leader.

A further challenge for the implementation of AfL was the retention of key personnel, particularly in small municipalities. One school leader explained how the local municipality level had cut back on central positions, and stakeholders with knowledge in the field had taken up new positions. There were also reports of several job changes among central leaders during the implementation phase, which was not easy for the participants interviewed. One school leader reported that it was hard to know which person to call if problems occurred, and as a result, he felt that no one was available to help him.

The teacher union perspective: Control vs. trust

The views of the teachers towards AfL in general, expressed through the National Teacher Union, varied greatly. One of the leaders in the National Teacher Union explained that members assessed the policy as being successful because it addressed key issues for the teaching profession: teaching practices and student learning. Then again, there were also teachers who did not like the new programme and felt it was too much work on top of what they were doing on a daily basis. Some teachers also feared that AfL would narrow students' learning, and they worried that attainment goals and criteria would reduce learning outcomes to the knowledge of mere facts. The union representative confirmed that the Directorate seemed to be more appreciated by the teachers involved in the AfL programme. In a way, the reform helped the Directorate to improve their status and gain trust among many teachers. These teachers did not experience the programme as a top-down measure, but instead believed that that their opinions were being heard and taken into account.

On the other hand, the union representative stressed the fact that in some municipalities school leaders and teachers did not dare to clearly express their views or disagree with the municipality level. In addition, some school leaders were not fully aware of the programme, for example stating that they did not know how to use the DET website and tools. Therefore, the Union was concerned that some municipalities could use AfL for control instead of trusting teachers to do their job.

The voice of norwegian researchers: Conflicting views on the AfL programme

Five Norwegian University and University College researchers from across the country with substantive experience in the area of assessment, including supervising schools, were interviewed about their views on the AfL programme. When asked about the AfL programme, expected results and communication, there were some positive comments welcoming the initiative and how it was implemented. As one of the professors said:

> We were so pleased that a project with Assessment for Learning actually was launched. It could have been a test programme, how to become better test producers, I mean, in Germany and the States, they have let this field being taken over by the private testing firms. Well, this is not the tradition in Scandinavian countries. (Professor 1).

This professor saw the implementation of AfL as a result of the dialogue between leading experts in assessment, pedagogical experts in the field, and the tradition of what kind of school is wanted in Norway.

However, other respondents were less positive. One professor focused more on the challenges and argued that the major problem with DET's implementation strategy was that it did not take the variation between schools in terms of their knowledge of assessment into account:

> What level are the schools starting at, and what level are they going to end at? Some of the schools included in the sample now, have already been working on AfL for two or three years and are quite advanced. And you have to put up with people being burned out. It is not so much that they are forced to do it. It is more that they think they have achieved a goal and then the problem is how do you change people who already think they have achieved this? (…) I think that it is a problem with the national level as well. I do not think they have differentiated enough between schools. It is more, here is your offer, it is one offer from the state to meet all the different schools. And I think that is a miscalculation from the policy level. They want to include everybody, but they are all on different levels. (Professor 2).

Another of the Norwegian researchers involved explained how the development of knowledge in assessment should be based on the needs of teachers and school leaders:

> It is positive that we now have more focus upon assessment, but it is more problematic that there are initiatives from DET which are not necessarily based upon what teachers and school leaders find challenging, but instead, it is built upon what are seen as the main challenges by the central government. And when the government and DET decides what the challenges are, it will easily be a summary of what is seen as the main problems in Norway overall, and not adapted to the specific school context. When I work together with schools, I have to start where the school is. I have to know their practice, and know about their experiences so far. (Professor 3).

When discussing the implementation with the Norwegian researchers who had been a part of this programme, another challenge raised was that the Directorate did not seem to acknowledge that there is no real consensus of what AfL is. Researchers themselves do not even seem to agree upon the concept (see for example the criticisms of Taras (2009), Dunn and Malvenon (2009) and Bennett (2011) on the looseness of AfL as a concept). Instead, the policy level had developed a programme that was going to be implemented as the AfL programme. The researchers feared that teachers and school leaders would thus have a superficial understanding of AfL and believe it is all about recipes (echoing the findings of international research, see Chapter 3):

> We had a commission from the Directorate to write a paper. Our main argument is that AfL has no consensus meaning. And it is very interesting if you compare it with classroom management, which is a big movement in Norway at the moment, you can basically say that everybody agrees theoretically, practically, politically how to do it. But you go to AfL and there is actually no agreement. I think part of the problem is that the policy level has not communicated that there is a number of ways to interpret it. It has been interpreted as "this is it, this is how it is interpreted. (Professor 2).

This statement captured one element of the tension found between researchers and participants from the DET. Researchers expressed a preference for focusing upon the complexity of AfL, while the DET had developed a programme that was seen by some teachers as the only way of doing AfL. The four principles of assessment were seen as the most important part of the work, and since it is now in fact a part of the Education Act, teachers have to implement these practices. Another of the Norwegian researchers stated

that there were many good intentions in Norway, but that they believed assessment was left to people without the deepest understanding of what assessment was. The researcher explained further:

> This is why I disagree with them [DET] strongly, that certain specific ways of doing Assessment for Learning can be transferred and prescribed for every teacher and every school and every context. And then they look at the activity, and they don't look at the purpose (…) you have to do this, and you have to do this, and you have to do this… Then there is a problem. [For example] when they talk about Assessment for Learning, then the legal aspects perhaps become more important than the pedagogical aspects. (Professor 4).

The researcher went on to explain how students' rights could result in parents suing teachers if they believed their sons and daughters had not received teaching according to the Education Act. The researchers feared that teachers would work within these constraints and thus limit themselves to doing what they were told. With reference to work in the UK, a researcher explained:

> Mary James said, in one of her papers where she had a big British project on this, that the moment you take what works well in one context and try to translate it into a larger context, like in the national context, it is not possible. Because Assessment for Learning then loses the moments of continuance. Some students need this type of feedback and other students need that type of feedback. One student needs to be followed and one student needs to get grades, and other students need no grades. We have to be very careful and it is not a prescription. It is about developing teachers' competence and awareness. Not the technicians. In development of understanding of what assessment for learning is. And that is what I'm missing. (Professor 3).

Professor 2 explained that there seemed to be a challenging tension between the policy level and researchers in this programme. There were signs that the knowledge of Norwegian researchers was being devalued, and some conflicting signals from the DET to the research community and vice versa:

> And the second challenge is that I do not think that the Ministry, it may sound like I am complaining, but I do not think the Ministry has decided how they would like to use the Higher Education sector. Sometimes they want to use the HE sector sometimes they don´t. And knowing how the inner dynamics of HE function, because they have been so unclear and changing, the HE sector is not really interested. (Professor 2).

On the other hand, some of the civil servants in the DET found it challenging to have to lead the AfL programme and be the change agents in Norway. One of them argued that it should have been steered by the Universities and Colleges but instead they felt they had to lead it from the DET, because they could not find enough expertise among the Universities and Colleges in Norway.

Knowledge production and implementation of AfL in Norway is, in other words, facing the challenge that different stakeholders have different views on what AfL is and how it should be practiced. Teachers are unsure what it is, and need time to develop their practice. The DET has suggested a programme for AfL which some teachers would like to use, but other teachers would like to develop more from their own ideas. Researchers in Norway disagree with some of the interpretations made by the DET on the AfL programme and some researchers also disagree about the implementation strategy used by the DET. These mixed signals are challenging for schools and reveals that there is a tension between a part of the DET and some of the Norwegian researchers. The different perceptions of the programme from the different stakeholders, particularly from some of the researchers in Norway, demonstrate challenges involved when

implementing the AfL programme. Some of these researchers supervised schools in the AfL programme, but did not agree with the implementation strategy developed by the DET, or the interpretation of the AfL programme. This is obviously a dilemma in the knowledge production of AfL and implementation of programmes.

These professors also saw the danger of teachers reducing AfL to just tools and techniques, without understanding the learning processes behind it:

> We see that some municipalities and schools pick and choose from the website what they would like to use, and then they leave the rest. And then the result is very dependent on the schools' capacity for development. (Professor 5).

The AfL programme has to be translated into practice, and one way of doing this is to use the tips from the website. One book[17] on the website has also turned out to be the best-selling book in the field, since it gives suggestions for classroom practice in easy language, with tips for teachers. Some of the researchers find this problematic, since the book neither takes into account the complexity of AfL nor is based upon research like some of the more academic books. Even though the author of the book tried to explain to teachers that this was not the "gold standard" for AfL, it was interpreted as the way of working with AfL in most of the schools included in this report. From the researchers' perspective, the concern was that the implementation of AfL has been superficial with no real understanding. Instead of acknowledging the complexity of the field, the DET offered a teacher friendly programme that could be implemented step-by-step, like a recipe. Even if the members of the DET strongly emphasised that these practices could be developed in several ways, some school leaders and teachers have interpreted the website version of AfL as the "truth" about AfL.

The AfL programme has been built upon the idea that municipalities and counties should pick and choose who they would like to collaborate with in this project. The DET has encouraged co-operation between municipalities, schools and the University and Higher Education sector, but different municipalities and counties have chosen different solutions. Some have also used people outside of the University sector, such as teachers with a specialization in classroom assessment practices and independent consultant companies. It is fair to say that it is a growing market in Norway where different actors offer AfL packages, and from a research point of view, not all of these packages are seen as high quality tools.

Three of the five Norwegian researchers interviewed feared that this belief about the policy programme would lead to further teacher de-professionalisation as well as a lack of trust. The policy programme could be seen as a sign of distrust of the teaching profession, as well as of Teacher Education, since it both suggested which material should be read, and offered examples of how to practice AfL in classrooms. This, according to the researchers, suggested that teachers were not able to develop assessment practices for themselves, and therefore needed recipes to know how to do it.

One of the professors suggested that those at policy level had not realized how much effort is needed to implement AfL. According to him, it was not enough to have big conferences with foreign experts, have teachers listen to lectures and expect them to implement AfL strategies. The professor argued it was necessary to work for years and to have teachers working together in their specific subjects such as history, language and science, to share knowledge and experience. This would be particularly important in secondary and upper secondary schools, where teachers trust subject matter experts more than others.

The interviews with the five professors in Norway showed expert knowledge in the field of assessment. However, as previously described, the choice of the DET to use primarily foreign experts in

the roll-out of the programme has led to some tensions between researchers and the Directorate. In a recent mapping of Norwegian researchers, the Norwegian Knowledge Center found more than 100 researchers in Norwegian universities, colleges and research institutes who had published on the topic of assessment (Lillejord, 2013, personal email). To build capacity and implement AfL in large scale, it would probably be wise to use the capacity these researchers represent and strengthen the dialogue between Norwegian researchers in the field of assessment and the DET. It is evident that there are challenges in how to transform the complex knowledge researchers possess into knowledge that teachers can use in developing their assessment literacy.

The media: Successes and failures of AfL

As part of this case study, searches of online newspapers with the largest readerships in Norway were conducted using the search term *Vurdering for læring*. Searches for articles were also conducted on the websites of Norwegian think tanks[18] and publications specifically targeted at teachers and school leaders. In general, AfL was positively described in the majority of the articles. Less positive coverage tended to be related to conflicts regarding the interpretation and use of the concept. According to the articles analysed, the closer AfL came to summative testing and controlling students and teachers, the more negatively teachers reacted to AfL.

