

Comparison and benchmarking as key elements in governing processes in Norwegian schools

Guri Skedsmo, Associate Professor

Department of Teacher Education and School Research

University of Oslo

guri.skedsmo@ils.uio.no

Abstract

Over the past 15 years, the assessment of the Programme for International Student Assessment (PISA), which is sponsored by the Organisation for Economic Co-operation and Development (OECD), has become a strategically important actor in international education. This chapter aims to demonstrate how the OECD, and particularly the PISA, has influenced assessment policies and school governing in Norway. First, I examine the key functions of international and national assessment policies. Second, I explore the ways of which the use of national test results is included in local quality assessment systems and how superintendents, principals and teachers perceive the use of these results. Based on this, I will discuss the functions of national achievement tests, how comparison and benchmarking are integrated in school governing and the possible consequences for accountability practices in which visibility creates pressure on key actors in the school system.

Introduction

Over the past 15 years, the assessment of the Programme for International Student Assessment (PISA), which is sponsored by the Organisation for Economic Co-operation and Development (OECD), has become a strategically important actor in international education policy debates. In a brochure delineating the PISA results in 2015, the Secretary-General, Angel Gurría, stated the following:

“Over the past decade [...], PISA has become the world’s premier yardstick for evaluating the quality, equity and efficiency of school systems. By identifying the characteristics of high-performing education

systems, PISA allows governments and educators to identify effective policies that they can then adapt to their local contexts” (OECD, 2016, p. 2).

The main function of PISA is to describe, monitor and benchmark important aspects of education systems all over the world (cf. Howie & Plump, 2005). This is a type of governing which is relatively new, but at the same time, it has a long history in terms of systematic collections of demographic and economic data used by states to monitor their populations (Ball, 2015). In education, the numbers are fundamental to the constitution of the modern school in the form of examination and tests where the results are used to categorize, compare, rank and position individuals, organisations and systems. As a consequence of neoliberal policies in many countries, the data have increased importance and represent management tools for constantly improvements and they are linked to mechanisms of reward and sanction to boost performance (Gunter et al., 2016). England, the US and Australia are often characterized as countries where neoliberal policies and managerialism were embedded at an early stage. These governing ideas and modernization efforts have needed more time to adapt to gain influence in other countries. Norway, for instance, has been characterized as a late comer (Møller & Skedsmo, 2013; Skedsmo, 2009).

Through PISA it is argued that the OECD will assume “a new institutional role as arbiter of global governance, simultaneously acting as diagnostician, judge and policy advisor to the world’s school system” (Meyer & Benavot, 2013, p. 9). The role of judge implies global accountability measures which classify and rank students, educators and school systems from diverse cultures and countries using the same standardised benchmarks. At the centre of this type of governing are data and data systems that construct policy problems and frame policy solutions across national contexts (Nóvoa & Yariv-Marshall, 2003; Ozga, 2009, 2012). Moreover, the data and the use of data are presented as politically and ideologically neutral and calculable (Pettersen et al. 2017).

Shortly after the first PISA results were published in 2001, a national quality assessment system (NQAS) was implemented in the Norwegian school system. The Norwegian PISA performance, which was below the average of OECD-countries, was important in legitimising these new assessment policies (Elstad & Sivesind, 2010), and both national tests and international comparative assessment studies represented new creations as the NQAS was introduced. Although the national tests take competency aims in the national curriculum as a point of departure, the use of test results shows that the national tests have a similar function nationally and in local municipalities and schools as PISA has for its member countries. According to

current national educational policies in Norway, the key to improvement lies in the use of performance data and output controls. Key actors, such as local authorities, school principals and teachers, are expected to use this information to improve their practice in ways that enhance student outcomes, particularly students' performance on national tests (Skedsmo, 2009).

This chapter aims to demonstrate how the OECD, and particularly PISA, has influenced assessment policies and school governing in Norway. First, I examine the key functions of international and national assessment policies. Second, based on analysis of policy documents and interview data, I explore the role of national test results in local quality assessment systems in three municipalities and how superintendents, principals and teachers perceive the use of the results.

Characteristics related to school governing in Norway

In the Norwegian context, the municipalities¹ have the responsibility for primary and lower secondary schools. In the early 2000s, they were defined as *school owners* (White Paper No. 30, 2003-2004). Local responsibility for quality assurance and development, and as part of this also continuous professional development of teachers, was regulated by law (The Education Act, 1998, Regulation to the Education Act, 2006). Different typologies have been used to characterise school governing in the Norwegian education context. Lundgren (1990) used a typology originally consisting of three steering systems: *the legal steering system*, *the economic steering system* and *the ideological steering system*. He later added *evaluation* as a fourth system. In the Norwegian context, evaluation emerged during the 1990s, along with the introduction of governing concepts such as *management by objectives* and *results* (Engeland, 2000; Engeland & Langfeldt, 2009; Engeland et al., 2008; Karlsen, 2006). While the other three systems can be characterised as *effecting tools*, evaluation represents an *information-gathering tool or instrument* (Hood, 2007). The information gathered can be used by educational authorities for different purposes. Looking at the relationship among the four different systems, Lundgren (1990) claimed that the more the first three systems loosen up and, thus, allow room

¹ Per 1. January 2018 there are 422 municipalities in Norway and many of them are quite small (approximately one fourth of the municipalities have more than 10.000 inhabitants).

for various interpretations, the more evaluation gains dominance as a governing system that provides interpretations in an operational way.

Traditionally, the ideological dimension has been strong in the Norwegian context. Building a comprehensive education system providing equal opportunities for everyone became a highly prioritised goal in Norway towards the end of the eighteenth century. The core of these ideas implies that regardless of gender, residence and socio-economic background, all students have the right to an equitable education that is adapted to their abilities and interests (Karlsen, 1993). Important tools in realising these political aims were, first of all, the Education Act, in other words, the *legal steering system*; the national curriculum, which refers to the *ideological steering system* and defined the overall purposes of public schooling, as well as the aims and content for the individual subjects; and finances, in other words, the *economic steering system* (Bachmann et al., 2008; Lundgren, 1990; Sivesind & Bachmann, 2008). With respect to the political aims for Norwegian schooling, there are questions of whether certain ideologies underpin all these steering systems and how they connect to the realities of knowledge and schooling.

