Assessment for Learning in Primary School Teaching: The Cases of Norway and Portugal[1]

Guri A. Nortvedt[a], Leonor Santos[b], and Jorge Pinto[c]

[a] Department of Teacher Education and School Research, University of Oslo

[b] Institute of Education, University of Lisbon

[c] School of Education, Polytechnic Institute of Setúbal

In this study, we aim to understand the driving forces underlying assessment for learning (AfL) in primary school teaching. By applying a case study design, including the cases of Norway and Portugal, and using mathematics teaching as an example, available policy documents and research reports are analysed to identify the differences and similarities that might explain the assessment practices previously observed in the two countries. Many similarities are found at the school and national levels. In particular, AfL is introduced as a national policy in both countries. Still, AfL practices are not common in primary mathematics classrooms in either country, although this is true for different reasons in each country. It is suggested that the assessment culture caused by national policies, such as curriculum reforms, national professional development projects, and teacher autonomy, explain the similarities in the observed outcomes.

Keywords: Primary school, mathematics education, classroom assessment practices, assessment for learning, national assessment policies

[1] Nortvedt, G., Santos, L., & Pinto, J. (2016). Assessment for learning in primary school mathematics teaching: The case of Norway and Portugal. *Assessment in Education: Principles, Policy & Practice*, 23(3), 377-395. (DOI: 10.1080/0969594X.2015.1108900)

1. Introduction

Several factors may influence student learning. One main factor is the assessment practices that students experience. For instance, external national assessment may have negative implications, such as reducing the curriculum to what is easiest to measure: knowledge that can be readily and reliably marked (Harlen, 2005). A similar risk is connected to the summative assessment practices that teachers develop and apply in the classrooms in which they teach (Boaler, 1997; Stobart, 2008); it is, for instance, easier to assess students' work on traditional tasks than their creative problem solving. In the last decades, special attention has been given to the kind of assessment implemented to aid student learning, which is called assessment for learning (AfL) (Black, Harrison, Lee, Marshall, & Wiliam, 2003), and AfL has been a growing industry (Black, 2013). This development has two origins: 1) *research* aimed at understanding whether the practices of AfL improve student performance and 2) *practice* that attempts to understand whether teachers can transform ideas from research into productive practices.

Based on research, it is well known that systematic AfL practice helps students to learn, even when teachers have access to only limited instruments, such as national examinations (McDonald & Boud, 2003). However, teachers seemingly struggle to develop and use AfL practices. This phenomenon is not country or culture specific (Black & Wiliam, 1998), nor is it subject specific; it is much more general. Even more importantly, research has not managed to explain the underlying reasons for this generalised situation. More focused attention has to be paid to local differences to understand the assessment traditions of each country and "to prospect for the development of effective formative assessment" (Black & Wiliam, 2005, p. 251).

The main aim of this paper is to understand the driving forces underlying AfL in primary school teaching. The research question reads: What similarities and differences can

be found in terms of the driving forces of AfL in primary school teaching at various organisational levels? Specifically, we will examine the driving forces of AfL in Norway and Portugal. The two countries—Norway and Portugal—were chosen because they share some similarities but also some differences. Most importantly, AfL has been adopted as a national strategy in both countries.

To help clarify our analysis, we will use the example of AfL in mathematics teaching and learning. We assume the importance of studying assessment issues within the context of specific subject matter (Bennett, 2011) because this may provide more context for the interpretations of the observed patterns. We chose the school subject mathematics due to the relevance of this topic in early education. In addition, projects aimed at supporting students in learning mathematics, such as the American Cognitively Guided Instruction (Carpenter, Fennema, Franke, Levi & Empson, 1999) or the United Kingdom Shell Centre's Diagnostic Teaching (Bell, 1993), might be seen as forerunners for adopting AfL in mathematics education. A common factor, for Norway and Portugal, is the substantial improvement in Grade 4 students' averages in the Trends in Mathematics and Science Study 2011 (TIMSS 2011) (Mullis, Martin, Foy, & Arora, 2012). However, while Norway scored at the TIMSS average in 2011, Portugal scored significantly higher, raising the average national score by 90 scale points from the first test implementation in 1995.

2. Theoretical Framework

In the research literature, the term "formative assessment" was used for many years (Gardner, 2012), but around 2002, the term "assessment for learning" was introduced, and with the work of the Assessment Reform Group (ARG) in England, it became the dominant term. Black and Wiliam (1998), in their review "Assessment and Classroom Learning", used the term "formative assessment" to describe assessment that could be used to modify teaching and learning activities and thus support student learning. They further claimed that no agreed-

upon and clear definition existed for “formative assessment”. Through their continuing work, ARG outlined and clarified what assessment to steer, improve, and support student learning was, stating that it was “the process of seeking and interpreting evidence for use by learners and teachers, to identify where the learners are in their learning, where they need to go and how best to get there” (ARG, 2002, pp. 2–3).