The articles found were divided into three categories: 1) articles explaining the Assessment for Learning programme and theories behind it; 2) articles from schools working with AfL and positive outcomes; 3) articles problematising the programme. The coverage in the two first categories was generally positive. With a few exceptions, these articles did not refer to any achievement standards for the programme. In contrast, the third category included some problematic findings arising from the AfL implementation.

The first category included a number of descriptive articles specifically targeted at teachers. These articles were generally written in a positive manner, explaining why AfL is a good practice. One example was found in *Dagbladet* (23 August 2007*)*, where four researchers wrote about the new AfL programme in Norway following the publication of their book on assessment. The authors specifically stressed the importance of focusing more upon learning, and less upon testing, and claimed that the new assessment programme could be seen as a way of doing this, giving Norway a new start after the introduction of the national tests, which they believed was part of an old-fashioned system.

The second category contained articles on different schools and municipalities working on AfL, with the local press writing about the new initiatives in their schools. In one article in *Fredrikstad Blad* (24 August 2012), the journalists described two schools sending their teachers to Dublin as part of the AfL programme in order to study how Irish students work with AfL. In an article in *Bedre Skole* three schools working on the Better Assessment project were presented and teachers, students and school leaders were interviewed about their experiences. The message was positive overall. A 14-year-old student was quoted as saying: "It is much easier to study for tests when we know what will be on the test and also know the criteria for each grade" (Ruud, 2009, p. 15). The teachers further described the importance of student participation and involving students in self-assessment. They even used cameras and allowed students to film each other as a part of peer-assessment activities. The teachers explained that they had been supervised by researchers from Lillehammer University College, and they explained that even if the new AfL practice was time-consuming, it was important (ibid., p.16).

Some articles described schools that had begun using AfL before the official programme started. In one article, the school leader at Øraker school in Oslo explained the new concept of Assessment for Learning and how the school worked systematically to improve students' learning outcomes through

feedback. The school leader also reflected upon the fact that Norwegian students were known to have good relationships with their teachers, and that they were satisfied with their school, something she believed was noticed abroad from the international test results. She believed this aspect was crucial also for learning (*Aftenposten*, 29 October 2008).

> Another positive outcome of the AfL programme was described in the newspaper *Bergens Tidende* (13 December 2012). The County Governor in Hordaland was interviewed about a decline in complaints from parents about their children's' final grades: 166 complaints in 2012 as opposed to 224 complaints in 2011. The County Governor suggested this could be due to the increased focus upon assessment in Norway and a better assessment practice.

The third category was the negative press. One example of this is from *Sandefjord blad*, where parents in a primary school wrote a letter to the editor and complained about the assessment system that had been introduced in schools. The municipality had interpreted the Assessment for Learning principles to involve assessment of students mid-year, using three categories in relation to expected attainment goals in the different subjects: above average, average or below average. Some of the teachers reacted strongly, and wrote a letter to inform the municipality that they would not participate or use the tools from the municipality to categorise their pupils. When interviewed by the local press, one teacher explained:

> This is an ethical dilemma that puts us teachers in a difficult position. On the one hand, we are supposed to make children believe in themselves and their ability to learn new skills; on the other hand, we are asked to categorise these children. We wish to start a constructive dialogue around what we believe is a way of introducing grades in primary school (…) This is a very serious issue which involves all the children in our city. I have never refused to do what I have been told as a teacher, but this time it is enough. Enough is enough (*Sandefjord Blad* 25 February 2012).

In the Norwegian school system, there is no grading of student work until secondary school level (age 13–16). There is a long tradition in Norway where it has been seen as counter-productive to use grades among primary school children (Lysne, 1984). Grading of young students has not had any support from the majority of the population, despite some attempted initiatives from the Conservative party to introduce grades in primary school. This is one of the reasons why teachers and parents reacted to the labelling of students in categories such as under and above average. It was perceived as too close to grading, something that is not legal in primary schools.

These teachers were not alone. Some parents also reacted strongly to this new system, writing a letter to the editor of *Sandefjord Blad* under the following headline: *What kind of feedback do we want?* :

> We do not need to know where our child is compared to other children. What is important for us is to know how we can help our children to learn more. We would like specific feedback and feed forward comments on how to do this, so we can help our children in the best possible ways. Categorising our children is demotivating, especially for children who are put into the 'below average' category. We do not want to tell our children that they are less good than others, we would like them to learn more (...) Categorising children increases comparability and can lead to more bullying and less enjoyment in schools (Sandefjord Blad 25 February 2012).

The leader of the municipality level was interviewed on the same day in the newspaper, explaining that the feedback offered to the students could actually help them improve, and that informing parents

whether their child was below, at, or above average would in the end help each student to learn. He also pointed to the fact that the DET had instructed all schools to give students mid-year assessments, and the municipality level had therefore decided to use a specific electronic tool for this solution. This case demonstrates that AfL practices have been interpreted quite differently in the various municipalities in Norway, and in some places teachers have reacted strongly to how it was developed and applied. In such circumstances, it is obviously difficult to implement a programme like AfL successfully.

Impact of the AfL programme: Students' learning outcomes

The analysis of policy documents for this case study showed that the central policy level did not have any clear expectations in terms of achievement goals for the participating schools in the programme. For example, the programme was not designed to compare participating municipalities and schools with other schools.

According to the AfL literature, AfL can be expected to increase students' learning outcomes (Black and Wiliam, 1998; Hattie, 2009). In the present study, national test scores from schools participating in the AfL programme in 2010-12 were analysed to see whether there were any changes in student achievement. Scores from Reading, English reading, and Mathematics in the years 2010, 2011 and 2012 are presented. Students in Norway are tested in Grades 5, 8 and 9, so these results include the primary and secondary schools participating in this case study.

Tables 4.3 through 4.8 show the scores from national tests in Reading, English and Mathematics from the year 2010, when the AfL started, to 2012. Approximately 145 primary schools and 80 secondary schools participated in the AfL programme (intervention schools). The test scores from these schools were compared to those from the rest of the schools in Norway, approximately 1 450 primary schools and 820 secondary schools (non-intervention schools). All differences in scores are non-significant.

In Grade 5, students are tested on a proficiency scale from 1 (lowest) to 3 (highest), and in Grades 8 and 9, students are tested on a proficiency scale from 1 (lowest) to 5 (highest). As can be seen from the tables, the mean scores for intervention schools are lower than those for the non-intervention schools in all subjects each year, with one exception. The scores for reading in Grade 9 in 2011 are 3.51 for intervention schools and 3.50 for non-intervention schools (Table 4.4). This is a non-significant difference.

The intervention schools show a decline in Grade 5 Reading scores from 2010 to 2012, from 1.95 to 1.90, and an increase in Mathematics scores from 1.95 to 1.97, but again these are not significant. English reading shows no difference in the period 2010-12. The 2011 English reading scores are missing due to technical problems with the computer-based tests in Norway.

In Grade 9 the same pattern emerges, with slight differences and no real increases in test scores. The mean student score for Reading in intervention schools was 3.44 in both 2010 and 2012, while the non-intervention schools show a decline from 3.49 in 2010 to 3.46 in 2012. It is also worth noting that overall, Grade 9 students from 2011 scored higher than Grade 9 students in 2010 and 2012.

Table 4.3. Results from national tests: Reading Grade 5

	Intervention			Non-Intervention		
Year	Mean	N	Std. Deviation	Mean	N	Std. Deviation
2010	1.95	141	.23	2.00	1427	.25
2011	1.96	143	.23	2.01	1495	.23
2012	1.90	142	.21	1.96	1463	.25

Table 4.4. Results from national tests: Reading Grade 9

	Intervention			Non-Intervention		
Year	Mean	N	Std. Deviation	Mean	N	Std. Deviation
2010	3.44	82	.31	3.49	822	.34
2011	3.51	82	.27	3.50	827	.32
2012	3.44	81	.27	3.46	828	.30

Table 4.5. Results from national tests: Mathematics Grade 5

	Intervention			Non-Intervention		
Year	Mean	N	Std. Deviation	Mean	N	Std. Deviation
2010	1.95	146	.27	1.98	1487	.29
2011	1.93	145	.25	1.95	1495	.28
2012	1.97	143	.24	1.97	1477	.27

Table 4.6. Results from national tests: Mathematics Grade 9

	Intervention			Non-Intervention		
Year	Mean	N	Std. Deviation	Mean	N.	Std. Deviation
2010	3.34	87	.38	3.41	846	.33
2011	3.43	82	.32	3.41	829	.33
2012	3.37	82	.29	3.43	828	.32

Table 4.7. Results from national tests: English Grade 5

	Intervention			Non-Intervention		
Year	Mean	N	Std. Deviation	Mean	N	Std. Deviation
2010	1.94	145	.24	1.99	1487	.26
2011	m	m	m	m	m	m
2012	1.93	141	.24	1.97	1463	.27

Source: Note: 2011 English reading scores are missing due to technical problems with the computer-based tests in Norway

Table 4.8. Results from national tests: English Grade 8

	Intervention			Non-Intervention		
Year	Mean	N	Std. Deviation	Mean	N	Std. Deviation
2010	2.99	59	.24	3.00	598	.31
2011	2.99	80	.28	3.01	824	.35
2012	2.92	81	.29	2.97	837	.33

In general, student achievement in intervention schools is not better compared to their peers in non-intervention schools. Although it might be tempting to conclude that AfL does not have an impact on student achievement, there are a number of reasons why it would be premature to draw this conclusion from these data alone, for the following reasons:

1. The sample of intervention schools has been challenging to define. As previously explained, some schools in Norway believe they are part of the intervention programme and have already integrated the Education Act and AfL principles in their schools even if they are not, in fact, officially part of the intervention programme.

2. As demonstrated in the interview data, some of the participating schools have not included all teachers in their AfL programme since teachers are allowed to volunteer to participate. This means that some students in intervention schools have teachers who still do not use AfL.

3. There is no process data from the AfL programme. Even though teachers reported that they used AfL, there is no classroom observation data to confirm their practice. In other words, there is little knowledge of how the teachers in intervention schools actually practice AfL.

4. Selection criteria for schools to participate in the programme varied between municipalities. Some invited schools with experience of assessment, while other municipalities chose to include schools that did not score as well as other schools on national tests. These municipalities used the AfL programme to include and support schools which had been struggling. This can be one possible explanation for the slightly lower national test results from AfL intervention schools.

- According to PISA data, Norway is one of the countries with the lowest variance in student performance between schools (OECD, 2010).

Overall, it is important to bear in mind that teachers need time to integrate AfL practices in their teaching. It would therefore be of interest to follow these participating schools over the next five years, and analyse national test data trends over this time span, to see whether AfL practices can have a positive impact on students' tests scores when observed over a longer period.

CHAPTER 5.

CONCLUSION AND RECOMMENDATIONS

Expectations of the programme – What next?

As seen through the analysis, the central policy level did not set clear expectations of achievement for the participants in the programme. In addition, the participants were allowed to interpret and outline their own implementation plans for the AfL programme. On the one hand, this has resulted in a rich dataset describing several ways of working with assessment. On the other hand, it makes it very hard to evaluate the effectiveness of the programme and its success. The main message is therefore that a better designed programme, with a baseline and specific targets, would have made it easier to examine its success and impact. This is also a conclusion that was reported in a recent publication synthesising Norwegian research reports (Olsen et al., 2012), and the recent OECD Review of Evaluation and Assessment in Education report on Norway (OECD, 2011). The lack of clear goals and expectations for implementation makes it even more difficult to govern these programmes in Norway.