Until the 1990s, curriculum guidelines were considered the most important tool for school governing in the Norwegian context (Bachmann et al., 2004; Sivesind, 2008; Sivesind et al., 2003). The curriculum provided the aims and guidelines for content and methods, as well as for student assessment. Until 1998, a national centre (In Norwegian: *Nasjonalt læremiddelsenter*) was responsible for approving the text books which were going to be used in schools. In other words, until the early 2000s, there has been no focus on testing student achievements and assessing outcomes according to performance standards. Instead, there has been a qualification system based on the examination system and overall assessment grades. These tools have served as sorting mechanisms for further education and working life (Hopmann, 2003; Lundgren, 2003; Sivesind & Bachmann, 2008; Werler & Sivesind, 2007; Tveit, 2014).

This emphasis on input-oriented school governing seemed to be the concern of the OECD reviewers in 1988. In their report, they noted that central authorities were working determinedly to attain national aims but raised questions concerning how these authorities could form an opinion of and influence the level of quality in a school system as strongly decentralised as the Norwegian school system (OECD, 1988). The report stated that monitoring and evaluation functions seemed largely absent from the Norwegian education system. It pointed out that the traditional tools, like the curriculum guidelines, were no longer sufficient to develop the

education system within the frameworks agreed upon in Parliament and that there was a need for national educational authorities to adopt an evaluative and monitoring function to fulfil their responsibilities. The reviewers emphasised that their concern was not to reintroduce national control but, rather, to consider ways in which “good norms of educational practice” (OECD, 1988, p. 45) could be established and disseminated.

This message from the OECD reviewers that more attention had to be given to the evaluation of educational processes and outcomes is also related to the introduction of management by objectives, or *målstyring* in the Norwegian context, towards the end of the 1980s. The concept of *målstyring* was linked to governing ideas in terms of *virksomhetsplanlegging* which refers to administrative policies aiming to reform public administration, in particular decision-making behaviour by the means of goals and programmes which influence the formal organisation, personnel and working methods (Christensen & Lægneid, 1998). Both management by objectives and the administrative reform policies implied a stronger focus on aims and goals as a way of defining direction to guide the work in schools. During the 1990s, such administrative policies were presented as New Public Management reforms. Following the OECD report from 1988, numerous policy documents were worked out by different committees at the behest of the Royal Ministry of Education and Research. They all discussed aspects related to the forms and functions of evaluation, and the discussions have been described as “muddling through” (Hopmann, 2007). Fifteen years after the OECD report, a national evaluation system was introduced in 2005 which, in many ways, can be described as a shift in the Norwegian educational policy away from the use of input-oriented policy instruments towards a more output-oriented policy.

Theoretical perspectives

In this governing regime, a lot of faith is put into the assessment tools that provide data and information to improve practice. It implies, in certain ways, that tools and devices put into play structure policies (Lascoumes & Le Gales, 2007). The implementation of new assessment policies depends, however, on how key actors interpret and respond to the meanings they carry (cf. Ball, 1994, 2008). In this chapter, I focus on national testing as a tool and how the use of its data by local key actors, such as superintendents, principals and teachers, is influenced by a benchmarking logic which connects to policies advocated by the OECD.

According to Lascoumes and Le Gales (2007), an analytic approach has certain advantages. First, it implies a stronger emphasis on the concrete procedures established to attain objectives, which makes it possible to study school governing processes in a more material form. Second, it considers that such instruments “organize specific social relations between the state and those it is addressed to” (Lascoumes & Le Gales, 2007, p. 4). By this, we acknowledge that every instrument constitutes a “condensed form of knowledge”, as it pertains to social control and the means of exercising it (Lascoumes & Le Gales, 2007, p. 3). This is particularly useful in a complex education context characterised by governing processes and interaction among multiple actors. Third, such a definition includes the fact that the effects the instruments produce depend on the aims and purposes ascribed to them. This means that instruments are not neutral devices or methods put into a system to accomplish aims. The tools may seem neutral, but they inherently entail underlying assumptions in terms of values, interpretations and meaning which influence their modes of regulation and possible effects (Lascoumes & Le Gales, 2007). Depending on how the data are perceived and used by the key actors in the school system, for example, by the superintendents at the municipal level, principals and teachers, and the extent to which they are tied to accountabilities, they represent rather strong means of school governing (cf. Hood, 2007; Lascoumes & Le Gales, 2007).

Methodological approaches

The analysis in this chapter mainly draws on data from interviews conducted with superintendents, principals and teachers in three municipalities. To depict some of the variation among the Norwegian municipalities, they were selected from three different counties in Norway according to the principles of maximum diversity (Yin, 2009). The municipalities differ according to structure, organisation and quality management systems, as well as the political steering in each municipality. Three schools were selected in each municipality according to criteria such as school type and locality. Key documents regarding municipal quality management systems and annual reports about the school sector were collected and analysed as well as school development plans and evaluation reports. Moreover, interviews with different key actors were conducted from January 2012 until May 2013, as shown in Table 1.

Table 1. *Overview of Conducted Interviews*

	Group interviews with teachers (approx. 45 min.)	Interview with principals (approx. 60 min.)	Interview with superintendents (approx. 60 min.)
Riverside municipality (3 schools)	7	3	1
Waterfalls municipality (3 schools)	3	3	1
Lakeview municipality (3 schools)	3	3	1
Total: 46 interviews	13	9	3

Altogether, three superintendents at the municipal level were interviewed. In each school, the principal was interviewed as well as one group of teachers (2–4 persons).² The overall aim of the study was to investigate how these key actors perceive and use the results of national tests. All the interviews were recorded and then transcribed verbatim. The transcripts were then coded and analysed.

The role of transnational bodies and international assessment studies

The large-scale comparative achievement studies, in particular the PISA, Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS), are present in educational policy debates as well as in the educational research community. It is, however, important to note that international comparative studies such as PISA are not new creations. Researchers have pointed to the International Examinations Inquiry from the 1930s, which was financed by the Carnegie Foundation in the US, as marking the beginning of international collaboration on quality assessment in education (Jarning, 2010). Since the foundation of the International Association for the Evaluation of Educational Achievement (IEA) in 1959, there has been formalised international collaboration on quality assessment in education. The IEA was founded by a small group of educational and social science researchers with the purpose of conducting international comparative research studies focused on educational achievement and its determinants. According to Gustafsson (2008), the development of the IEA was characterised by two phases. In the first phase, from around 1960

² Riverside has more interviews than the other schools because this school was also used as the pilot for the study.

until 1990, the researchers pushed the development forward. The second phase was characterised by strong influence from policymakers and administrators. I argue that the transition of the international studies from phase one to phase two, with respect to the organisation of the IEA and its main focus, is important to understand what characterises these studies as policy tools and the ways in which they influence school governing.