AfL has the potential to strongly impact student learning, and it may be one of the most significant factors influencing growth in mathematical knowledge (Hattie, 2009; Wiliam, 2007). However, it is challenging for teachers to adopt and practise AfL in their mathematics teaching (Black & Wiliam, 2012). National policies and external assessment might constrain or enable AfL (Black & Wiliam, 2005), displaying, for instance, trust or lack of trust in teachers and beliefs about the relationship between teaching, learning, and assessment.

Often, the emphasis of assessment in mathematics classrooms is that it is a formal process (Wiliam, 2007). Many teachers struggle to provide students with feedback that will help them to understand what they know about mathematical concepts and procedures and where to proceed with their work. Should they, for instance, understand a new concept, learn heuristics for problem solving, or should they develop some skill or learn a new algorithm? If teachers do not focus their feedback, the mathematical point of view might be lost for the students (Boistrup, 2010). Many mathematical concepts are difficult to understand, and many students are likely to develop misconceptions, such as “multiplication makes bigger”. Likewise, when students work on solving addition problems, some strategies, such as “count all” strategies, are less effective than “count on” strategies (Geary & Hoard, 2003), and teachers need to be able to help their students reveal and recognise these matters through teaching–learning activities. In her research, Boistrup (2010) found that, for teacher feedback to students to be effective (i.e., offer opportunities for learning mathematics), the focus of the

assessment practices should be on mathematical processes rather than on the student (self) or the task. In addition, the openness of teacher questions was vital. Good intentions and providing feedback to students are not enough for assessment actions to be AfL (Stobart, 2008).

Continuing professional learning is a condition for teachers' adoption of AfL (Peddar & James, 2012). While Black and Wiliam (2012) argue that professional development might be implemented across school subjects, focusing on three processes and five key strategies only, we argue that professional development needs to be rooted in the subjects within which the teachers are to implement the AfL practices. Teachers not only need to change the focus of their assessment practices but also their beliefs about assessment and mathematics. If teachers, for instance, believe that learning mathematics is about memorising and applying rules, focusing strongly on assessing what is easily assessed, they will give very different feedback from teachers who have a more competence-oriented view of mathematics. In addition, according to Boistrup (2010), assessment practices that focus on guiding students through closed questions or accepting any mathematical solutions to problems as solutions of equal mathematical value, hinder learning. Consequently, we can assume that to engage effectively in AfL, teachers require deep and principled knowledge not only of mathematics but also didactical[2] knowledge regarding the mathematical topics they teach.

AfL can be introduced at various levels of an educational system. At the national level, AfL might be introduced in the form of a national policy for education in schools or a policy for teachers' professional development. AfL can also be implemented at the local level as part of teachers' assessment practices in their classrooms, part of a professional development course, or as a component of cooperation between teachers and researchers in a collaborative project to improve teaching and learning. However, for AfL to be effective,

[2] We use "didactics" and "didactical" as in the German tradition.

some alignment between assessment and curriculum is necessary. Such an alignment is a question of whether the current policy maintains key aspects of AfL, such as definition, progression, and procedures (Daughterly, Black, Ecclestone, James, & Newton, 2012). For instance, a strong focus on summative assessment or the national implementation of tests that have multiple purposes could hinder the adoption of AfL (Black & Wiliam, 2005; Newton, 2007; 2012). Just as alarming is that the extensive use of external tests could limit what is valued in teaching and learning (Harlen, 2012). This could also be a question of validity. Using the evidence collected for summative purposes for formative assessment can affect the validity of the assessment. In conclusion, the potential for tension between educational policies and classroom assessment strategies, such as AfL, exists.

3. Methodology

This paper aims to understand the driving forces underlying AfL in Norway and Portugal by using mathematics teaching as an illustrative example. A case study design (Yin, 2003), in which each country is one case, has been applied to first give a description of each case before discussing their similarities and differences in order to identify possible underlying driving forces.

The data were gathered through document analysis (Creswell, 2012). Documents might be seen as unobtrusive measures (Robson, 2002) in that they are non-reactive—they are not affected by our attention and analysis. However, the documents that have been used cover a wide range of authors, institutions, and purposes. Only a few of the data sources represent the researcher community and are peer-reviewed, published "high-quality research". Other data sources are policy documents, such as national white papers and curricula prepared by national ministries, or directorates for education. A third group of documents provides empirical evidence, reported by national agencies on the educational system. In other words, official documents permit us to describe national assessment policies, and the research

literature provides analyses of the assessment cultures. Table 1 lists the different kinds of text sources and gives some examples of the data sources for each category. As many of the documents are national reports in the national language, a traditional literature search with prototypical search words has not been possible. Rather, the included documents are what might be termed recognized and credible documents that are mainly from the national level. Another inclusion criterion was availability—all documents should be available to the public. The final criterion was that the documents should be policy documents or documents related to policy implementation, for instance, by addressing or assessing the uptake of a policy.