This report was written at the programme's mid-way point, as it is due to continue until 2014. It is not yet known what will happen in the schools after they have taken part in this programme, but one of the threats to the implementation of AfL is the likelihood of teachers returning to old practices and routines if there is no plan in place for sustaining the changes. The government has decreed by law that all students have the right to know how they will be assessed, what the goal of the learning process is, and what criteria will be used in the assessment process. There are, however, few tools that can be used to monitor whether students are actually receiving what they have been promised. The national Pupil's Survey in Norway is mandatory for all schools to give to their students in the 7th and 10th level[19] and the first year of upper secondary, and surveys students on their participation in assessment. However, there are differences in how the data from these surveys are followed up by municipalities, school leaders and teachers (Rambøll, 2013). It is likely that strong parents will be able to claim these rights for their children, but those from families who are not able to have such a dialogue with the school may lose out. This is a delicate matter, though: if any quality control system is introduced to monitor teachers' assessment practices, it will be interpreted as implying distrust. In other words, governing through laws can easily be interpreted as lack of trust in teachers, which may in turn lead to feelings of being controlled, which could potentially harm their creativity, motivation and self-respect. On the other hand, if the quality of the teaching is not good enough, and parents and students do not feel that teachers are following the Education Act, public trust in schools may decrease. The way forward seems, therefore, to be to continue to support those teachers who seem to have taken up the four principles in the Education Act, and support the good schools and teachers as examples for other schools.

Overall, there is a need for a strategic plan and vision: what will these schools look like in 2014, 2017, 2020? What is the goal and how would one get there? There are different perspectives here that are challenging. Practitioners work in schools for years and years, some even their whole working life. Politicians function in shorter time spans: ministers come and go and are often judged on policies implemented by those before them. These differing outlooks and timelines need to be acknowledged as

one of the challenges in the governance of education. It is reasonable for a minister or politician to want results in their political lifetime. It is equally reasonable for practitioners to need more than 16 months to implement new teaching practices and establish new teacher and student roles as a result of the AfL programme.

Recommendations

Implementation needs clear goals

Implementation in education systems is complex and demands clearly stated goals from the Ministry as well as a clear process for implementation. Clear communication, alignment of the different levels, and agreement on expected short to medium-term outputs must also be ensured in order to optimize the process.

Implementation in co-operation with students

If education programs and reforms are to be successful, they need to focus upon the learner. Student feedback in implementation processes can fuel change and be a valuable reality check on what is really going on in the classroom.

Balancing accountability and trust

The balance between trust and accountability can be seen as a challenge in Norway. In general there is a strong sense of trust in the system, and although Norway has been developing a new quality assessment system (NQAS) in recent years, there are relatively few accountability mechanisms in Norway compared to the US and England. However, when the system relies wholly on trust and thus has few incentives (or sanctions) for the actors, long-term implementation in the face of resistance becomes problematic. School leaders must involve the teachers in the process of developing school cultures based on a real understanding of the intentions and principles of AfL. Yet the majority of the teachers interviewed struggled with interpreting the AfL programme and what would be considered as "correct" practice. There is still a lack of understanding regarding the government's intentions and teachers have not developed a common understanding on how to transform the theory underlying AfL into high quality teaching practices.

In order for the implementation of AfL to continue successfully, it seems crucial to work closely with teachers, balancing the dilemma of trust and accountability. In other words, AfL practice should be expected, but without controlling too much and risking losing teachers' trust and motivation. It would also be useful to show teachers how the different centrally initiated programmes can support each other, instead of competing with each other. It might also be useful to rethink how many programmes it is efficient to have running in Norway at the same time.

Recommendations

Implementation of new practices is complex and requires competence and leadership

The municipalities that succeeded with implementation demonstrated both leadership skills and knowledge of the content of the AfL programme. As such, it is critical that these elements be fostered as the programme progresses.

Implementation through trust

Policy can be seen as persuasion. Implementation of new practices in education is more likely to be sustainable if the leaders are able to keep trust in the system. Clear communication between the different levels and a high degree of trust amongst all stakeholders are thus necessary for successful implementation.

Dealing with "reform fatigue"

Some of the teachers interviewed with more than 25 years of experience, expressed feeling tired of reforms and were concerned that the new programme would be forgotten within a couple of years, just as previous initiatives had been. Teachers also claimed that they felt a sense of loyalty and obligation to do what they had agreed to do, even if sometimes they were not very motivated to put the change into practice. As one teacher said, "We know we do not have a choice, so then there is no need to complain". From the perspective of the teachers, one of the main challenges is that the Directorate of Education and Training implements too many competing programs. Any particular school could be participating in several programs at the same time, and not all schools are able to see how these programs can complement and strengthen each other. As a result, school leaders and teachers, especially those from the smaller municipalities, can be overwhelmed by these competing programmes, initiatives and reforms.

Co-operative networks between those schools and municipalities that are struggling and more outstanding teachers, school leaders and schools, could be better utilised in order to strengthen the smaller municipalities and help them meet the challenges of implementation.

The recent OECD Review of Evaluation and Assessment in Education report on Norway (OECD, 2011) looks at the larger evaluation and assessment framework that AfL is operating in, and provides key recommendations for improving whole-system alignment in this area. This study, while focused on one specific element of that work, underscores the importance of focusing both on the broad picture (i.e. the system) as well as looking deeply into one domain (here, AfL) in order to make strategic decisions that will work to improve the Norwegian education system in the long term. The importance of strategic planning and alignment of goals between governance levels cannot be understated.

Recommendations

Prioritise – What is most important?

For smaller municipalities, there seem to be too many competing programmes and initiatives. There is a need to prioritise these goals and enumerate which are the most important to implement, as well as communicate clearly how those specific elements work together with other initiatives to form a whole.

Strengthen the small municipalities in networks with outstanding schools

Working in networks between schools and sharing knowledge is a proven implementation strategy. Struggling schools and municipalities should continue to work in co-operation with outstanding schools or teachers and school leaders that can support them.

The gap from theory to practice

The implementation of AfL in Norway is facing the challenge that different stakeholders have different views on what AfL is and how it should be practiced. Teachers are unsure what it is, and they

need time to develop their practice. The DET has suggested a programme for AfL which some teachers would like to use, while others would prefer to develop their own ideas more. Researchers in Norway disagree with some of the interpretations made by the DET on the AfL programme and some researchers also disagree about the implementation strategy used by the DET. These mixed signals are challenging for schools.

The implementation of AfL is not only connected to new methods and strategies for teaching, it also has to do with changing professional attitudes towards research and knowledge and adjusting deeply rooted professional knowledge. The programme initiated by the Directorate for Education and Training has, to some extent, succeeded in setting up an infrastructure for knowledge development in the teaching profession. School leaders and teachers who have been involved in the project know some of the theoretical literature and research on assessment, and are able to find new literature and build on theories that can give new insights in their daily work with students. However, it seems that most teachers pursue their work with a singular view of what is considered 'right' in terms of assessment and often do not know about, or know how to handle, differing approaches.

One important element, which has not been sufficiently addressed, is the lack of research consensus of what AfL is. In order to smooth implementation, the concept has been simplified and the website tools streamlined. However, previous international research warns against an overly superficial understanding of AfL on the part of teachers and school leaders, as a lack of deeper understanding not only hinders the sustainability of the practice in the long-term, but also its effectiveness and impact.

Some schools are able to connect the new practice to existing knowledge by considering AfL as a different paradigm for learning and teaching, but there is still a lot more work to be done. It usually takes several years for a whole school to be able to really change perspective and establish new practices. In some subjects, for example Norwegian, teachers have already worked with process-oriented writing, which is based on learning theories similar to AfL. Making this connection allows them to see how to work with AfL in the classroom and how to experiment with new methods and strategies. In other subjects, this can be more difficult. In general, it is not possible to say that AfL has been integrated into teachers' knowledge base or accepted as a core practice in the profession at present.

As mentioned in Chapter 3, Norway has nine national centres[20] that successfully support teachers and schools in reading, writing, science, language, mathematics and the learning environment, yet there is at present no assessment centre in Norway. A national centre of expertise in assessment could work to build capacity on a large scale. It could also support cooperation between researchers and facilitate the development and implementation of assessment models.

Recommendations

Continue to work to clarify the concepts and understanding of what AfL is

In order to reap the full benefits of AfL, is important to continue to develop a deep understanding of what AfL is and how that translates to practice in a multiplicity of ways. As knowledge and capacity grows, new high-quality tools and networks can be developed and promoted by and for teachers and school leaders.

Establish a national assessment centre to build capacity on large scale

As previously recommended by the OECD (OECD, 2011), establishing Centres with expertise on assessment evaluation would allow for capacity building on the system level. Such a centre would need to

work closely with practitioners in order to identify needs for capacity building and how best to encourage teachers and school leaders to become involved. It would also need to connect closely to relevant research and institutes.

Change takes time

Even in the schools that have worked with AfL for 6 years, teachers report that school leaders need to make a real effort to ensure that it becomes a core practice. This is not surprising: changes in practice take time and will not succeed unless teachers are supported and trained in both the theoretical and practical aspects of their work over the course of their career. Yet sustainable development is seldom considered in AfL implementation (Gardner et al., 2011). Data from this case study indicate that half of the municipalities were struggling with implementation of AfL, despite all the work and effort to ensure success.

Recommendations

Implementation needs sustainable development over time

School leaders need support to be able to keep focus over time. Governance actors and school leaders require training and support in the acquisition of research knowledge to develop greater competency in this area and enable them to facilitate change. They need knowledge about the content of the change process, about what works and what are the theoretical assumptions underlying the new teaching paradigm. These processes develop over years, and implementation therefore needs to be planned strategically, with a particular focus on what to do *after* the official programme is over.

> *(...) you actually have to engage people. They have to come with you (...). People see research, policy and practice as a hierarchy. However, you can turn that model on its side and say: collectively what is it that we are all trying to achieve and what are the different roles we play if we are to achieve our aspiration? You can't do that without dialogue. There has to be space both for dialogue and for the networks necessary to support change across communities.*
>
> Prof. Louise Hayward, September 2012n

NOTES

1. www.oecd.org/newsroom/g20labourministersmustfocusonyoungjobseekers.htm (accessed 5 December 2012).

2. www.ssb.no/en/befolkning/statistikker/folkendrhist (accessed 19 March 2013, Statistics Norway).

3. www.aftenposten.no/meninger/kommentatorer/madsen/Den-nye-larerrollen6609219.html#.UUcaQ2iTeS0.

4. In the Swedish calendar, there is a name for each day. The EMIL project started on the 14th of November, which is the day of Emil (Lundgren 2003).

5. The Storting is the Norwegian Parliament, and these reports are White Papers for the Storting.

6. www.udir.no/Upload/Rapporter/2009/5/Utd_speilet_09_u_blanksider_nett.pdf?epslanguage=no (accessed 10 December 2012).

7. www.udir.no/Upload/Statistikk/Karakterer/2012/Karakterstatistiskk_2011_2012_analyse.pdf?epslanguage=no.

8. Opplæringslova (Norwegian for Education Act) short title for the Act relating to primary and secondary education.

9. Dated 25 August 2010, ref 2010/1396.

10. Each school owner chooses one or several resource persons who will have 20 % of a full-time equivalent position throughout the whole participation period. This person runs the school development process locally, both when it comes to content and organisation. Resource persons can be administrators, school leaders, teachers, university employees or from other agencies. The school owner can hire the resource person at a higher percentage of a full-time equivalent, and for municipalities the resource person(s) can also serve more than one municipality.