According to Gustafsson's (2008) summary of the first phase of the history of the IEA, the researchers' aim was to understand the great complexity of factors influencing student achievement in different subject fields. They used the popular metaphor of the world as an "educational laboratory" to investigate the effects of school, home, student and societal factors, and they argued that an international comparative approach was necessary to investigate the effects of many of these factors. In 1990, a new organisation of IEA was set up with a permanent secretariat in the Netherlands and a data-processing centre in Hamburg. This second phase was, according to Gustafsson (2008), characterised by the following changes: First, there was a dramatic increase in the volume and frequency of the studies, as well as in the number of countries participating in the studies. Second, the involvement of national administrative and policy institutions became stronger. At the same time, the researcher presence was less marked. Researchers were still involved in the design, analysis and reporting of the studies, but the level of ambition of the international reporting was limited. The task of analysing the factors behind the outcomes for the different countries was left to each participating country, and the databases were made available to the research community for secondary analysis. Third, the focus of the studies shifted away from explanations to descriptions, which were used as a basis for national policy discussions and decisions.

When the Organisation for Economic Cooperation and Development launched the PISA study in 2000 (OECD, 2001), the emphasis of the international comparative studies related to educational policy became even stronger. An information sheet about the history and the future of PISA from 2007 states that PISA is a "collaborative effort" which brings together scientific expertise from the participating countries and which is "steered jointly by their governments on the basis of shared, policy-driven interests" (OECD, 2007, p. 10). The information sheet emphasises that an important aim is to produce and describe results which can be used to inform national policy makers. The role of the policy makers in further developments of PISA is also emphasised, as illustrated in this quote: "Above all, this evolution is guided by the priorities of educational policy makers, who want to ensure that further changes in education systems are

firmly rooted in good evidence” (OECD, 2007, p. 17). This implies that the priorities and decisions related to the development of test design are policy driven, and it can be argued that the data produced reflect policy interests. However, the international and comparative aspects in the descriptions imply the comparison of school systems from diverse cultures and countries using the same standardised benchmarks. The main function of PISA is to describe, monitor and benchmark important aspects of education systems all over the world.

The Norwegian National Quality Assessment System

Although the move towards more output-based forms of school governing started long before PISA, it has been argued that the publication of test results, in which Norwegian students scored below the average of students in the other OECD countries, helped the argument for introducing national testing in the Norwegian context (Elstad, 2008; Langfeldt, 2008; Skedsmo, 2009).

The National Quality Assessment System (NQAS), as it was introduced in 2005, comprised a mix of new and traditional tools. The national tests and the international comparative achievement studies, such as PISA, TIMSS and PIRLS, were new inventions. To some extent, its screening tests and information material could also be defined as new tools, while its formative and summative assessment of students in terms of local tests could be characterised as traditional. The School-leaving Examination and the Craft Certificate were also part of the NQAS. Since these elements had constituted the examination system in secondary school for quite some time, they could be categorised as traditional tools. However, it should be noted that they had not, until recent years, been used to provide educational statistics for governing purposes or to legitimise political decision making in a systematic way.

The overall aim of the comprehensive national evaluation system was “to contribute to quality development on all levels of compulsory education with respect to adapted teaching and improved learning outcomes for the individual student” (The Norwegian Directorate for Education and Training, 2005. My translation). In addition, the national system was supposed to provide information for the education sector about the national and local state of progress, which could be used to form the basis for general decision making and for local work on evaluation and development. At the same time, the system should contribute to increased

openness, transparency and dialogue about the school's practice (White Paper No. 30, 2003-04).

The guidelines formulated for the different tools, such as the national tests, state that their aim is to investigate the extent to which students' achievements align with the aims of competencies in the national curriculum. However, national tests are also supposed to inform the students and other key actors about the level of achievement as a foundation for improvement (The Norwegian Directorate for Education and Training, 2012). These two purposes exemplify what has been the focus in all the policy documents discussing frameworks and questions related to establishing a national quality assessment system (Skedsmo, 2009).

The international comparative achievement tests have two functions. First, they make it possible to evaluate and compare Norwegian students' levels of achievement with those of students in other countries. Second, they are intended to provide insight for policy formulation and to develop national quality indicators (The Norwegian Directorate for Education and Training, 2012). This last purpose seems to follow up on what the European Report on the Quality of School Education addressed as the "challenge of data and comparability" of PISA, which was identified as "the need to set quantifiable targets, indicators and benchmarks as a means of comparing best practice and as instruments for monitoring and reviewing the progress achieved in order to provide a basis for educational policy making" (European Commission, 2000, p. 16).

This benchmarking logic is also emphasised within the national policy context. The stated purpose for the use of national tests pertains to evaluating the students' basic competencies as they relate to national aims. It is also about evaluating the extent to which the schools succeed in developing basic competencies among their students. The way the results are used indicates that comparisons are rather strong driving motivational forces. For instance, when summing up the results from the national tests of 2007, the Norwegian Directorate for Education and Training emphasised the comparison of boys' versus girls' results (2008, 2017a, 2017b). Nothing was mentioned about the extent to which national aims had been reached. The gender differences in level of achievement within the national context were also compared to the gender differences on the PISA results. Due to widespread critique about the publication of the results and ranking of schools in the first round of national testing, it was decided in 2007 that the results should not be publicly accessible, at least not at the school or individual level, and the tests should be taken at the beginning of the school year (5th and 8th grades) to strengthen the tests' formative aspects. The results were, however, still used to compare counties and

municipalities and to rank schools if the media got hold of the results. Further, studies have shown that schools use the results for comparison with other schools within the same municipality and between municipalities within regions (e.g. Elstad, 2009). In 2015, it was again decided to make the results publicly accessible. The way the data are currently presented on the online platform promotes the comparisons of counties, municipalities and, in particular, schools within a municipality.

Use of national test results by key actors in municipalities and schools

All three municipalities in this study have implemented *management by objectives* and the use of balanced scorecards as the key elements in their governing systems, but they differ largely regarding how the systems are put into play, particularly in how performance management and accountability are played out. Based on local policy documents and interviews with three superintendents, I will first describe how national testing is integrated in local quality assessment systems (LQAS) and how superintendents at the municipal level make use of these data. Then I move on to how principals and teachers perceive the use of these data on school level.