Table 1

Overview of Document Types and Status Including Examples Used for This Paper

Document type	Purpose of author	Examples
Legal document	Steering (i.e., educational acts, curriculum etc.)	Educational Act, 1998 MEC 2013 NMER & NDET, 2006 Decreto-Lei nº 137/2012 Despacho normativo nº 98-A/92
White paper	Policymaking	NMER, 2007 CNE, 2012
Official documents and web documents	Policy (teacher education programmes)	NDET, 2014b NMER, 2010 Abrantes, 2002
Research report group 1	Evaluate policy	Hagen & Nyen, 2009 Hopfenbeck et al., 2013 Santiago et al., 2012
Research report group 2	Monitoring educational systems, informing policy	Breiteig, 2013 Mullis et al., 2012 GEP, 2010
Research article	Research	Elstad, Nortvedt, & Turmo, 2009 Nordahl & Hausstätter, 2009 Santos et al., 2010 Barreira & Pinto, 2005 Fernandes, 2006

Although not triangulation by data or methods, this can be seen as triangulation by perspectives (Yin, 2003). Nevertheless, there are several restraints and issues regarding the transparency of the available information related to the two cases. Many available

documents—such as government reports, national analysis, and steering documents, such as white papers and webpages—are only available in the national language.

The data analyses were performed using content analysis (Robson, 2002) of the documentary sources by applying the following categories:

- Teacher education programmes
- Primary education curriculum
- Assessment systems
- Assessment practices at the classroom level

4. Results

4.1 The Educational Systems of Norway and Portugal

Although Norway is large in terms of area—385,186 km²,—the population is only 5 million. Each child cohort comprises approximately 60,000 children. Due to the population-to-area ratio, Norway has many small schools, thus raising the cost of schooling. On average, primary and lower secondary schools have 200 students in total, with 33% of schools having 100 students or fewer and another 40% of schools having between 100 and 299 students in total. Over the past 10 years, the number of students attending large schools has increased, and altogether, although there are still many small schools, only 8% of students attend a school that has fewer than 100 students in total (NDET, 2012). Norway spent a little more than 60 billion NOK (more than 7 million euros) on education in 2012 (Statistics Norway, 2014), spending approximately 5% of the yearly national budget on compulsory education.

Portugal comprises a mainland and two autonomous regions in the North Atlantic: the archipelagos of the Azores and Madeira. The total area is 92,090 km^2, with a population of 10,562,178 individuals in 2011 (INE, 2012). The area-to-population ratio in Portugal is very different from that of Norway, and starting in 2003, most schools were grouped into clusters[3]

[3] A school cluster is an organisational unit with its own administration and management organs, and it consists of the integration of preschool educational institutions and schools at different levels of the

that share a single school administration (financial and pedagogical) (Decreto-Lei nº 137/2012). Most clusters cover all levels of schooling, from primary to upper secondary, and a cluster can have as many as 2,500 students in total. All teachers answer to the same headteacher. However, they only have day-to-day responsibilities within their own school units. Portugal currently spends less than 4% of its national budget on education (INE-BP, DGO/MF, 2012).

Norwegian schools are inclusive, and students are not streamed until Grade 11. Compulsory education consists of seven years of primary school and three years of lower secondary school, followed by three or four years of upper secondary school (see Figure 1). All students attend primary and lower secondary education[4], and all students enter upper secondary education either in the academic (three years) or vocational programme (three or four years). However, dropout rates in upper secondary school are high, with only 70% of the students completing the programme within five years (NDET, 2012). Dropouts mainly occur in vocational education, and most are male students.

educational system. Its main propose is to ensure and enhance the quality of schools, as well as the proficient management of the human and material resources of schools.

[4] Parents can apply for a permit to homeschool their children. However, this is rare.