11. www.udir.no/PageFiles/35141/UDIR_VFL_Sluttrapport%20pulje%201.pdf: Nasjonal satsing på vurdering for læring (2010-14). Mer systematisk arbeid med vurdering for læring på alle nivåer. Fra forskning og forskrift til felles forståelse, forankring og faglig fokus. Oppsummering av sluttrapporter fra pulje 1 i satstingen Vurdering for læring.

12. Ofsted is the Office for Standards in Education Children's Services and Skills in England. Ofsted report directly to Parliament and are independent and impartial. Ofsted inspect and regulate services which care for children and young people, and those providing education and skills for learners of all ages. Ofsted regularly publish reports on schools in England(see http://www.ofsted.gov.uk/about-us).

13. DfES is the Deparment of Education in England, and responsible for education and children's services.

14. H = Conservative Party, KrF = Christian Democrats, V +Social Liberal Party, AP= Labour Party, SV=Socialist Left Party, SP=Centre Party.

15. The Knowledge Promotion reform in 2006 introduced a new outcomes-based curriculum with competence goals for key stages of education.

16. The Nordic Institute for Studies in Innovation (NIFU) conducted a survey on behalf of the Directorate asking school leaders from 667 schools whether they were participating in the AfL programme; 283 responded in the affirmative. When investigated further, it was found that 82 of these 283 schools were not actually participants in the programme. The researchers therefore suggested that some schools have started their own school development programme in AfL, despite not being formally involved in the initial Directorate project. These schools thus believed they were part of the official DET programme (Vibe, 2012). It is important to clarify that it is the municipalities with their project leaders who participate in the official programme, but they also invite and choose schools that are then offered places at the conferences and subsequently implement the AfL programme. In some of the municipalities participating in this study, it seems as if some of the schools who are not officially members of the programme believe that they are. One possible explanation for this is that some of the municipalities use strategies for sharing knowledge through conferences for all schools in their municipality. Nevertheless, it is challenging to evaluate the impact of the AfL programme on participating schools through test scores when so many schools were confused about their status/degree of involvement in the project. It must be noted, however, that this situation commonly occurs in AfL implementation initiatives in other contexts, with planning for the impact of the policy programme only being seen as relevant and addressed, if at all, at the end of the whole process (see Gardner et al., 2011).

17. Slemmen, T (2009), *Vurdering for læring i klasserommet* [Assessment for Learning in the Classroom], Gyldendal Akademisk, Oslo, www.udir.no/Vurdering-for-laring/Litteraturtips/Norsk-litteratur/.

18. Civita and Manifest Analyse, representing two of the most active think tanks in Norway with different political and philosophical perspectives.

19. When the students are 12-13 years-old, 15-16 years-old, and the first year of upper secondary.

20.**Error! Hyperlink reference not valid.**.

REFERENCES

Aasen, P. et al. (2012), Kunnskapsløftet som styringsreform – et løft eller et løfte? Forvaltningsnivåenes og institusjonenes rolle i implementeringen av reformen [The Implemention of the Curriculum Reform: Knowledge Promotion and the Roles of Different Stakeholders], Report 2012/20, NIFU/Universitetet i Oslo, Oslo.

Alexiadou, N. (2010), "Education policy convergence through the open method of coordination: Theoretical reflections and implementation in 'old' and 'new' national contexts", *European Educational Research Journal,* Vol. 9, No. 3, p. 345-358.

Al-Samarrai, S. (2009), "The impact of governance on education inequality: Evidence from Bangladesh", *Public Administration and Development*, Vol. 29, No. 3, pp. 167-179.

Allerup, A. P. et al.(2009), *Evaluering av det Nasjonale kvalitets- vurderingssystemet for grunnopplæringen* [Evaluation of the Norwegian System for Quality Assessment], Agderforskning, Report FoU-rapport.

Apple, M. W. (2005), "Are Markets in Education Democratic? Neoliberal Globalism, Vouchers and Politics of Choice", in M. W. Apple, J. Kenway and M. Singh (eds.), *Globalizing Education Policies, Pedagogies, and Politics*, Oxford, Peter Lang, New York/Washington/ D.C ./Baltimore/Bern /Frankfurt am Main/ Brussels/Vienna.

ARG (2006), *The Role of Teachers in the Assessment of Learning*, Assessment Reform Group, http://assessmentreformgroup.files.wordpress.com/2012/01/asf_english.pdf.

Azúa, X. and M. Bick (2009), 'Formación de profesores en evaluación para el aprendizaje' [Teacher Education on Assessment for Learning], in Sotomayor, Carmen and Walker, Horacio (eds.) *Formación Continua de Profesores*, Santiago de Chile Editorial Universitaria, pp. 263-283.

Bakken, A. and J.I. Elstad (2012), For store forventninger? Kunnskapsløftet og ulikehter i grunnskolekarakterer [Too High Expectations? The Curriculum Reform, Knowledge Promotion and Differences in School Grades] ,Norsk institutt for forsknig om oppvekst, velferd og aldring, NOVA Report.

Ball, S. J. (1990), Politics and Policy Making in Education. Explorations in Policy Sociology, Routledge, London and New York.

Ball, S.J.and C. Junemann (2011), "Education policy and philanthropy-the changing landscape of English educational governance", *International Journal of Public Administration*, Vol. 34, No. 10, pp. 646-661.

Ball, S., M. Maguire and A. Braun (2012), *How Schools Do Policy: Policy Enactments in Secondary Schools*, Routledge, London.

Baird, J. and L. Lee-Kelley (2009),"The dearth of managerialism in the implementation of national assessment policy", *Journal of Education Policy*, Vol. 24, No.1, pp. 55-81.

Barrett, A. M. (2011), "An education millennium development goal for quality: Complexity and democracy", *Compare*, Vol. 41, No. 1, pp. 145-148.

Bennett, R. E. (2011), "Formative assessment: a critical review", *Assessment in Education: Principles, Policy and Practice*, Vol.18, No.1, pp. 5-25.

Bergesen, H. O. (2005), *Kampen om Kunnskapsskolen* [The Battle for the Knowledge School], Universitetsforlaget, Oslo.

Berryhill, J, et al. (2009), "The effects of education accountability on teachers: Are policies too stress - provoking for their own good?", *International Journal of Education Policy and Leadership*, Vol. 4, No. 5, pp. 1-14.

Biesta, G. J. (2004), "Education, accountability, and the ethical demand: Can the democratic potential of accountability be regained?", *Educational Theory,* Vol. 54, No. 3, pp. 233-250.

Black, P.et al. (2003), *Assessment for Learning, Putting It Into Practice*, Open University Press, London.

Black, P.et al. (2004), The Nature and Value of Formative Assessment for Learning (Draft), King's College, London.

Black, P. and D. Wiliam, (1998), "Assessment and classroom learning", *Assessment in Education: Principles, Policy and Practice*, Vol. 5, No. 1, pp. 7 -74.

Blackmore, J. (2011), "Bureaucratic, corporate/market and network governance: Shifting spaces for gender equity in education", *Gender, Work and Organization*, Vol. 18, No. 5, pp. 443 - 466.

Brookhart, S. M. (2001), "Successful students' formative and summative uses of assessment information", *Assessment in Education*, *Principles, Policy and Practice*, Vol. 8, No. 2, pp. 153-169.

Brown, G. T. L. et al. (2009), "Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment", *Assessment in Education: Principles, Policy and Practice*, Vol. 16, No. 3, pp. 347-363.

Bunar, N. (2011), "Urban development, governance and education: The implementation of an area-based development initiative in Sweden", *Urban Studies*, Vol. 48, No.13, pp. 2849-2864.

Campbell, C. and B. Levin (2009), "Using data to support educational improvement", *Educational Assessment, Evaluation and Accountability*, Vol. 21, No. 1, pp. 47–65.

Carless, D. (2005), "Prospects for the implementation of assessment for learning", *Assessment in Education: Principles Policy and Practice*, Vol.12, No.1, pp. 39-54.

Condie, R. et al. (2005), *Evaluation of the Assessment is for Learning Programme: Final report*, Quality in Education Centre, University of Strathclyde, Glasgow.

Dale, R. (ed.) (1985), Education, Training and Employment. Towards a New Vocationalism?, Pergament Press, Oxford.

The Directorate for Education and Training (2012), The Education Mirror. Analysis of Primary and Secondary Education and Training in Norway.

Dysthe, O. (2008), ”Klasseromsvurdering og læring”, [Classroom Assessment and Learning], *Bedre skole*, Vol. 3, pp. 16-23.

DfES (Department for Education) (2007), *Assessment for Learning: Eight schools Project Report*, DfES Publications, London.

Dobbins, M. and K. Martens (2012), “Towards an education approach a la finlandaise? French education policy after PISA”, *Journal of Education Policy,* Vol. 27, No. 1, pp. 23-43.

Dokka, H. J. (1975), ”Enhetsskoleproblemet i den norske skoles historie” [The Problems with the Unified School in the Norwegian School History], in Harbo et al. (eds.) ”*Skolen og samfunnet. En pedagogisk essaysamling* ”[School and Society. An Essay Collection.], Fabritius Forlag, Oslo.

Dori, Y. J. (2003), “From nationwide standardized testing to school-based alternative embedded assessment in Israel: Students’ performance in the Matriculation 2000 project”, *Journal of Research in Science Teaching*, Vol. 40, No.1, pp. 34-52.

Dunn, K. and S. Mulvenon (2009), “A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education”*, Practical Assessment, Research and Evaluation*, Vol. 14, No. 7, pp.1-11.

Ecclestone, K and R. Daugherty (2006), “Constructing assessment for learning in the UK policy environment”, in Gardner, J. (ed.) *Assessment and Learning*, SAGE, London, 149-167.

Eklöf, H., et al. (2012), “Hva vet vi om elevers testmotivasjon? Erfaringer fra internasjonale og nasjonale undersøkelser i Norge og Sverige*”* [What do we know about students’ test motivation? Experiences of international and national tests in Norway and Sweden]. Chapter 6, in T.N. Hopfenbeck, M. Kjærnsli and R.V. Olsen (eds.) “Kvalitet i norsk skole. Internasjonale og nasjonale undersøkelser av læringsutbytte og undervisning*”* [Quality in the Norwegian School. International and National Tests of Learning Outcomes and Teaching], Universitetsforlaget, Oslo.

Elstad, E., Nortvedt and Turmo. (2009), “The Norwegian assessment system: An accountability perspective”, *Cadmo*, Vol. 17, No.1, 89-103.

Ertl, H. (2006), “Educational standards and the changing discourse on education: the reception and consequences of the PISA study in Germany”, *Oxford Review of Education*, Vol. 32, No. 5, pp. 619-634.

Fazekas, M. and Burns, T. (2012), “Exploring the complex interaction between governance and knowledge in education”, *OECD Education Working Papers*, No. 67, OECD Publishing. doi.org/10.1787/5k9flcx2l340-en

Gardner, J. (2011), *Assessment and Learning*, Second Edition, Sage.

Gardner, J. et al. (2011), "Engaging and empowering teachers in innovative assessment practice", in Berry, R. and B. Adamson (eds), *Assessment Reform in Education: Policy and Practice*, Springer, London, pp. 105-119.

Gioka, O. (2006), "Assessment for learning in physics investigations: Assessment criteria, questions and feedback in marking", *Physics Education*, Vol. 41, No. 4, pp. 342-346.