Characteristics of LQAS and the superintendents' use of national test results

Over the last decade, Riverside, the first municipality examined in this study, has developed a quality management system which consists of two main areas: evaluation of the school as a learning organisation and evaluation of the students' learning outcomes. The most important assessment tools and data sources are the annual national student survey³ and the national standardised tests. "Together", the superintendent stated: "these two information sources provide a good picture of the current status of school quality". Compared to the other two municipalities, Riverside has changed practice profoundly in terms of developing tighter control as a consequence of having national test results.

In Waterfall, test results were already an important part of local school governing and quality management before the national standardised testing was launched. Over the last 15 years, they have developed a quality management system which includes different types of quality

³ The national student survey (*Elevundersøkelsen*) is organised by the Norwegian Directorate for Education and Training and conducted annually to collect data about students' perceptions of their learning environment.

indicators: 1) key numbers and facts about the schools, their resources and working conditions, 2) school results in terms of student outcomes on national standardised tests, local and national screening tests, the National Student Survey and the municipal survey on the students' learning environment and 3) self-evaluation of efforts and results on strategic development areas. The national tests have been added to the system they already had in place, and the school administration has developed an extensive plan for conducting all the tests throughout the school year.

In contrast, Lakeview has followed a different course of evaluation and school development. Earlier they had a peer review approach to school evaluation in which they collaborated with six other municipalities that are all part of the Seven Star Network⁴. In this approach, a team of teachers from one municipality had the responsibility of conducting school evaluation according to given requirements in another municipality. However, this proved to be costly, and, due to cuts in state transfers, the network could not continue this approach after the pilot project ended. Since 2010, Lakeview has had a quality management (QM) tool which requires that the schools follow up on documenting and reporting on targeted areas, such as school economy, work on quality development, school start and transition between schools, adapted teaching, student assessment, including national test results and other assessment tools integrated in the NQAS, learning environment and collaboration with parents.

The results on the national tests are mainly used by the superintendents for the following purposes:

- To monitor the schools' results over time
- To compare 1) results of schools within the municipality and 2) the municipal average score with the other municipalities and the national average.
- To hold the principals accountable for the level of school achievements
- To decide on improvement efforts

However, the process of assessing quality mainly includes monitoring and comparing current and previous results, where the comparisons with the results of other schools and municipalities, etc., represent the reference points. In two of the municipalities, which have quite different local quality management systems in place, similar patterns can be identified regarding the role of comparison in quality assessment. In the third municipality, the superintendent has another

⁴The Seven Star Network is an inter-municipal collaboration on specific areas such as school evaluation and continuous professional development for teachers.

approach which also includes how the school results are used and monitored and the ways in which accountabilities are put into play.

Waterfall has aimed towards establishing clear performance expectations for all actors throughout the system and a highly transparent and evidence-based system characterised by school results that are publicly accessible in order to make “good practice” wanted and required, as illustrated by the following quote:

My aim was always to establish a system with clear expectations and a practice where we can demonstrate effects on student learning. (Superintendent, Waterfall Municipality)

The superintendent in Riverside had another viewpoint. According to her, making school results publicly accessible leads to frustration and additional pressure for schools with weak results. She pointed out that this could actually hinder school improvement:

The principals in our municipality know the results of all the schools. If we publish the results, we can easily get the wrong focus in terms of ranking schools. Last year, the Municipal board decided that I should present the school results to them and that, in this part of the meeting, the media would not be allowed to be present. But the local press complained about this decision, and they were allowed to be present and report on the results. On the first page, they presented the best and the worst schools. This kind of publicity has a negative effect no matter how professionally we make use of the results. No schools improve by telling them publicly that they are bad. (Superintendent, Riverside Municipality)

In other words, this superintendent clearly objects to ranking schools publicly. The results of each school are presented in her meetings with the principal group, and the principals as well as the teachers are aware of the informal ranking list within the municipality. The superintendent stated that she does compare the results of Riverside with other municipalities and the country average but expressed that her aspiration is not that Riverside should climb the ranking list with other municipalities. She expressed her worries about the consequences if 70 percent of the students in a specific age group in a school do not have sufficient reading skills. To integrate data use in school development work, she has established *result meetings* on different levels of the local governing system. On the municipal level she has result meetings with all the principals in the municipality where the informal ranking list is used as a means to emphasise transparency about the results achieved as well as to hold the principals accountable. On school level, the principals have result meetings with the teacher teams to discuss national test results and what they can learn from them. In addition, a former network meeting between the primary schools and lower secondary schools located in the same area has been transformed. The previous aim of this meeting was to ensure smooth transitions of the students. In the result meeting, the principals and teachers from the different school types meet to discuss the students’

results on national tests in 5th grade (primary school) and 8th grade (lower secondary school), since the tests in 8th grade measure students' competencies as they have just entered lower secondary education.

The superintendent in Lakeview said he sees the national test results as one out of many quality indicators and that he pays more attention to how schools follow up on prioritised improvement areas formulated in their school development plans. He was somehow unconvinced about the strong focus on national test scores in other municipalities:

I don't know... All municipalities in our region aim towards having the best national test results. This is not possible since this always depends on the achievements of the other schools and municipalities. (Superintendent, Lakeview Municipality)

Here, the superintendent pointed out that with the previous test procedures, achievement levels differed from year to year depending on the overall performance of students. If all the schools put great effort into improving their results, he was not certain that this would change the municipalities' positions on the ranking list. Like the superintendent in Riverside, he also has an annual dialogue meeting with the principals and the teachers who are part of the school development group and, in small schools, he involves the whole school staff. In these dialogues, he stresses the school outcomes over time on national tests, the exam results (in lower secondary schools), the results from the Student Survey and results from an annual survey among parents. He stated that, instead of comparing results, he is concerned about how the principals reflect on the achieved results as well as professional arguments related to the schools' actions to improve prioritised areas and the evaluation of the effectiveness of these actions.

Principals' perceptions of national test results

All nine of the principals interviewed regard the results of national testing as an important quality indicator and agree that it is important to do well on the tests. Generally, they use the results for the following purposes:

- To monitor the results of the school compared with the results of neighbouring schools, the average of the municipality and the national average
- To check that "we are on the right track"
- To decide on improvement efforts
- To prioritise areas for professional development
- To legitimate decisions
- To hold teachers accountable and commit them to school priorities

The principals in Riverside regard the outcomes of the national standardised tests as important quality indicators. They use the results to monitor progress over time, which is illustrated by the following quote:

I think the results of the national testing are more exciting than I expected they would be. It is exciting to see if we are moving in the right direction and – to be honest – whether we are above or below the average results in the municipality as well as the average country results. (Interview with the Principal at Pine school, Riverside)

Since national test results are presented by the superintendent, there is an unofficial ranking list among the principals. They are concerned about the local press since there are examples of negative publicity for schools which perform worse than expected, and they find it unnecessary for school leaders and staff to have such an additional burden.