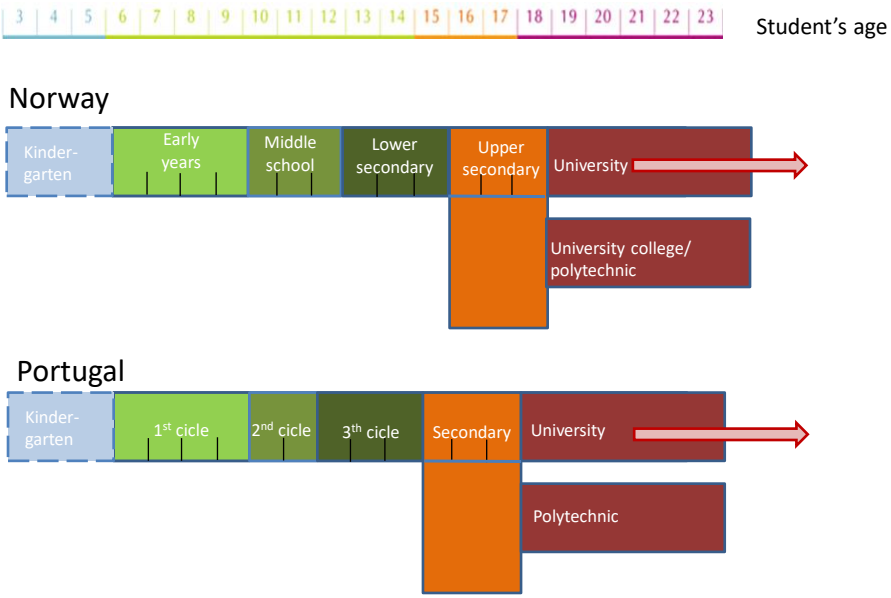

Figure 1. The structure of the educational systems in Norway and Portugal

Portuguese schools are also inclusive, with a single pathway for what is termed basic education (the first nine years, see Figure 1). The Portuguese educational system is structured in five levels: early childhood education, basic education, secondary education, post-secondary education, and higher education (see Figure 1). After basic education, a diversity of pathways are offered from secondary education onward, one track being more oriented towards further education and the other being vocational. Starting from a very precarious situation, in which Portugal showed a significant lag in relation to its European partners, Portugal has experienced very positive development in recent decades. In 2011, the enrolment ratio in the first cycle of basic education was 100%; 95% in the second cycle, 92% in the third cycle, and 73% in secondary education (CNE, 2012).

In both countries, students start school the year they turn six (see Figure 1). Norwegian primary school students attend their local or neighbouring schools. In the first years of schooling, mathematics is one of the few subjects taught. The class teacher usually teaches all

subjects. Schools are allowed to have specialist teachers, although few are mathematics teachers. On average, students receive four mathematics lessons of 60 minutes a week for 39 weeks a year throughout the first four years (approximately 156 hours of mathematics lessons per school year) (NDET, 2014a).

In Portugal, the teaching in the first cycle of basic education can be considered more global. However, without focusing on isolated subjects, the aim is still to develop basic skills in the subject areas, one of which is mathematics. The teaching part of the school day usually lasts for six hours. The students are taught by a class teacher, while expert teachers are used for certain areas, particularly physical–motor expressions and special needs education. On average, students receive at least seven hours of mathematics teaching a week for 35 weeks during each of the first four years (approximately 245 hours of mathematics lessons in one school year) (Despacho nº 8248/2013).

In summary, many similarities can be found between the structuring of the early education of Norwegian and Portuguese children. The first level of formal schooling is four years in both counties, and mathematics is one of a few subjects that students are taught by a class teacher. Students attend inclusive schools in systems offering a single pathway until the end of compulsory education.[5] Nevertheless, the different organisation of the teaching may represent differences in, for instance, the orientation towards global education or separate subjects. A strong emphasis on particular subjects may reflect an idea that some subjects are more important than others, while a global orientation gives quite different signals. The same applies to the number of hours devoted to each subject. To be specific, mathematics lessons occupy more hours in the Portuguese school year than in Norway, but at the same time, the teaching is more global. Consequently, different signals are given to teachers and students,

[5] For the students who entered the educational system in 2010 and later, compulsory education will be 12 years.

but in both countries, mathematics is identified as an important school subject and is given some priority.

4.2 Teacher Education Programmes

In Norway, primary school teachers are educated in a teacher or kindergarten teacher programme at a teacher training college, qualifying them to teach all subjects for Grades 1–7. All teacher education programmes are designed to train future teachers for the teaching profession and consequently include training in schools in addition to lectures and other learning activities at the college. In the current programme, students must complete 30 ECTS in a combined course, including both mathematical and didactical components (NMER, 2010). Prior to 2010, the teacher education programmes certified teachers for Grades 1–10, including a compulsory mathematics component that has varied over the years from consisting of a few lessons to 30 ECTS (Breiteig, 2013). As a consequence, teachers have very different formal qualifications for teaching mathematics, and The Norwegian government is concerned about this issue. Before 2017, the teacher education programmes will be revised once more, as teacher education will then become integrated master programmes (MER, 2014a). It is not yet known how this will affect student teacher preparation to teach subject matter such as mathematics.

Table 2

Overview of Current Teacher Education Programmes for Primary School Teachers in Norway and Portugal

Country	Implemented	Duration of full study programme (Semesters/years)	Compulsory mathematics component (mathematics + didactics) (ECTS)	Qualifies for teaching in grades
Norway	2010	8/4	30 (integrated)	1–7
Portugal	2007	9/4.5	30 + 6.25/7.5	1–4 + kindergarten

	8/4	30 + 3.75/5	1–4
	10/5	(30 + 5)+ 4.15	1–6

Table 2 provides an overview of the current teacher education programmes for primary school teachers in both countries. Prospective primary school teachers in Portugal complete six semesters (180 ECTS) at the bachelor's level before entering a master's programme lasting two, three, or four semesters. Depending on the length of the master's component, teachers are certified to teach from Grades 1–4 or 1–6. All programmes educating professional teachers include training components in which students visit schools to observe and/or practise teaching for at least 10 weeks per semester. While the mathematics courses that students complete as part of their training at the bachelor's level focus on mathematics only, the courses at the master's level are integrated courses containing mainly didactical[6] components. At the end of the bachelor's level, students must have completed at least 30 ECTS in mathematics (Decreto Lei 43/2007). Access to all master's programmes are conditional upon the attainment of sufficient mathematics credits at the bachelor's level.