Gipps, C. et al. (2005), "From TA to Assessment for Learning: The impact of assessment policy on teachers' assessment practice", paper presented at the British Educational Research Association Annual Conference, University of Glamorgan, 14−17 September 2005.

Grek, S. (2010), "International organisations and the shared construction of policy 'problems': Problematisation and change in education governance in Europe", *European Educational Research Journal*, Vol. 9, No. 3, pp. 396-406.

James, M. et al. (2007), Improving Learning How to Learn: Classrooms, Schools and Networks, Routledge, London.

Harlen, W. (2007), "Criteria for evaluating systems for assessment", *Studies in Educational Evaluation,* Vol. 33, No.1, pp. 15-28.

Hattie, J. (2009), Visible Learning: A Synthesis of Over 800 Meta Analyses Relating to Achievement, Routledge, Oxon.

Hayward, L. and E. Spencer (2010), "The complexities of change: Formative assessment in Scotland", *Curriculum Journal*, Vol. 21, No.2, pp.161-177.

Haug, P. (2004), "Sentrale resultat frå evalueringa av Reform 97" [Key Results From the Evaluation of Reform 97], *Norsk Pedagogisk Tidsskrift*, Vol. 88, No.4, pp. 248-263.

Haug, P. (2003), "The research council of Norway Evaluating Reform 97", in P. Haug and T. A. Schwandt (eds.), *Evaluating Educational Reforms Scandinavian Perspectives*, Information Age Publishing, Connecticut.

Hernes, G. and K. Knudsen (1976), "Utdanning og ulikhet" [Education and Inequality], NOU 1976/46, Universitetsforlaget, Oslo.

Hertzberg, F. (2003), "Arbeid med muntlige ferdigheter" [Work on Oral Skills], in Klette, K. (ed.), *Klasserommets praksisformer etter Reform 97* [Classroom Practices After Reform 97], Unipub, Oslo, pp.137-171.

Hodgson, J. Rønning, W., and P. Tomlinson(2012), "Sammenhengen Mellom Undervisning og Læring, En studie av læreres praksis og deres tenkning under Kunnskapsløftet" [On the Relationship Between Teaching and Learning, A Study of Teachers' Practice and Their Thinking Under the Curriculum Reform Knowledge Promotion], Nordlands Forsking.

Hooge, E., T. Burns and H. Wilkoszewski (2012), "Looking beyond the numbers: Stakeholders and multiple school accountability", *OECD Education Working Papers*, No. 85, OECD Publishing.

Hopfenbeck, T.N., et al. (2012), "Assessment with distinctly defined criteria: A research study of a National project", *Policy Futures in Education*, Vol. 10, No. 4, 421-433.

Hopfenbeck, T.N. (2011), "Fostering self-regulated learners in a community of quality assessment practices", *CADMO*, Vol. 1, pp. 7-21.

Hyslop-Margison, E. J., and A. M. Sears (2006), Neo-Liberalism, Globalization and Human Capital Learning. Reclaiming Education for Democratic Citizenship, Springer, Netherlands.

Jensen, K. (2008), "Rapport til Forskningsrådets KUL programme 2008. Profesjonslæring i endring" (PROLEARN) [Report to the Norwegian Research Council on the programme KUL in 2008], University of Oslo.

Jensen, K. (2012), "The desire to learn: An analysis of knowledge-seeking practices among professionals", in Jensen, Lahn and Nerland (eds.) *Professional Learning in the Knowledge Society*, Sense Publishers, Rotterdam/Boston/Taipei.

Kellaghan, T. (2004), *Assessing Student Learning in Africa,* World Bank, Washington, D.C.

Kellard, K. et al. (2008), Evaluation of the Developing Thinking and Assessment for Learning Development programme, Welsh Assembly Government.

Kirton, A. et al. (2007), "Revolution, evolution or a Trojan horse? Piloting assessment for learning in some Scottish primary schools", *British Educational Research Journal*, Vol. 33, No. 4, pp. 605-627.

Klette, K. and Carlsen (2012), "Knowledge in teacher learning: New professional challenges", in Jensen, Lahn og Nerland (eds.), *Professional Learning in the Knowledge Society,* Sense Publishers, Rotterdam/ Boston/Taipei.

Klette, K. (2003), "Lærernes klasseromsarbeid; Interaksjons- og arbeidsformer i norske klasserom etter Reform 97" [Teachers in the Classroom; Interaction and Working in Norwegian Classrooms After Reform 97], in K. Klette (ed.) "*Klasserommets praksisformer etter reform 97"*, [Classroom Practices After Reform 97] *,* Unipub, Oslo , pp. 39-76.

Klenowski, V. (2009), "Assessment for learning revisited: an Asia−Pacific perspective", *Assessment in Education*, Vol. 16, No 3, pp. 263−268.

Kogan, M, Lundgren and O'Donoghue (1990), Part One. The Examiner's Report. Review of National Policies for Education, Norway, OECD Publishing, Paris.

Kogan, M. et al. (1989), "*Del 2 Ekspertvurdering fra OECD. OECD-vurdering av norsk utdanningspolitikk.* Norsk rapport til OECD Part 2" [Expert Evaluation from OECD of Politics for Education in Norway], Kirke-og undervisningsdepartementet: Kultur-og vitenskapsdepartementet, Aschehoug, Oslo.

Lawn, M. and J. Ozga (2009), The Sleep of Reason Breeds Monsters: Data and Education Governance in England, Centre for Educational Sociology, Edinburgh.

Levin, B. (2011), "Mobilising research knowledge in education", *London Review of Education*, Vol. 9, No. 1, pp. 15.

Lillejord, S. (1997), "Målstyring, skoleutvikling og skoleledelse" [Steering, School Development and School Leadership], in O. L. Fuglestad and S. Lillejord (eds.), "Pedagogisk ledelse*"*[Education Leadership], Fagbokforlaget, Bergen.

Lillejord (2003), "Ledelse i en lærende skole*"* [Leadership in a Learning School]*,* Universitetsforlaget.

Lillejord, S. and T.N. Hopfenbeck (2013), "Vurdering og læring i skolen" [Assessment and Learning in School]*,* in Lillejord et al., "Livet i skolen" [Life in School], 2, Grunnbok i pedagogikk og elevkunnskap. Lærerprofesjonalitet, pp. 231-259.

Lie, S. et al. (2004), "Nasjonale prøver på prøve. Rapport fra en utvalgsundersøkelse for å analysere og vurdere kvaliteten på oppgaver og resultater til nasjonale prøver våren 2004" [Testing the national tests. Report of a Survey to Analyze and Evaluate the Quality of National Tests in Spring 2004], University of Oslo/Namsmatsstofnun, Educational Testing Institute, Oslo/Rekykjavik.

Lie et al. (2005), "Nasjonale prøver på ny prøve*"* [Re-testing the National Tests], University of Oslo, Oslo.

Lund, T. (ed.) (2000), "Nye perspektiver på elevvurdering. Erfaringer fra forsøk med nye vurderingsformer i ungdomsskolen" [New Perspectives on Student Assessment. Lessons learned from Experiments with New Forms in Secondary Schools], En rapport fra Ungdomstrinnsprosjektet i Troms, Statens Utdanningskontor i Troms.

Lundgren, U. P. (2003), "The political governing (governance) of education and evaluation", in P. Haug and T. A. Schwandt, (eds.), *Evaluating Educational Reforms: Scandinavian Perspectives,* Information Age, Greenwich, CT., pp. 99-110.

Lysne, A. (1984), "Grading of student's attainment: Purposes and functions", *Scandinavian Journal of Educational Research*, Vol. 28, No.3, pp.149-165.

Marshall, B. and M. J. Drummond (2006), "How teachers engage with Assessment for Learning: Lessons from the classroom", *Research Papers in Education,* Vol. 21, No. 2, pp. 133-149.

MacPhail, A. and J. Halbert (2010), "We had to do intelligent thinking during recent PE: students' and teachers' experiences of assessment for learning in post-primary physical education", *Assessment in Education Principles Policy and Practice*, Vol. 17, No. 1, pp. 23-39.

The Ministry of Education and Research (1988), "Med viten og vilje "[With Knowledge and Intent], Avgitt til Kultur- og vitenskapsdepartementet 9, September 1988, Oslo.

The Ministry of Education and Research (1991), Proposition to the Storting No. 37, (1990 – 1991) *Concerning Certain Aspects of Upper Secondary Education,* Oslo, Norway.

The Ministry of Education and Research (1996), Proposition to the Storting No. 47, (1995-1996) *Concerning Student Assessment, School self-evaluation and the National Quality Assessment System*, Oslo, Norway.

The Ministry of Education and Research (1999), Proposition to the Storting No. 28 (1998-1999) Towards Richer Goals: Concerning the Comprehensive Schools, Equal Opportunity In Educational and a National Strategy for Assessment and Quality Development in Compulsory and Upper Secondary Education, Oslo, Norway.

The Ministry of Education and Research (2002), *First Class from First Grade*, Official Norwegian Report, 2002:10, Oslo, Norway.

The Ministry of Education and Research (2003), *In the first row. Enhanced Quality in Education for all*, Vol. 16, Oslo, Norway.

Mikkelsen, R. et al. (2011), "Morgendagens samfunnsborgere. Norske ungdomsskoleelevers prestasjoner i den internasjonale demokratiundersokelsen ICCS*"* [International Civic and Citizenship Education Study 2009]*,* Institutt for lærerutdanning og skoleforskning, University of Oslo.

Monsen, L. (2003), "Challenges and dilemmas for the evaluation of Reform 97", in Haug, P. and T. Schwandt, (eds.) *Evaluating Educational Reforms: Scandinavian Perspectives*, Information Age Publishing.

Olsen, R., et al. (2012), Elevenes læringssituasjon etter innføringen av en ny reform Synteserapport fra evalueringen av Kunnskapsløftet [Pupils' Learning and the Introduction of a New Reform. A Synthesis Report], Institutt for Lærerutdanning og Skoleforskning, Universitetet i Oslo.

OECD (2012), *OECD Economic Surveys: Norway 2012*, OECD Publishing, Paris, http://dx.doi.org/10.1787/eco_survey-nor-2012-en.

OECD (2011), Reviews of National Policies for Education: Improving Lower Secondary Schools in Norway 2011, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264114579-en.

OECD (2010), PISA 2009 Results: What Makes a School Successful, *Resources, Policies and Practices (Volume IV)*, PISA, OECD Publishing http://dx.doi.org/10.1787/9789264091559-en.

OFSTED (2008), Assessment for Learning: The Impact of National Strategy Support, Office for Standards in Education, London.

Olsen, R. et al. (2012), "Elevenes læringssituasjon etter innføringen av ny reform. Synteserapport fra evualeringen av Kunnskapsløftet" [The Pupils' Learning Situation After the Introduction of New Reform. Report from evualeringen of Knowledge Promotion], Acta Didactica, January 2012.

Ozga, J. (2009), "Governing education through data in England: from regulation to self-evaluation", *Journal of Education Policy*, Vol. 24, No 2, pp. 149-162.

Peterson, J., (2003), *Policy Networks,* Institute for Advanced Studies, Vienna.

Phillips, D. (2004), "Toward a theory of policy attraction in education", in Steiner-Khamsi, Gita (ed.), *The Global Politics of Educational Borrowing and Lending*, Teachers College, Columbia University, New York /London, pp. 54-67.