The principals in Waterfall clearly distinguish between different the assessment tools and their functions. Regarding the national tests for the 8th grade, they pointed out that these results must be judged against the competency aims for primary schools and that they should concentrate on improving the basic competencies of the students from 8th to 9th grade. However, they admitted that they are, to some degree, competition oriented, which is illustrated in the quote below:

I think we are not that concerned about comparing...and beating other schools... However, I have teachers and leaders in this school who are very competition oriented. I am too, actually... It is always nice to do well...also compared to others. (Interview with the Principal at Elm School, Waterfall)

Similar to the superintendent in Lakeview, the principals in this municipality regard national test scores as one out of many quality indicators. The principals in the two lower secondary schools find the examination results in 10th grade to be better quality indicators than the national test results. The principal in the primary school is more concerned about the national test results compared to the other principals, and the school has invested money in an ICT-based quality system which provides a good overview of the different types of results where she can compare groups of students and monitor the results over time.

In the dialogues about national test results with teachers, all three principals in Lakeview emphasised the importance of providing support for their teachers in terms of professional development courses or collaboration with other teachers. Unlike the other two municipalities, they have not established any additional arenas to discuss the results but use the meeting structure they already have in place.

The principal in Hazel school in Lakeview pointed out that the cohorts of the students can vary largely. Last year, the 5th-grade students in this school had the best results in the whole region

in numeracy. She was quite proud of these results, but, at the same time, she emphasised that this would not have been possible for the cohort one year earlier. Compared to the other two municipalities, the schools in this municipality highlight school profiles where practical esthetical subjects and outdoor activities are central.

In our municipality, we have many parents who do not have higher education, and the academic achievements of the different cohorts differ largely. In our school, we emphasise that all students should experience success with something, and we make this possible by offering a range of activities. This is also part of the responsibility of schools, namely focussing on the whole student and not only the academic achievements. With the increased emphasis on comparing schools, it might be a danger that we define school quality narrowly and lose sight of the general part of the national curriculum (Principal at Hazel School, Lakeview).

In Waterfall, the principals agreed with the local policies and the strong performance orientation, but they also reflected on the implications of such a one-sided focus. The principal in the schools with the most heterogeneous student group has long experience as a school leader. He stated that he thinks the schools in this municipality have improved their practices profoundly through the focus on test results and comparison between schools, which creates pressure, especially on the principals. At the same time, he also reflected on other important aspects of education which, as a consequence, receive less attention:

You know, it is like this flowery rhetoric...the aspects you focus on, they get better... The general part of the curriculum, for instance, gets less attention, and you only hear about these aspects in speeches. In our school, we struggle with social issues among the students and we need to focus on these issues at the same time. If not...then the consequences could be bad for the students involved. (Principal in Willow School, Waterfall)

Teachers' perceptions of national test results

All the teachers interviewed reported expectations from the principal to perform well on the national tests and that there is a greater focus on students' academic achievements as a consequence of national testing. Generally, the teachers in the study reported that they use the national test results for the following purposes:

- To check that the students are "on track"
- To check which kinds of tasks the students managed and which types of test exercises proved to be difficult

In all three municipalities, the teachers emphasised that national test results confirm what they already know about the students' achievement levels, rather than provide new information.

We discuss the results within the team and the extent to which the score fits with our assessment of the students' achievement level... To be honest, I don't care much about these results... The national tests are not really popular... Maybe they could inspire us to work more systematically and efficiently... No one likes to be at the bottom of a ranking list... The principal in our school was very happy this year since

we improved our results from last year, and she brought chocolate for the staff meeting to celebrate. (Group interview with teachers, 8th grade, Pine School, Riverside)

In Pine school, the teachers were not worried about saying that they did not really care about the national test results. Instead, the teachers laughed and made jokes about the results, such as how they noticed their principles were pleased with the results when she brought chocolate to the school. In contrast, the teachers in Waterfall municipality were much more concerned about the results of the national tests. Here, the teachers explained that the principal is held accountable for the results by the superintendent, and improvement efforts in the school are tighter coupled with the results achieved:

Based on the national testing, you get a report for each student in which the achievement level is stated, as well as how many points the student accomplished. In addition, there is a description of the student's competence. However, these results do not really provide me with any new information about the student. Rather, the tests have a control and accountability function which I see quite clearly in the annual performance appraisals with the principal. (Interview with teacher, 8th grade, Spruce School, Waterfall)

In the lower secondary school with the most heterogenous student group and where the students perform poorly, the national test results seemed to matter more for the teachers in terms of accountability pressure from the superintendent. The teachers expressed that they are happy that the principals have so much experience in handling this pressure and in setting local priorities which are important for the various student groups in the school.

The teachers in all three municipalities stated that they find using the results from the national testing problematic; in their experience, it does not add any new information they did not already have about the students. They compare the data from the different tests that they have to use, and they emphasise the results from the screening tests (e.g. *Kartlegger'n*), which include a pre-test at the start of the school year and a post-test during the spring, as much more appropriate for use in adapting their teaching.

Compared to the teachers in Lakeview and Riverside, several teachers in Waterfall mentioned that some schools go easy on the rules for exempting students from the tests to get better results.

We are doing quite well, even on a national basis. But I think this is actually against all odds, especially when I know that some schools are rather sloppy with respect to the rules for exempting students from the tests. Only four students who just came from the "introduction class"⁵ (*mottaksklasse*) were exempted for the reading test in this school. Apart from them, they all took the test. I think the rules should be stricter. It is not fair when the school can make this decision, especially because the results are published.

⁵ The concept of "introduction class" (In Norwegian: *mottaksklasse*) refers to a class for students from immigrant families who have just come to Norway.

It would not have had such implications if the results were not published. (Teacher at Willow School, Waterfall)

Teachers in both Waterfall and Riverside, which both have more emphasis on consequences if the level of student achievement is considered low, expressed concern about the test results only measuring certain aspects of the students' competencies. They accentuated the need to look at the overall performance of the students throughout the school year in a more holistic way. The teachers in lower secondary schools considered the test results in 9th grade as more useful for them because they can compare the results for the same group of students over time.