In addition to the compulsory mathematics components in the teacher programmes, most Norwegian teacher training colleges offer optional courses, allowing their students to complete 60 ECTS in mathematics and didactics. However, the students entering the teacher education programmes have very low levels of prior knowledge of mathematics, and typically struggle with fractions, percentages, and so on (Nortvedt, Elvebakk, & Lindstrøm, 2010). Consequently, they might be expected to struggle with the mathematical content in the teacher training courses as well. According to Lagerstrøm (2007), only 10% of Norwegian primary school teachers had completed 60 ECTS or more in mathematics or mathematics education. In total, by 2011, 95% of the Norwegian teachers had completed a degree at the bachelor level or the equivalent (Mullis et al., 2012). Approximately 40% of the teachers who taught primary mathematics in 2006 took no mathematics courses during teacher training, and another 26%

[6] We use "didactic" as in the French and German traditions.

took fewer than 30 ECTS (Lagerstrøm, 2007). In 2014, the proportion of primary school teachers without any ECTS in mathematics is one of four (Lagerstrøm, Moafi, & Revold, 2014). Some of the primary teachers are kindergarten teachers by education, who, since 1997[7], have been allowed to teach the first four years of primary school after attending an additional year of teacher training that includes courses in mathematics education, reading instruction, and pedagogy (NMER, 2009). Few teachers have a major or specialisation in mathematics. According to Mullis et al. (2012), 30% of Grade 4 students in Norway are taught by teachers with such qualifications. Clearly, many of the Grade 1– 4 students in Norway are taught mathematics by a teacher who lacks the formal training needed to become a proficient mathematics teacher. Recently, awareness has been raised regarding the lack of formal training, and teachers might be offered a stipend to attend courses at universities and colleges (NDET, 2014d). Indeed, in 2015, 5,050 teachers will have the possibility to receive some support in the form of a stipend or other arrangements to reduce teaching hours (MER, 2014b).

In Portugal, by 2011, 94% of teachers teaching mathematics in Grade 4 had completed a degree at the bachelor's level or higher (Mullis et al., 2012). Most teachers, 71%, reported having a major in primary education but no major specialisation in mathematics. However, 25% hold a major in both primary education and mathematics (Mullis et al., 2012). Thus, while the percentages of teachers with a teaching qualification are comparable in Norway and Portugal, Portuguese teachers are better prepared, judging from the number of teachers holding a specialisation in primary education or mathematics education. This is reflected in the percentage of teachers reporting that they feel "very well" prepared to teach TIMSS mathematics topics: 92% of Portuguese students are taught by a confident teacher, compared to 78% of Norwegian students. Table 3 presents the percentages of students taught

[7] Compulsory school lasted nine years prior to 1997. In 1997, one year was added to primary school, extending basic education to 10 years. Consequently, many more teachers were needed, and kindergarten teachers were allowed to teach early primary school.

by a teacher who reported having participated in professional development in some aspect of mathematics education during the past two years.

Table 3

Attendance at Professional Development Courses as Reported in the Teacher Questionnaire for TIMSS 2011, Reported as the Percentage of Students Taught by a Teacher Who Has Participated in Professional Development in the Past Two Years (Based on Mullis et al., 2012).

	Didactics and pedagogy	Mathematics	Curriculum	Assessment	Technology
TIMMS average	46%	44%	41%	37%	33%
Norway	30%	25%	11%	16%	11%
Portugal	54%	58%	61%	25%	36%

The Portuguese *Mathematics Action Plan* included several national strategies that were implemented for a period of seven years. For instance, a new mathematics curriculum for basic education that was grounded in recent research in mathematics education was implemented as part of this policy (MEC, 2007). A training programme for mathematics teachers of students aged 6 to 11 was developed, which was the responsibility of the providers of pre-service teacher education (i.e., universities and schools of education) (Serrazina, 2013). An innovative national support programme was developed that covered around 95% of the basic education in public schools for the basic education level that were situated in the continental part of the country (Santos, 2008). At the school level, mathematics teachers were to establish their project work and thus enhance the mathematics learning of their students. The projects were scrutinised at monthly regional meetings. In the second stage of the mathematics plan, the implementation of the new curriculum was supported. Overall, taking into account the last results of TIMMS 2011 and PISA 2012 (Mullis et al., 2012; OECD, 2013), this educational policy seems to have had positive outcomes. Judging by the comments made in the PISA 2012 report (OECD, 2013), this is also recognised by the OECD.

Portugal is among the countries in which Grade 4 teachers have the highest participation in professional development courses (Mullis et al., 2012). All teachers in Portugal must participate in continuous professional training to pursue their teaching careers (Decreto-Lei, nº 22/2014). Approximately 50% of the courses are in didactics for the topics the teachers teach, which might explain the high numbers of teachers who report having participated in courses.

Judging by the responses to the teacher questionnaires in TIMSS, Norwegian teachers have had fewer opportunities for professional development (see Table 3). Only one in four students is taught by a teacher who has participated in professional development courses in mathematics. Prior studies have revealed that Norwegian teachers most likely have many opportunities to participate in more general professional development courses (Hagen & Nyen, 2009). For instance, many Norwegian teachers have had the opportunity to participate in AfL courses (NDET, 2014b). However, these are general courses that are not tied to the assessment of mathematical knowledge.