Rambøll (2013), "Evaluering av årlige tilstandsrapporter" [Evaluation of Annual Status Reports. Subreport 2. Report to Directorate for Education and Training], Delrapport 2, Rapport til Utdanningsdirektoratet.

Resh, N. and A. Benavot (2009), "Educational governance, school autonomy, and curriculum implementation: diversity and uniformity in knowledge offerings to Israeli pupils", *Journal of Curriculum Studies*, Vol. 41, No. 1, pp. 67.

Rizvi, F. and B. Lingard, (2009), *Globalizing Education Policy*, Routledge, London.

Ruud, M (2009), "Mer konkret vurdering " [More Specific Assessment], *Bedre Skole*, pp. 14-17.

Sandberg N. and P. Aasen (2008), "Det nasjonale styringsnivået. Intensjoner, forventninger og vurderinger, Delrapport 1 Evaluering av Kunnskapsløftet" [The National Level of Government Intentions, Expectations and Assessments. Subreport 1 Evaluation of the Curriculum Reform Knowledge Promotion], NIFU Step, UiO, Oslo.

Sandvik, L. V. et al. (2012), "Vurdering i skolen. Intensjoner og forståelser" [Assessment in School. Intentions and Perceptions] Trondheim, "Delrapport 1 fra prosjektet Forskning på individuell vurdering i skolen (FIVIS)" [Subreport 1 of the Project Research on Individual Assessment in Schools].

Shepard, L. (1992), *Will National Tests Improve Student Learning?,* CSE Technical Report 342, CRESST, University of Colorado, Boulder.

Sivesind, K. (2012), "Kunnskapsløftet: Implementering av nye lærplaner i reformen. Synteserapport fra evalueringen av Kunnskapsløftet" [The Curriculum Reform Knowledge Promotion and Implementation: Synthesis Report of the Evaluation of Knowledge Promotion], Acta Didactica Oslo, Institutt for lærerutdanning og Skoleforskning, Universitetet i Oslo.

Skarheim, P. (2010), "Assessment policies and research in Norway", Keynote by Petter Skarheim, Norwegian Directorate for Education and Training, Association for Educational Assessment in Europe, 11th Annual Conference, Oslo, Norway.

Skaalvik, E. and S. Skaalvik (2007), "Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout", *Journal of Educational Psychology,* Vol. 99, No. 3, pp. 611-625.

Skedsmo, G. (2011), "Formulation and realisation of evaluation policy: Inconsistencies and problematic issues", *Journal of Educational Assessment, Evaluation and Accountability*, Vol. 23, No. 1, pp. 5-20.

Sookrajh, R. and J. Joshua (2009), "Language matters in rural schools in South Africa: Are educators making the implementation of the Language in Education Policy", *Language Learning Journal*, Vol. 37, No.3, pp. 323-338.

Stiggins, R. and J. Arter (2002), *Assessment for Learning: International Perspectives. The Proceedings of an International Conference,* (Chester, England, United Kingdom, September 18, 2001). Corporate source: Assessment Training Inst., Inc., Portland.

Stobart, G. (2008), *Testing Times: The Uses and Abuses of Assessment,* Routledge Taylor and Francis group, Oxon.

Tapan, S. (2001), A Study of Present Learning Assessment Practices to Ensure Quality of Primary Education. Basic Education Studies, 2000-2001, UNESCO Office, Dhaka.

Taras, Maddalena (2009), "Summative assessment: The missing link for formative assessment", *Journal of Further and Higher Education*, Vol. 33, No.1, pp. 57-69.

Telhaug, A. O. (2007), Kunnskapsløftet i et utdanningshistorisk perspektiv [Knowledge in an educational historical perspective] in H. Hølleland (ed.) På vei mot Kunnskapsløftet, Begrunnelser, løsninger og utfordringer, Cappelen Akademisk Forlag: Oslo, pp. 47-65.

Telhaug, A. O. et al. (2006), "The Nordic model in education: Education as part of the political system in the last 50 years", *Scandinavian Journal of Educational Research*, Vol. 50, No. 3, pp. 245-283.

Telhaug, A. O. (1997), "Utdanningsreformene Oversikt og Analyse" [Education Reforms: Overview and Analysis], Didakta Norsk Forlag, Oslo.

Telhaug, A. O. (1990), "Den nye utdanningspolitiske retorikken: bilder av internasjonal skoleutvikling" [The New Education Policy Rhetoric: Pictures of International Educational Development], Universitetsforlaget, Oslo.

Theisens, H. (2012), "Effective governance from the centre", paper presented at the First Thematic GCES Conference, 21-22 November 2012, The Hague, OECD/CERI.

Thompson, M. and D. Wiliam (2008), "Tight but Loose: A Conceptual Framework for Scaling Up School Reforms", in Wylie, C. (ed.) *Tight but Loose: Scaling Up Teacher Professional Development in Diverse Contexts*, Educational Testing Service (ETS), Princeton, New Jersey.

Throndsen, I. et al. (2009), "Bedre vurdering for læring. Rapport fra. Evaluering av modeller for kjennetegn på måloppnåelse i fag" [Better Assessment of Learning. Report from "Evaluation of Models for Achievement Criteria in Subjects"], Universitetet i Oslo.

Tierney, R. and J. Charland (2007), "Stocks and prospects: Research on formative assessment in secondary classrooms", online submission, paper presented at the Annual Meeting of the American Educational Research Association, Chicago, 11 April.

Titeca, K. and T. de Herdt (2011), "Real governance beyond the 'failed state'. Negotiating education in the Democratic Republic of the Congo", *African Affairs*, Vol. 110, No. 439, pp. 213-231.

Tolo, A. (2011), "Hvordan blir lærerkompetanse konstruert? En kvalitativ studie av PPU-studenters kunnskapsutvikling" [How is Teacher Competence Developed? A Qualitative Study of Teacher Students' Knowledge], Phd. avhandling/monografi, Det psykologiske fakultet, Universitetet i Bergen.

Torrance, H. and J. Pryor (2001), "Developing formative assessment in the classroom: Using action research to explore and modify theory", *British Educational Research Journal*, Vol. 27, No. 5, pp. 615-631.

Trippestad, T. A. (1999), "Som man reder, ligger man", in Meyer S. and T. Sirnes (ed.) "Normalitet og identitetsmakt i Norge. Makt og demokratiutredningen 1998-2003*"* [Power and Democracy. A Report], AdNotam Gyldendal, Oslo.

Vibe, N. (2012), "Spørsmål til Skole Norge høsten 2011: Resultater og analyser av Utdanningsdirektoratets spørreundersøkelse blant skoler og skoleeiere" [Questions for School Norway fall 2011: Results and Analysis of the Directorate's Survey of Schools and School Owners.], (Report5/2012) NIFU, Oslo.

Webb, M. and J. Jones (2009), "Exploring tensions in developing assessment for learning", *Assessment in Education*, *Policy and Practice,* Vol. 16, No. 2, pp. 165-184.

Weeden, P. and Winter. (1999), Learners' Expectations of Assessment for Learning Nationally: Report for the Qualifications and Curriculum Authority, Graduate School of Education, University of Bristol, Bristol.

Welle-Strand A. and A. Tjeldvold (2002), "The Norwegian unified school – a paradise lost?", *Journal of Education Policy*, Vol. 17, No. 6, pp. 673-686.

Willis, J. (2008), "Assessment for learning: a socio-cultural approach", in AARE 2008 International Education Research Conference, Brisbane: Papers collection (Conference of the Australian Association for Research in Education, 30 November-4 December 2008), compiled by P. L. Jeffrey, Australian Association for Research in Education, Melbourne.

Xaba, M. I. (2011), "The possible cause of school governance challenges in South Africa", *South African Journal of Education*, Vol. 31, No. 2, pp. 201-211.

ANNEX A.

METHODOLOGICAL NOTES

Researchers' relationships with participants can influence both the interview situation and the analysis. Hopfenbeck has known several of the stakeholders from previous work in Norway, in different roles during the last 10 years, both on the municipal level, as a former teacher and as a school researcher at the University of Oslo. On the one hand, previous knowledge from the field can be seen as a positive from a research perspective, and serve to enhance the interview situation. On the other hand, the researchers needed to keep a professional distance to avoid becoming too close to the interview participants. The role of the second author was therefore partly to maintain the critical focus of the project and to avoid bias in interpretation of the data.

In some of the interviews, participants shared additional information "off the record". This was useful background material for interpretation of the findings.

An interview schedule was developed based on previous work conducted in Norway (Throndsen et al., 2009), and from work in Chile by Florez (2011). The interview method was further inspired by the work of Baird and Lee-Kelley (2009) who conducted an interview study on implementation of national examinations policy in the United Kingdom. Most of the interviews in the present study (50 out of 56) were conducted at the participants' workplaces and schools, except for six interviews that had been conducted by telephone or Skype. The interviews were recorded and transcribed, and the researchers took notes after each meeting to keep track of important questions raised. The interviews were conducted in both Norwegian and English. The Norwegian interviews were transcribed by the first and second authors, who are native speakers of Norwegian. In addition, one research assistant from Norway transcribed the majority of these interviews. The English interviews were divided and transcribed by all four authors. After coding the interviews using Nvivo, parts of the Norwegian interviews were translated into English by Hopfenbeck for this report.

Interviews lasted between 30 to 60 minutes. One participant agreed to be interviewed twice. The reason for this decision was that the participant had had a key role in the Directorate of Education and Training in the development of the AfL programme in Norway and contributed information that enabled us to ask for new interviews with important stakeholders. The second interview was needed to pursue clarifying questions relating to the emerging findings.

The interview schedules started by asking interviewees about their role in the reform and what they understood as the main aspect in the Assessment for Learning reform (see Appendix A). They were asked about what they found to be problematic about implementing this reform, as well as what they thought would be key to a successful implementation. The researchers maintained fieldwork notes and memos between interviews. This allowed the researcher to adapt the schedule for the next interview, based upon previous findings. Due to a strike in Norway among the teachers and researchers in May 2012, the original interview plan had to be changed, and several school interviews were postponed and carried out in late August and the beginning of September, when the school started after summer vacation. Originally, the

plan was to have two interview days every week and use the rest of the week for analyzing and planning as informed by Baird and Lee-Kelley (2009). Due to the strike, this plan was changed and many interviews were postponed a few days after the strike ended.

Coding of interviews

The coding has been developed based on the GCES framework and the AfL programme in Norway, and informed by the first literature review of educational research in the area of policy, implementation, governance, complexity and educational practice. The two first authors coded all interviews conducted in Norwegian. The first and third author coded the interviews in English. The code manual shows the main codes and their definitions.