Discussion and conclusion

The main functions of the PISA are to describe, monitor and benchmark important aspects of education systems all over the world and use the information provided, the evidence, as a basis for policy making and improving education on national levels. According to the current national educational policies in Norway, the key to improvement lies in the use of national test results and output controls. Key actors, such as local authorities, school principals and teachers, are expected to use the information to improve their practice in ways that enhance student outcomes, particularly the students' results on national tests. The information is presented in tables where it is possible to compare the results of schools and municipalities over time. As such, the data appear to summarise rather complex phenomena and dimensions across different sites and time (cf. Hacking, 1983) and they also indicate a certain neutrality and objectivity. As mentioned earlier, tools such as national tests may seem neutral, but they inherently entail underlying assumptions in terms of values, interpretations and meaning which influence their modes of regulation and possible effects (Lascombes & Le Gales, 2007). For instance, findings from this study show that test data have become a central feature of development work in schools. The way these tests are constructed and the data they provide embody particular representations of teaching, learning and students' success which enable users to see certain aspects related to teaching and learning processes (Spillane, 2012). At the same time, we must be aware that attention to other aspects is constrained. The findings in Riverside municipality with respect to establishing *results meetings* as an arena to link student achievement data to development work show that the national tests have the potential to establish new patterns of interactions between actors in the local governing system which may have concrete implications for teaching and learning in schools.

In many municipalities, new approaches to school governing have been developed which are responsible for the quality of schools, and in policy documents defined as school owners, along with new national expectations about using performance data to enhance educational quality. The introduction of comparative studies and national and municipal testing of student performance has, in many ways, led to new concepts of educational quality in terms of comparison and benchmarking and increased focus on competition among schools within and between municipalities (see also Skedsmo, 2009). By means of comparison, two of the superintendents stated that they aim to increase transparency and openness about school results. Before it was decided by national authorities to make the results publicly accessible again, there was still an official ranking of schools in Waterfall, while the superintendent in Riverside operated with an informal ranking list which was only presented to the principals. The benchmarking seems to represent a means of holding the principals accountable, but the intention is also to drive improvement. In Lakeview, the superintendent has a more holistic, long-term approach in which professional reflection on the results, development aims of the school and improvement efforts are expected. Nevertheless, the use of national test results seems to organise new types of social relations on local levels, first between the superintendents and the principals, and particularly in the two municipalities where the use of results is accompanied by performative accountability. New relations also emerge among principals and schools due to the focus on comparisons and benchmarks, which lead to increased competition among school in the two larger municipalities.

The principals in all three municipalities expressed concern with comparisons where the school's position on the ranking list – if the results are good – represents an important incentive for further work at the same time as it indicates “luck” with the student cohort. If the results are not so good, they like to compare their own school with the performance of similar schools in the municipality or the average results of the municipality. In contrast, the teachers seemed to pay less attention to national test results but, at the same time, understand that the principals are put under pressure to perform well.

Benchmarking was defined by Åkerstrøm-Andersen and Thygesen (2004) in their article *Governing Governing Tools* (In Danish: *Styring av styringsværktøjer*) as making a judgement about the parameters in which different schools, regions or countries differ to be able to close the gap. In certain ways, the results of the international comparative achievement studies seem to form a benchmarking system for assessing the quality and effectiveness of the Norwegian education system in relation to other countries. This can be seen as a consequence of the

knowledge economy and seems to be based on the assumption that there is a close link between a nation's educational achievement and its economic competitiveness (Linn & Baker, 1995). In an article from 1995, Linn and Baker demonstrated how international comparative achievement studies are used to set benchmarks for US performance. This was even suggested by the National Academy of Education panel on the Evaluation of the NAEP Trial State Assessment in 1993. In this article, Linn and Baker also described a movement towards using international assessment results to create world-class standards.

It can be argued that this benchmarking aspect is linked to what Benveniste (2002) referred to as the instrumental function of assessments related to school governing. The instrumental function implies that student tests are devices used to collect "objective" data and are designed to support rational decision making at the classroom as well as policy levels to implement change. However, Benveniste also pointed out that the measurement of student achievement is not necessarily linked to attempts to implement change. Assessment systems can also have symbolic functions. This perspective implies that the primary purpose of assessment is not to uncover deficiencies in education but to appear as if they do, which is motivated by a drive for legitimisation.

In the Norwegian context, comparison with others and benchmarking has emerged as a new concept of assessing educational quality and progress. Even if the results used as a basis are standardised, it implies a normative, fluent concept of quality driven by the monitoring of the positioning of schools and municipalities. The results are used to legitimise national educational policy and local improvement efforts. Interestingly, when national test results are used for accountability purposes, accountability seems, in many ways, to subsume visibility. If student performance on national tests does not meet expectations, it is attributed to the work of teacher teams and schools. As such, the pressure seems to go downwards in the governing chain, and discussions about results seldom involve national and local authorities for their decision making. Compared to other countries, such as England and the US, the schools in Norway are not in danger of getting closed, etc. However, the visibility can represent a serious consequence in terms of school reputation and the public's trust in the local schools. To date, elements linked to a market ideology are only evident in larger cities and, foremost in upper secondary education. However, even for a lower secondary school, it can imply changes in the student population if the parents of high-performing students choose to send their children to another school in the city. In a smaller municipality, negative visibility can have consequences for

individual teachers because it does not necessarily foster professional dialogues on how to improve teaching and learning in a classroom context with specific challenges.

On the national level, it is a question of whether testing and benchmarking creates its own dynamic by introducing new reforms which, in turn, require more testing and assessment to determine their effects (cf. Baker & LeTendre, 2005). On the local level, educators will constantly have to look for ways to help improve education. The question is what kinds of improvement can take place, depending on the types of professional reflection and discussion that take place among teacher and principals when they interact with national test results. So far, studies have shown that there is a tendency towards finding short-term solutions and efforts which aim to improve the next year's test results (Mausethagen et al., 2016). There is a tendency towards the increased standardisation of teachers' work in terms of using reading programmes and establishing routines for how to start and end a lesson, etc. This may help reduce the complexity in the work of professionals but could also reduce professional autonomy.

Although professional autonomy was still emphasised in the latest reform, the Knowledge Promotion (K06), there was a shift in how trust in education was communicated. Trust in the profession itself was replaced by trust in the results (Uljens et al., 2013). While it was argued that the managerial approach to education aimed at ensuring a basic standard for all, presumably equalising disadvantages, it was also a push for de-bureaucratisation and de-centralisation aimed at allowing for more differentiation and specialisation (Møller & Skedsmo, 2013; Paulsen & Høyer, 2016).