In summary, in both systems, student teachers are exposed to teaching in schools in addition to the education delivered on campus. The duration of the teacher education programme is a major difference; although the short master's programme in Portugal is only one semester longer than the Norwegian programme, Portuguese student teachers spend two to four semesters on more in-depth work when completing a master's programme. When Norwegian student teachers opt to take additional mathematics courses, these courses are also available at the bachelor's level. However, this is likely to change in 2017. Another striking difference regards the opportunities to attend professional development courses in mathematics education, with there being many more possibilities in Portugal than in Norway.

4.3 Primary Education Curriculum

Both countries have a national curriculum, including a mathematics curriculum, which was developed following a top-down logic. In Norway, the national curriculum consists of three parts. First, a general section retained from the 1997 curriculum provides an elaborate description of the students as ready and eager to learn and states that the aim of education is to teach the whole person, educate for lifelong learning, and help children to meet their full potential (MER, 1996). Next, the curriculum contains "the bridge", which is a set of principles for education. Finally, it contains the curriculum for the subjects. With the introduction of a new curriculum, called "Knowledge Promotion", in 2006, a striking development was observed. With this new curriculum, Norway moved from learning goals to achievement goals (NMER & NDET, 2006). In addition, five basic skills were defined and integrated into all subjects (reading, writing, oral skills, numeracy, and digital skills). Achievement goals were stated for Grades 2, 4, 7, 10, 11, 12, and 13 (NDET, 2014a). A few sentences were provided regarding summative assessment. However, the choice of teaching methods and approaches are left to the individual teacher, allowing for some autonomy.

A similar movement towards achievement goals can be observed in Portugal. With regard to mathematics, with the introduction of the Mathematics Action Plan in 2007, the focus was directed towards the processes of learning mathematics, such as number sense, problem solving, communicating mathematically, and mathematical thinking (MEC, 2007). Much autonomy was given to the teachers, allowing them to decide how to organise the teaching of each topic to enable the students to reach the goals stated in the curriculum. With the introduction of the latest curriculum in 2013, the focus was shifted from learning goals to achievement goals, as in Norway. However, one major difference is the detailing and breakdown of the goals students should reach. In the section related to learning outcomes in mathematics, the Portuguese mathematics curriculum published in 2013 (MEC, 2013) presents around 190 objectives and 900 descriptors that students should reach during the first

nine years of schooling. Achievement goals are stated for each grade level (Grades 1, 2, 3, etc.), massively reducing the autonomy of teachers. As in the Norwegian curriculum, very little is stated about assessment. Rather, teachers are referred to the educational acts.

In summary, in both countries in recent years, we can observe a tendency to develop standardised learning outcomes or descriptors in the form of competency goals, which might serve as national standards albeit without defining levels of proficiency. Norwegian teachers have more autonomy than Portuguese, as goals are not stated for each school year, allowing the teacher him/herself to break down curriculum goals into annual goals. Moreover, with respect to teaching methods, Norwegian teachers have more autonomy. Although in Portugal, in the first decade of the twenty-first century, we may say that there was an effort to reduce the characteristics of a curriculum regulation situation (Kuiper, Nieveen, & Berkvens, 2013), new outcomes and examinations were recently introduced, thereby reinforcing it again.

4.4 Assessment Systems

In Norway, the latest regulations explaining the Education Act[8] of 1998 identify four principles influencing assessment practices. These indicate that students should: 1) understand what to learn and what is expected of them (§ 3-1), 2) obtain feedback that provides information on the quality of their work or performance (§ 3-11), 3) be given advice on how to improve (§ 3-11), and 4) be involved in assessing their own learning process (§ 3-12) (Lovdata, 2006). Students in Grades 1–7 should be assessed without using grades or marks.

The same orientation in assessment policy can be found in Portugal. Since 1992, a high level of importance has been assigned to AfL. For instance, AfL has been assumed as the main assessment modality for students in basic education (Despacho Normativo, 98-A/92). Several assessment principles have been established since 2001, such as *coherence* between the objectives of learning and assessment practices, the *primacy of formative assessment,* the

[8] Opplæringsloven (1998).

relativity of the assessment decisions (the need to consider the evolution of each student), the *transparency* of the assessment processes, and the *diversification* of processes (Abrantes, 2002). This policy orientation has remained in place for more than two decades.