Code manual: Governing Complex Education Systems

	Code	Reference to
1	Accountability -	Accountability, monitoring results and test scores, how stakeholders are held accountable.
2	AfL - VfL	The programme in Norway, and with particular focus on the four principles to the students, as well as peer and self assessment, assessment criteria, learning goals etc.
3	Governing and Steering	Governing, steering, school leadership, municipality level and steering, counties and steering, strategic thinking of leadership.
4	Implementation	Implementation in practice
5	Knowledge production, Knowledge use	How different stakeholders develop their knowledge, and how they share and use knowledge, can also include the conferences and reference to research evidence
6	Original ideas behind the project	ARF, or Scotland or any other country where it all started, which inspired them, or any other ideas which are used for explanation
7	OECD	OECD only, without specific reference to any of the studies
8	PISA and international testing	PISA; IEA studies, testing and international testing
9	Policy reflections	Political history, reflections, ideas or ideology, interpretations of what has happened based upon political view such as debates about private and state schools.
10	Research	Research in general, as used for policy implementation, evidence or about Norwegian research
11	Steering through Acts	The Education Act, “Forskriften” Acts, Laws
12	Success criteria	Criteria for success in the implementation examples that have been done, personal experiences and success stories.
13	The Norwegian way	Norway, different actors’ beliefs and what is important when implementing AfL in the Norwegian context
14	Trust and Dialogue	Trust among the participants, respect, dialogue, consensus, shared understanding
15	Challenges and problems	Problems or challenges in the project, struggles, fights, short comings, misunderstanding, distrust, confusion etc.
16	Supervision and support from HE	External support such as guiding, tutoring, co-operation
17	National tests	National tests as they are used in Norway
18	Supervision and support from the municipality level	Help and supervision teachers and school leaders have received from the municipality level, regarding the AfL

	Code	Reference to
		programme.
19	Assessment in general	Any assessment practice explained by the students that do not have AfL elements.

One meeting was set up between the leader and the research assistant to check the inter-rater reliability of the coding of two of the English interviews. Based upon analysis in Nvivo9, differences in the use of 2 codes were found. When investigating these differences, it was found that Code 3, Assessment in general, had not been defined clearly enough, and resulted in misinterpretations. The definition was then improved and the researchers agreed on how to adjust the coding. The second difference here was found in the interpretation of a dialogue in the transcripts where the first researcher had an implicit understanding of the meaning from the interview, since she had conducted this particular interview, but the meaning was not clearly seen through the written transcripts. After agreeing about the codes, the researchers went through four of the codes to see how they had been used in the chosen interview.

Checks were specifically conducted on Codes 1, 14, 15 and 16. The reason for checking these codes was that one of them was about the Norwegian context, and could be challenging to define, another code was about accountability, and was really more easy to decide how it would be interpreted, while the code for problems and challenges turned out to be challenging to define, since it was subjective and thus open to interpretation. The transcripts can be interpreted differently, and what would be viewed as a problem for implementation for some, would not necessarily be seen as a problem for others. It was therefore decided that challenges and problems in the analysis should specifically be linked to the implementation process, and only what participants described as being a problem. From the Nvivo analysis, it was shown that the first researcher has chosen to include less text than the second researcher, but overall, they had chosen the same utterances in the coding.

National test scores

The research team was supplied with a list of schools in the intervention programme and the datasets for the national test datasets for Maths, Reading and English in 2010, 2011 and 2012. In order to identify the intervention schools in the main test datasets, the schools taking part in the programme were first matched with the 2012 national test datasets for Maths, Reading and English. For each of the intervention schools, unique identifying codes were sourced from the 2012 national test dataset. This process was carried out using the Vlookup function in Microsoft Excel, and then manual checking to ensure the school was identified correctly. Variations in school name and other irregularities were checked with the Principal Researcher. At this stage, a significant minority (20%) of schools remained without a unique identifier. Next, the school names were checked against the 2010 and 2011 national test spreadsheets. However, these datasets did not contain school unique identifying codes. Therefore, in order to be able to assign unique identifying codes to both the national test datasets and the intervention schools, a master schools list was created. The concatenate function in Excel was used to generate a new unique school-municipality code for each institution in 2010, 2011 and 2012. Duplicates were eliminated, and the remaining schools were checked manually. If schools in the national test spreadsheets could not be matched to an entry with a unique identifier, they were assigned one manually. The intervention schools list was then checked against this master schools list, using a colour-coding system to highlight special cases; for example, schools for which there was no test data and schools that appeared to have closed. The Norwegian Ministry of Education's online database, *skoleporten.no*, was used to check this information and to source the unique identifying codes for those schools that were still lacking. The Principal Researcher conducted a final check of the irregular cases.

ANNEX B. LIST OF INTERVIEWEES

	Position	Interviewer	Date (2012)
1	1 Municipality leader	Therese N. Hopfenbeck	21 May
2	1 Norwegian Researcher	Therese N. Hopfenbeck	21 May
3	1 Researcher, UK	Therese N. Hopfenbeck	24 May
4	1 School leader	Therese N. Hopfenbeck	30 May
5	2 Teachers, 8th grade	Therese N. Hopfenbeck	30 May
6	4 Pupils, 8th grade	Therese N. Hopfenbeck	30 May
7	1 School leader	Therese N. Hopfenbeck	31 May
8	2 Teachers, 5th grade	Therese N. Hopfenbeck	31 May
9	6 Pupils, 5th grade	Therese N. Hopfenbeck	31 May
10	1 School leader (Headteacher)	Therese N. Hopfenbeck	1 June
11	1 Teacher, 6th grade	Therese N. Hopfenbeck	1 June
12	6 Pupils, 6th grade	Therese N. Hopfenbeck	1 June
13	1 Directorate for Education and Training	Therese N. Hopfenbeck	5 June
14	Minister of Education	Therese N. Hopfenbeck	12 June
15	1 Municipality leader	Therese N. Hopfenbeck	15 June
16	3 Teachers, 5th grade	Therese N. Hopfenbeck	15 June
17	7 Pupils, 5th grade	Therese N. Hopfenbeck	15 June
18	1 School leader	Therese N. Hopfenbeck	15 June
19	4 Pupils, 9th grade	Therese N. Hopfenbeck	15 June
20	7 Teachers, 9th grade	Therese N. Hopfenbeck	15 June
21	1 School leader	Therese N. Hopfenbeck	15 June
22	1 Researcher, UK	Therese N. Hopfenbeck	18 June
23	1 Directorate for Education and Training	Therese N. Hopfenbeck	22 June
24	1 Researcher, Norway	Astrid Tolo	25 June
25	1 Supervisor in the project	Therese N. Hopfenbeck	25 June
26	1 Directorate for Education and Training	Therese N. Hopfenbeck	26 June
27	1 Directorate for Education and Training	Therese N. Hopfenbeck	26 June
28	1 Directorate for Education and Training	Therese N. Hopfenbeck	2 July
29	1 Researcher, Norway	Therese N. Hopfenbeck	6 July
30	1 Minister of Education (former)	Therese N. Hopfenbeck	14 August
31	1 School leader	Astrid Tolo	22 August
32	2 Teachers, 6th grade	Astrid Tolo	22 August
33	6 Pupils, 6th grade	Astrid Tolo	22 August
34	1 Municipality leader	Astrid Tolo	27 August
35	1 School leader	Astrid Tolo	27 August
36	4 Pupils, 8th and 9th grade	Astrid Tolo	27 August
37	3 Teachers, 8th and 9th grade	Astrid Tolo	27 August
38	2 Pupils, 5th grade	Astrid Tolo	28 August
39	2 Pupils, 6th grade	Astrid Tolo	28 August
40	2 Teachers, 5th and 6th grade	Astrid Tolo	28 August
41	2 School leaders	Astrid Tolo	28 August
42	2 School leaders	Astrid Tolo	6 September
43	2 Leaders, Municipality level	Astrid Tolo	11 September
44	1 Researcher, Norway	Astrid Tolo	13 September
45	1 Teacher Union, leading position	Therese N. Hopfenbeck	18 September
46	1 Minister of Education, previous	Therese N. Hopfenbeck	18 September
47	2 Teachers, 9th grade	Astrid Tolo	18 September
48	2 Teachers, 5th, 6th and 7th grade	Astrid Tolo	18 September
49	4 Pupils, 6th grade	Astrid Tolo	18 September
50	1 Journalist and writer	Therese N. Hopfenbeck	19 September
51	1 OECD leader	Therese N. Hopfenbeck	19 September
52	4 Pupils, 9th grade	Astrid Tolo	21 September
53	1 Researcher, UK	Therese N. Hopfenbeck/Astrid Tolo	25 September
54	1 Researcher in UK	Therese N. Hopfenbeck	28 September
55	2 Parents	Astrid Tolo	28 Sepembter
56	2 Parents	Astrid Tolo	28 September

ANNEX C. INTERVIEW GUIDE

Interview schedule

Aims of the interview:

- To understand the strategies and motivations/interests behind the implementation work in the Assessment for Learning project.
- To understand the view of policymakers, school leaders, teachers and students about other actors and systems involved in assessment reform processes and the degree of responsibility they assume as part of the process.
- To detect, in the context of their narrative about the process, conflicting points, potential contradictions and aspects perceived as positive by the actors in assessment reform processes and the academics who were working with the schools.

This interview guide is basedon the interview guide used in the Norwegian project by Throndsen, Hopfenbeck, Dale and Lie (2009) and the interview guide developed by Teresa Florez, (2011); it is informed by the OECD Governing Complex Education Systems framework..

Questions and guidelines for the interview:

The interviewer will inform the interviewee about the aims of the interview and the study, and how data will be used in the OECD report and journal articles. The interviewee will be informed about the expected length of the interview, use of recorder, transcription and confidentiality, and will also be given time to clarify questions. All participants will sign a letter of consent about the project.

Interview guide: Policy makers and researchers

1. Can you briefly describe your background and your role in this project?
2. In your view, can you explain when and why Assessment for Learning has become a focus in the Norwegian policy?
3. (What do you think has been the driving force behind this project?)
4. What have been the key ideas in AfL as you understood them?
5. Before implementing this project, what kind of research evidence on AfL did you look at? Was it from research or the field of practice?
6. Could you briefly explain the process, starting from the idea of using AfL, to a programme designed to be implemented in practice?
7. What were, in your opinion, the main challenges experienced in the process? To what causes do you attribute these challenges? Were they overcome? How?

8. What level of interaction with other units of the DET/Ministry did the process involve? Were there differences with some units in terms of their perspective or their practices on assessment?

9. What do you think of the current situation? How would you describe the implementation process so far? Do you have any indications or research evidence so far of success or failure?

10. In your view, what has been the most important and most challenging factor in the implementation of the project?

11. Given the experience in Norway, what are the major challenges of implementing a strategy programme such as Better Assessment?

12. In your view, who are the main stakeholders in education that are being held accountable for their actions and how is this done? To what extent do you believe school leaders are held accountable for implementing the new reform (Accountability)

13. What are the main competencies that are needed and lacking in Norway regarding how to implement new BV reforms? (Capacity building)?

14. How are the local and central level working together to implement new strategies and what are the major challenges for these collaborations?

15. How are the successes of the new implementations measured?

16. What is happening with the schools that are not succeeding with this project?

17. Looking back, if you could start all over again, would you have changed anything about the project?

18. What do you predict will happen within the next few years with this reform?

Interview guide: School leaders

1. Can you briefly describe your background and your role in this project?

4. In your view, can you explain when and why Assessment for Learning has become a focus in the Norwegian policy?

5. What do you think has been the driving force behind this project?

6. What have been the key ideas in AfL as you understood them?

7. Before implementing this project, what kind of research evidence on AfL did you look at?

8. Could you briefly explain the process, starting from the idea of using AfL, to a programme designed to be implemented in practice?

9. What were, in your opinion, the main challenges experienced in the process? To what causes do you attribute these challenges? Were they overcome? How?

10. What level of interaction with the municipality level did the process involve?

11. What do you think of the current situation? How would you describe the implementation process so far? Do you have any indications or research evidence so far of success or failure?

12. In your view, what has been the most important factor in the implementation of the project?

13. In your view, what has been most challenging in the implementation of this project?

14. Given the experience in Norway, what are the major challenges of implementing a strategy programme such as Better Assessment?