At the same time, the welfarist legacy, which emphasises education for the public good, is still strong, and it mediates the reading, interpretation and shaping of international trends. However, conflicting rationales identified as neo-liberalism or technical-economic rationality are gaining terrain. One of the main tensions is between discourses rooted in socially democratic ideologies linked to notions of equity, participation and comprehensive education and discourses of accountability and competition which underpin managerial forms of governing schools. In many ways, national testing and other types of evaluation represent technologies which are used by both superintendents and principals to monitor student outcomes, and this can be read as a shift toward what has been termed *organisational professionalism*, which incorporates standardised work procedures and relies on external regulation and accountability measures (cf. Evetts, 2009). It echoes the management discourse promoted by the OECD, where performance orientation represents a main pillar closely connected to output control.

References

- Bachmann, K., Sivesind, K., Afsar, A., & Hopmann, S. (2004b). *Hvordan formidles læreplanen? En komparativ evaluering av læreplanbaserte virkemidler - deres utforming, konsistens og betydning for læreres praksis*. Kristiansand: Høyskoleforlaget.
- Bachmann, K., Sivesind, K., & Bergem, R. (2008). Evaluering og ansvarliggjøring i skolen (Evaluation and school accountability). In Langfeldt, G., Elstad, E. & Hopmann, S. (Eds.), *Ansvarlighet i skolen. Politiske spørsmål og pedagogiske svar* (p. 94-122). Fagernes: Cappelen Akademisk Forlag.
- Baker, D. P. & LeTendre, G. (2005). *Global Similarities and National Differences*. Stanford, CA: Stanford University Press.
- Ball, S. (2015). Education, governance and the tyranny of numbers. *Journal of Education Policy*, 30(3), 299-301.
- Benveniste, L. (2002). The Political Structuration of Assessment: Negotiating State Power and Legitimacy. *Comparative Education Review*, 46(1), 89-118.
- Christensen, T. & Læg Reid, P. (1998). Administrative reform policy: the case of Norway. *International Review of Administrative Sciences*, Vol. 64(3), 457-475.
- Elstad, E. (2008). Ansvarliggjøringsmekanismer: Når skoleeier eller pressen stiller skoler til ansvar In G. Langfeldt, E. Elstad & S. Hopmann (Eds.). *Ansvarlighet i skolen. Politiske spørsmål og pedagogiske svar* (p. 204-234). Fagernes: Cappelen Akademisk Forlag.
- Elstad, E. (2009). Schools which are named, shamed and blamed by the media: school accountability in Norway. *Educational Assessment, Evaluation and Accountability*, 21(2), 173-189.
- Elstad, Eyvind & Sivesind, Kirsten (Eds.) (2010). *PISA sannheten om skolen?* Universitetsforlaget.
- Engeland, Ø. (2000). *Skolen i kommunalt eie - politisk styrt eller profesjonell ledet skoleutvikling?* PhD thesis. Faculty of Education: University of Oslo.
- Engeland, Ø., & Langfeldt, G. (2009). Forholdet mellom stat og kommune i styring av norsk utdanningspolitikk 1970-2008. *Acta Didactica*, 3(1), 1-16.
- European Commission. (2000). *The European Report on Quality of School Education*.
- Gunter, H., Hall, D., Serpieri, R. & Grimaldi, E. (Ed.): *New Public Management and the Reform of Education: European lessons for policy and practice*. London. Routledge.
- Gustafsson, J.E. (2008): Effects of International Comparative Studies on Educational Quality on the Quality of Educational Research. *European Educational Research Journal*, Vol. 7(1), 1-17.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.
- Hood, C. (2007). Intellectual Obsolescence and Intellectual Makeovers: Reflections on the Tools of Government after Two Decades. *Governance: An International Journal of Policy, Administration, and Institutions*, 20(1), 127-144.
- Hopmann, S. (2003). On the Evaluation of Curriculum Reforms. In Haug, P. & Schwandt, T. A. (Eds.), *Evaluating educational reforms. Scandinavian perspectives* (pp. 111-132). Greenwich, Connecticut: Information Age Publishing.
- Howie, S. & Plump, T. (2005). International Comparative Studies of Education and Large-scale Change. In: Bascia, N., Cumming, A., Datnow, A., Leithwood, K. &

- Livingstone, D. (Eds.) *International Handbook of Educational Policy*, vol. 1. (p. 75-99). Dordrecht: Springer.
- Jarning, Harald (2010). Resultater som teller - Kunnskapskontroll og resultater i skolens århundre. In: Eyvind Elstad & Kirsten Sivesind (red.) *PISA sannheten om skolen?* (p. 199 – 221). Oslo: Universitetsforlaget. ISBN 978-82-15-01487-6.
- Karlsen, G. E. (1993). Desentralisert skoleutvikling. En utdanningspolitisk studie av norsk grunnskole med vekt på 70- og 80-tallet. Oslo: Ad Notam Gyldendal.
- Karlsen, G. E. (2006). Utdanning, styring og marked: norsk utdanningspolitikk i et internasjonalt perspektiv. Oslo: Universitetsforlaget.
- Langfeldt, G. (2008). Resultatstyring som verktøy og ideologi. Statlige styringsstrategier i utdanningssektoren. In G. Langfeldt, E. Elstad & S. T. Hopmann (Eds.), *Ansvarlighet i skolen*. Oslo: Cappelen.
- Lascoumes, P., & Le Gales, P. (2007). Introduction: Understanding Public Policy through Its Instruments - From the Nature of Instruments to the Sociology of Public Policy Instrumentation. *Governance: An International Journal of Policy, Administration, and Institutions*, 20(1), 1-21.
- Linn, R., & Baker, E. L. (1995). What do International Assessments Imply for World-Class Standards? *Educational Evaluation and Policy Analysis*, 17(4), 405-418.
- Lundgren, U. P. (1990). Educational policy-making, decentralisation and evaluation. In Granheim, M., Kogan, M. & Lundgren, U. P. (Eds.), *Evaluation as Policymaking. Introducing evaluation into a national decentralised educational system* (pp. 23-41). London: Jessica Kingsley Publishers.
- Lundgren, U. P. (2003). The Political Governing (Governance) of Education and Evaluation. In P. Haug & T. A. Schwandt (Eds.), *Evaluating Educational Reforms: Scandinavian Reforms* (p. 99-110). Information Age Publishing Inc.
- Mausethagen, S., Skedsmo, G. & Prøitz, T. (2016): Ansvarliggjøring og nye organisasjonsrutiner i skolen – rom for læring? *Nordiske organisasjonsstudier*. Vol. 18/2, 79-97.
- Meyer, H-D. & Benvot, A. (2013). PISA and the Globalization of Education Governance: some puzzles and problems. In: Meyer, H-D. & Benvot, A. (Eds.) *Pisa, Power and Policy – the emergence of global educational governance*. Southampton: Symposium Books, 9-26.
- Møller, J. & Skedsmo, G. (2013): Centralization and Decentralization as Twin Reform Strategies. In: Moos, L. (Ed.): *Transnational Influences on Values and Practices in Nordic Educational Leadership - Is there a Nordic Model*. (p. 61-72). Dordrecht. Springer.
- Nóvoa, A. & Yariv-Mashal, T. (2003). Comparative research in Education: a mode of governance or a historical journey? *Comparative Education*, 39(4), 423-438.
- OECD (1988). *Reviews of National Policies for Education Norway*. Paris: Department of the Examiners.
- OECD. (2007). *PISA - The OECD Programme for International Student Assessment*. Retrieved March, 21, 2009, from <http://www.pisa.oecd.org/dataoecd/51/27/37474503.pdf>
- OECD. (2016). *PISA 2015. Results in Focus*. Retrived February, 2, 2017, from <https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf>
- Ozga, J. (2009). Governing Education Through Data in England: From Regulation To Self-Evaluation. *Journal of Educational Policy*, 24(2), 149-163.
- Ozga, J. (2012). Assessing PISA. *European Educational Research Journal*, 11(2), 166-171.