In both countries, at the end of each term, students are assessed summatively. As Norwegian primary school students do not receive any grades or marks, their teachers produce and send home a written evaluation to be presented to the students' parents. Assessment practices have traditionally been weak, with teachers focusing on effort rather than the quality of students' competence and curriculum mastery (Hopfenbeck, Tolo, Florez, & El Masri, 2013). Teachers were not good at making their expectations clear to the students (Klette, 2003). In addition, when students experience difficulties learning to read or in mathematics, teachers have tended to "wait and see", stating that the children will most likely overcome their difficulties by themselves (Nordahl & Hausstätter, 2009; van Daal, Solheim, Gabrielsen, & Begnum, 2007). However, with the introduction of Whitepaper 16,2006–2007 (NMER, 2007), early intervention became a national policy, and in 2007, national mapping tests in reading and numeracy were implemented to help primary school teachers identify the weakest 20% of students. No other external assessments for Grades 1–4 currently exist in Norway. When national tests were first introduced in 2004, Grade 4 students were tested in numeracy, Norwegian, and English reading in the second half of the last semester, but this was only done for the purposes of monitoring the school system. Tests were low stakes both for the schools and students, without any negative consequences for low results (Elstad, Nortvedt, & Turmo, 2009). After criticism regarding test frameworks, national tests were stopped for one year. When they were reintroduced in 2007, the Norwegian government decided to make the tests entrance tests for Grade 5 (NDET, 2008), consequently removing all external testing for the lower primary level.

Portuguese students are graded by their teachers. There is a particular emphasis on grading students at the end of Grade 4, when students are graded on a scale from 1 to 5, with 3, 4, and 5 being passing grades. In earlier grades, student assessment is translated into a qualitative description of the difficulties and learning achievements of the students. Traditionally, Portugal has had a large number of national exams, but in 1968, they were removed. Since the end of the 1980s, students take any external examinations prior to entering university. More recently, in 2000, national tests in mathematics and the Portuguese language were re-introduced at the end of the 4th, 6th, and 9th grades, but these were only intended to evaluate the educational system and did not have implications for the individual student. However, in 2005, these tests were transformed into formal exams for the 9th grade. Following that, in 2013, the same choices were made for Grades 4 and 6. This change, although made in the name of objectivity and accountability, can be seen as reinforcing a process of social exclusion in Portugal, as the transformation of national tests into formal exams has high-stake consequences for the individual student.

Another contrast indicating very different national educational policies concerns transition. In Norway, students are automatically transferred to the next grade level each autumn. Students can, on rare occasions, skip a school year, but they cannot redo one. In contrast to the Norwegian situation, Portuguese students require a passing grade from their teachers to be transferred from one grade level to the next, starting with the transition from Grade 2 to 3, although more emphasis is placed on the transition from Grade 4 to 5 because this is the transition from the first to the second cycle. In fact, during the 2010/2011 school year, 3.3% of the Portuguese students failed a teacher assessment and had to retake a school year at some point during the first cycle (DGEEC, 2012).

To assist in the implementation of the guidelines outlined in national policies and laws, each country has developed several national strategies. In Norway, two major

assessment policies have been implemented at the primary level—the Quality Assessment System (QAS)[9] (NMER, 2002) and Assessment for Learning[10] (NDET, 2014b). In Portugal, several national educational programmes have been implemented concerning mathematics, science, and Portuguese. The most complete programme focusing on mathematical learning was the educational programme Mathematics Action Plan.[11]

The QAS was introduced after the OECD stated that Norway lacked a culture of assessing young students. The results from TIMSS 1995 and PISA 2000 clearly demonstrated that Norwegian students scored at a much lower level than expected (see, for instance, Bergersen, 2005). The QAS identified key tests at various educational levels, the purpose of each assessment, and for what and whom the results were intended (Elstad, Nortvedt, & Turmo, 2009). Primary education Grades 1–4 are involved on two levels. The results from international comparative tests, such as TIMSS and PIRLS, are to be used to monitor the quality of education and inform educational policy by providing information about the goals reached at the end of the lower primary level. In addition, numeracy and reading mapping tests are to be used to inform the classroom and school levels. Following the implementation of the QAS and a white paper on early intervention (NMER, 2007), the first generation of mapping tests was introduced in 2008, when a compulsory numeracy test for Grade 2 students was implemented. The purpose of the assessment was to identify the weakest 20% of students (Throndsen & Alseth, 2013). Following this, optional tests for Grades 3 and 1 were introduced in 2009 and 2011, respectively. In 2014, the second generation of mapping tests was launched (NDET, 2014c). The main development is that the new tests are even more targeted towards the weakest students and have a strong ceiling effect by design (NDET, 2014c). Consequently, the tests will not provide any information about the majority of the

[9] Kvalitetsvurderingssystemet
[10] Vurdering for læring
[11] Plano de Ação da Matemática

students. The tests should be used for AfL and to plan interventions for weak students, although this will mainly be accomplished within the frame of inclusive classrooms and general education (NDET, 2014c).

AfL, as a national strategy in Norway, was first implemented in a national project called Better Assessment Practices[12] (Hopfenbeck et al., 2013). The aim of the project was to develop distinct national assessment criteria that teachers could use to assess to what extent students attained curriculum achievement goals. Participating schools tested different sets of criteria in, for instance, mathematics. The project evaluation revealed that the suggested criteria frameworks were too complicated to be useful for teachers, and many of the documents produced within the project revealed misconceptions among teachers regarding assessment criteria (Hopfenbeck, Throndsen, Lie, & Dale, 2012). The evaluating team suggested that new projects should include professional development aspects, such as collaboration between assessment specialists, researchers, and teachers. Following this, although it was considered a national policy, AfL was launched as an optional project involving several cycles, with more municipalities and schools entering the project during each cycle (Hopfenbeck et al., 2013), taking into consideration the recommendations from the evaluation of Better Assessment Practices. For the first implementation, strategies such as seminars with presentations from renowned assessment specialists from England (e.g., Gordon Stobart), learning networks within the municipality, and a webpage with resources for teachers and school leaders were used (NDET, 2014b).