15. In your view, who are the main stakeholders in education that are being held accountable for their actions and how is this done? (Accountability)

16. What are the main competencies that are needed and lacking in Norway regarding how to implement new reforms? (Capacity building)?

17. How are the local and central level working together to implement new strategies? What are the major challenges in these collaborations?

18. To what extent do you believe school leaders are held accountable for implementing the new reform?

19. How are the successes of the new implementations measured?

20. What is happening with the schools that are not succeeding in this project?

21. Looking back, if you could start all over again, would you have changed anything in the project?

22. What do you predict will happen within the next years with this reform?

Interview guide: Teachers

Start by asking about general information, years of experience, type of school, type of initial education. Volunteered or not to be part of this project?

1. Can you briefly describe your background and your role in this project?

2. In your view, can you explain when and why Assessment for Learning has become a focus in the Norwegian education policy?

3. What do you think has been the driving force behind this project?

4. What have been the key ideas in AfL as you understood them?

5. Before implementing this project, what kinds of research evidence on AfL did you look at?

6. Could you briefly explain the process, starting from the idea of using AfL, to how you implement your classroom practice?

7. What were, in your opinion, the main challenges experienced in the process? To what causes do you attribute these challenges? Were they overcome? How?

8. What do you think of the current situation? How would you describe the implementation of Assessment for Learning so far? Do you have any indications or research evidence so far of success or failure?

9. In your view, what has been the most important factor in the implementation of the project?

10. In your view, what has been most challenging in the implementation of this project?

11. Given the experience in Norway, what are the major challenges of implementing a strategy programme such as Better Assessment?

12. In your view, who are the main stakeholders in education that are being held accountable for their actions and how is this done on the basis of the BV project? (Accountability)

13. What are the main competencies that are required and lacking in Norway regarding how to implement new BV reforms? (Capacity building)?

14. How are the local and central level working together to implement new strategies? What are the major challenges in these collaborations?

15. To what extent do you believe school leaders are held accountable for implementing the new reform?

16. How are the successes of the new implementations measured?

17. What is happening with the schools that are not succeeding with this project?

18. Looking back, if you could start all over again, would you have changed anything in the project?

19. What do you predict will happen within the next years with this reform?

Interview guide: Students (approximately age 14)

1. Can you in a few words explain what you think are the main ideas of this project?

2. (Follow-up questions: assessment criteria, peer-and self-assessment, autonomy)

3. When did you start with this project in your school?

4. Do you think this project is important? Why or why not?

5. What are the main challenges with Assessment for Learning in the classroom?

6. What are your class and your teacher doing to solve these problems?

7. What do you think of your classes today and the project? How would you describe your learning activities in this process so far? What have you learned from this project? Are there any problems with this project?

8. In your view, what was the most important thing you learned from this project?

9. In your view, what has been most difficult in this project?

10. If you were going to advise students in another school about this project, what would you tell them to do?

11. If you were asked to give teachers in another school some advice, what would you tell them?

12. Looking back, if you could start all over again, would you have done anything different in this project?

13. What do you think will happen next year? Will you continue working with Assessment for Learning in your class?

Thank you for your time and contribution to this research as participants.

ANNEX D.
WEBSITES FOR THE MEDIA ANALYSIS: EXAMPLES FROM DIFFERENT CATEGORIES

1) Information about AfL

Demokraten, [DET, Vivi Bjelke explaining why feedback is important], www.demokraten.no/lokalnytt/article3897247.ece

Utdanning, [A teacher journal in Norway. Introducing the new programme in AfL], http://utdanningsnytt.no/4/Meny-B/Grunnskole/Utviklingsarbeid/Ny-vurdering-proves-ut/.

[School leader describing the principles behind AfL in a text to the school magazine], http://utdanningsnytt.no/4/Meny-A/Meninger/Innspill/Vurderingspraksis/.

[A teacher using the arguments from Stobart, Black and Wiliam to explain why grades among young students will not be good for their motivation and learning], www.utdanningsnytt.no/4/Meny-A/Debatt/Debattinnlegg-2012/Karakterar-pa-7-trinn/ (accessed 27 March 2012).

Dagbladet Elevvurdering krever kompetanse, www.dagbladet.no/kultur/2007/08/23/509747.html.

2) Schools working on AfL

Aftenposten, [Descriptions from schools working with AfL before the official programme started], www.aftenposten.no/nyheter/iriks/Stiller-toffere-krav-til-skoleelevene-6589420.html.

Fredrikstad blad Hvalerlaerer til Dublin, www.f-b.no/nyheter/hvaler-lererne-til-dublin-1.7506245, (accessed 24 August 2012).

Verdens Gang Krever bedre tilbakemeldinger til eleven, [The Minister speaking to teachers at their annual teacher union conference about the need for better feedback from teachers to students. She was rewarded with a 45 second applause], www.vg.no/nyheter/innenriks/elevavisen/artikkel.php?artid=582520.

3) Problems in relation to assessment

Dagens Næringsliv, Høyreskolen [Editorial discussing the claims on cheating with the national test scores in Oslo].

Aftenposten, Nasjonale prøvelser, by Astrid Søgnen, [Defending use of national tests in Oslo], www.aftenposten.no/meninger/debatt/Nasjonale-provelser-7097451.html (accessed 21 January 2013).

Store variasjoner ved standpunktkarater-praksis, by Kristine Gru Langset ,[Accusations of variation in grading in Norway: The Directorate explain that AfL programme can be one way of strengthening

teachers' assessment literacy], www.aftenposten.no/nyheter/iriks/Stor-variasjon-ved-standpunktkarakter-praksis-5109799.html.

Bergens Tidende Flest klager på karakteren 4, by Per Lindberg, [Complaints on grading in Norway. But, increased focus upon assessment raised less complaints in one of the counties?], www.bt.no/jobb/Flest-klager-pa-karakteren-4-2812106.html (accessed 13 December 2012).

Vil ha nasjonale prøver [Former Minister of Education Øystein Djupedal introducing the new programme], www.bt.no/nyheter/innenriks/Vil-ha-nasjonale-prover-1833251.html.

Nordlys Ideologisert skoleforskning, *Magnus Maeland* [Tory party suggesting a debate about the different advice given by Norwegian researchers], www.nordlys.no/kronikk/article6080513.ece.

Moss Avis [The 7th grade students wanted grades, but DET said no, students should learn from feedback, not grades], www.moss-avis.no/nyheter/direktoratet-sa-blankt-nei-til-tallkarakterer-1.5030387 .

Tønsberg blad [Researcher warning about the misuse of tests in classrooms and how this discourages students. He argues that teachers and school leaders have misinterpreted DET and argues for using AfL without too much testing], http://tb.no/meninger/debattartikler/kutt-ut-fredagsprovene-1.6840495.

[Answer to the researcher from a teacher who disagrees with his views about the misuse of tests in classrooms], http://tb.no/meninger/debattartikler/kutt-ut-fredagstestene-1.7069713.

Vårt land [National tests. Professor Peder Haug is concerned about the results on national tests], www.vl.no/samfunn/forsker-skoler-kan-bli-taperfabrikker/ .

Skolen besatt av kontroll ,Stig Johannesen NTNU, [Researcher warning that there is too much control in Norwegian schools, www.vl.no/samfunn/skolen-besatt-av-kontroll/.

Adresseavisen, Åpenhet gir bedre skole [About publishing school results in Trondheim, the fourth-biggest city in Norway], www.adressa.no/nyheter/innsyn/article3237673.ece#cxrecs_s (accessed 9 July 2012).

[Teacher Union arguing against publishing of school results. The Student Union agrees with this point], www.adressa.no/nyheter/trondheim/article6491681.ece, (accessed 19 December 2012).

Bergensavisen, Økning i klagesaker i Bergen: Kraftig økning i karakterklager [Complaints about grading], www.ba.no/nyheter/article5835804.ece.

Grades in school, [Researcher Åge Diseth commenting upon research showing that grades do not motivate students], www.ba.no/nyheter/politikk/valg2011/article5710194.ece.

Budstikka, Best paa klasseledelse- daarligst paa vurdering [Budstikka, Best in class management; highest in assessment], www.budstikka.no/nyheter/best-pa-klasseledelse-darligst-pa-vurdering-1.7096511.

Bygdebladet Randaberg og Rennesoy, [About national tests], http://bygdebladet.no/index.php?page=vis_nyhetandNyhetID=1446andsok=1.

Sandefjord blad, [Interpretations of national tests scores over three years claiming the municipality increased the achievement], www.sb.no/nyheter/meninger/gode-skoler-utdanner-og-stimulerer-1.7587855.

Nrk. no, Prøvepress fører til skolejuks, [Norwegian broad cast claiming test pressure leads to cheating] www.nrk.no/nyheter/innenriks/valg/valg2011/1.7760062.

4) Think tanks

Civita, Vedens beste skole, www.civita.no/2008/01/14/verdens-beste-skole .

Manifest Analyse, "Testing Testing" [movie interviewing Robin Alexander and Diane Ravitch about testing in U.K. and U.S.

RECENT OECD PUBLICATIONS OF RELEVANCE TO THIS WORKING PAPER

Fazekas, M. and T. Burns (2012), "Exploring the complex interaction between governance and knowledge in education", *OECD Education Working Papers*, No. 67, OECD Publishing, doi.org/10.1787/5k9flcx2l340-en

Hooge, E., T. Burns and H. Wilkoszewski (2012), "Looking Beyond the Numbers: Stakeholders and Multiple School Accountability", *OECD Education Working Papers, No. 85*, OECD Publishing, doi.org/10.1787/5k91dl7ct6q6-en

OECD (2011), Reviews of National Policies for Education: Improving Lower Secondary Schools in Norway 2011, OECD Publishing. doi.org/10.1787/9789264114579-en.

OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing. doi.org/10.1787/9789264190658-en

Snyder, Sean (2013). "The simple, the complicated, and the complex: Education reform and governance through the lens of complexity theory", *OECD Education Working Papers*, No. 96, OECD Publishing, doi.org/ 10.1787/5k3txnpt1lnr-en

Van Twist, M, M. van der Steen, M. Kleiboer, J. Sscherpennise and H. Theisens (2013), "Coping with very weak primary schools: Towards smart interventions in Dutch education policy", A Governing Complex Education Systems Case Study, *OECD Education Working Papers*, No. 98, OECD Publishing, doi.org/10.1787/5k3txnpnhld7-en

Waslander, S., C. Pater and M. van der Weide (2010), "Markets in Education: An Analytical Review of Empirical Research on Market Mechanisms in Education", *OECD Education Working Papers*, No. 52, OECD Publishing, doi.org/10.1787/5km4pskmkr27-en

THE OECD EDUCATION WORKING PAPERS SERIES ON LINE

The OECD Education Working Papers Series may be found at:

- The OECD Directorate for Education website: www.oecd.org/edu/workingpapers
- The OECD's online library, SourceOECD: www.sourceoecd.org
- The Research Papers in Economics (RePEc) website: www.repec.org

If you wish to be informed about the release of new OECD Education working papers, please:

- Go to www.oecd.org
- Click on "My OECD"
- Sign up and create an account with "My OECD"
- Select "Education" as one of your favourite themes
- Choose "OECD Education Working Papers" as one of the newsletters you would like to receive

For further information on the OECD Education Working Papers Series, please write to: edu.contact@oecd.org.