- Paulsen, J. M. & Høyer, H. C. (2016). External Control and Professional Trust in Norwegian School Governing. Synthesis from a Nordic Research Project. *Nordic Studies in Education*, Vol. 36(2), 86-102.
- Petterson, D., Popkewitz, T. & Lindblad, S. (2017). In the grey zone: large scale assessment-based activities betwixt and between policy, research and practice. *Nordic Journal of Studies in Educational Policy*, 3(1), 29-41.
- Sivesind, K., Bachmann, K., & Afsar, A. (2003). *Nordiske læreplaner (Nordic Curricula)*. Oslo: Læringscenteret.
- Sivesind, K., & Bachmann, K. (2008). Hva forandres med nye standarder? Krav og utfordringer med Kunnskapsløftets læreplaner. In G. Langfeldt, E. Elstad & S. Hopmann (Eds.), *Ansvarlighet i skolen. Politiske spørsmål og pedagogiske svar* (pp. 62-93). Fagernes: Cappelen Akademisk Forlag.
- Skedsmo, G. (2009). "School Governing in Transition. Perspectives, Purposes and Perceptions of Evaluation Policy." Doctoral thesis. Department of Teacher Education and School Development, Faculty of Education, University of Oslo.
- Skedsmo, G. (2011). Formulation and Realisation of Evaluation Policy: Inconcistencies and Problematic Issues. *Journal of Educational Assessment, Evaluation and Accountability* 23(1), 5-20.
- Skedsmo, G. & Møller, J. (2016): Governing by new performance expectations in Norwegian schools. In: Gunter, H., Hall, D., Serpieri, R. & Grimaldi, E. (Ed.): *New Public Management and the Reform of Education: European lessons for policy and practice*. London. Routledge, 53-65.
- Spillane, J. P. (2012). Data in Practice: Conceptualizing the Data-Based Decision-Making Phenomena. *American Journal of Education*, 118(2), 113-141.
- Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic Model in Education: Education as part of the political system in the last 50 years *Scandinavian Journal of Educational Research*, 51(3), 245-283.
- The Norwegian Directorate for Education and Training (2017a). Analyse av nasjonale prøver på 5. trinn. Retrived 1st of December, 2017, from <https://www.udir.no/tall-og-forskning/finn-forskning/tema/nasjonale-prover/resultater-fra-nasjonale-prover-pa-5.-trinn/>
- The Norwegian Directorate for Education and Training (2017b). Analyse av nasjonale prøver på 8. og 9. trinn. Retrived 1st of December, 2017, from <https://www.udir.no/tall-og-forskning/finn-forskning/tema/nasjonale-prover/resultater-fra-nasjonale-prover-pa-8.-og-9.-trinn/>
- The Norwegian Directorate for Education and Training (2012). Rettleiing og retningslinjer: Om nasjonale prøver (Framework for national tests). Retrieved 1st of November, 2016, from <http://www.udir.no/Vurdering/Nasjonale-prover/Rettleiing-og-retningslinjer-nasjonale-prover/1-Om-nasjonale-prover/>
- The Norwegian Directorate of Education and Training. (2008a). Analysis of results from the national tests. Retrieved February 21, 2009, from http://www.udir.no/upload/Nasjonale_prover/Analyse_av_resultater_fra_nasjonale_prover_2007.pdf
- The Norwegian Directorate of Education and Training. (2008b). Hva måler nasjonale prøver? (What do the national tests measure?). Retrieved 01.02.09, from http://www.utdanningsdirektoratet.no/templates/udir/TM_Artikkel.aspx?id=3546
- The Norwegian Directorate for Education and Training. (2005b). Nasjonalt kvalitetsvurderingssystem for grunnsopplæringen (A national system for quality assessment in comprehensive education). February 14, 2008, from

- http://udir.no/Artikler/_Vurdering/Nasjonalkvalitetsvurderingssystem-for-grunnopplaringen/
- The Education Act. Regulation (2006).
- Tveit, S. (2014) Educational assessment in Norway, *Assessment in Education: Principles, Policy & Practice*, 21(2), 221-237.
- Uljens, Michael; Møller, Jorunn; Årlestig, Helene & Frederiksen, Lars Frode (2013). The Professionalisation of Nordic School Leadership, In Lejf Moos (ed.), *Transnational Influences on Values and Practices in Nordic Educational Leadership. Is there a Nordic Model?* (p. 133-158). Dordrecht. Springer.
- Werler, T., & Sivesind, K. (2007). Norway. In Döbert, H., Hörner, W., von Kopp, B. & Mitter, W. (Eds.), *The Education System of Europe*. Dordrecht. The Netherlands: Kluwer.
- White Paper No. 30. (2003-2004). *Kultur for læring (Culture for learning)*: The Ministry of Education and Research.
- Åkerstrøm Andersen, N., & Thygesen, N. T. (2004). *Styring af styringsværktøjer*. *Nordisk Administrativ Tidsskrift* (1), 28-37.