In summary, we conclude that, in both countries, a national assessment policy that considers the importance of AfL exists. This recognition has implications for the further development of national strategies to help schools and teachers improve teaching and learning—particularly the latter. One of the ways to increase the efficiency of these strategies

[12] Bedre vurderingspraksis

is to understand assessment practices that are implemented at the classroom level. The main difference lies within student assessment practices and policies, with the Norwegian educational system focusing on less formal and "soft" procedures and having no external measures and the Portuguese system focusing on formal, summative procedures and using external measures to evaluate both the system and the students. This has consequences not only for the national accountability systems but also for students because they risk either not being provided with crucial information about their progress (Norway) or about their failing (Portugal). Both practices affect weak students and may focus on aspects of education other than AfL. Moreover, in Portugal, the tradition in external assessment and in a transition system is maintained even with the recent recognition of the importance of AfL. It is probable that teachers experience a great deal of tension regarding the necessity to prepare their students for examinations and that this makes them maximise the use of instruments that are more focused on external assessment and more likely to enhance test scores. This would leave very little time for AfL practices.

4.5 Assessment Practices at the Classroom Level

AfL, as an assessment policy, is implemented at the local classroom level by the teachers. In doing this, the teachers must translate national policies into local practices in which they engage with their students. This can be very challenging, as is often demonstrated in the research literature on AfL (see, for instance, Torrance & Pryor, 2001; Suurtamm, Koch, & Arden, 2010).

Very little research on Norwegian primary school mathematics teachers' classroom assessment practices exists. However, the OECD has performed extensive reviews of assessment in education in Norway, both of AfL (Hopfenbeck et al., 2013) and of assessment in general (Nushe, Earl, Maxwell, & Shewbridge, 2011). Hopfenbeck et al. (2013) evaluated the implementation of the national AfL programme. Their evaluation demonstrated that many

of the participating schools struggled to embed AfL in their everyday practice. One possible explanation is that some schools and school leaders lack support at the local (i.e., the municipality) level.

The OECD country note on equity in education in 2004 (Mortimer, Field, & Pont, 2004) can be interpreted as suggesting that Norwegian teachers lack an assessment culture. This may still be the case. Hopfenbeck et al. (2012) found that when asked to develop criteria for assessing whether students had reached curricula goals, Norwegian mathematics teachers often divided the goals into smaller units, making "very specific and fine-tuned competence descriptors" (p. 427). By adopting practices such as these, teachers are at risk of developing criteria such that they only assess and provide feedback regarding skills that are easily observed and measured rather than focusing on the more global competence described in the curriculum goals. This concern was also raised by Nushe et al. (2011). In addition, prior research has revealed that Norwegian primary school teachers tend to praise effort rather than the quality of student learning or knowledge (see, for instance, Mortimer, Field, & Pont, 2004). This is still evident from the practices that Hopfenbeck et al. (2012) observed. They suggest that Norwegian teachers do not understand how to develop and use assessment criteria. However, according to Hopfenbeck et al. (2012), more students in primary schools reported that their teachers told them what and how to improve their learning as compared to older students, who received grades. If this feedback is based on the unstructured observation of mathematics teaching and learning activities or breaking curriculum goals into very specific mathematical skills that can be easily assessed (for instance, learning the multiplication tables by heart), this practice cannot be termed assessment for mathematical learning.

In 2012, the Portuguese evaluation and assessment systems were reviewed by the OECD (Santiago, Donaldson, Looney, & Nusche, 2012). It was found that although teachers

and schools have a large degree of autonomy in terms of student assessment, there is a lack of AfL in teachers' practices. This prevents a culture of AfL:

> There is little emphasis in assessment practices on providing student feedback and developing teacher–student interactions about student learning. In classrooms and schools, the formative seems to be increasingly displaced by the summative and a focus on the generation of summative scores. (Santiago et al., 2012, p. 4)

In terms of what one may read in most of the schools' documents, diagnostic, formative, and summative assessment are all included in teachers' practices (Looney, 2011). However, the way in which these intentions are operationalised in practice provides support to the interpretation of Portuguese assessment practices presented in the OECD evaluation. There is no clear relationship between what teachers think, what they do, and what they wish to do. Formative assessment has been observed to be a residual practice, despite the conceptual adherence of many teachers (Fernandes, 2006). In particular, primary school teachers emphasise tests less than their colleagues who teach subsequent years. Teachers often use observation to collect information about their students, but this is typically unstructured observation (Barreira & Pinto, 2005).

Teachers have previously been seen to change their assessment practices. However, this may occur only in a few very particular situations, such as in teachers' training programmes and collaborative work with researchers (Peddar & James, 2013; Santos et al., 2010). Further, these teachers tend to articulate the formative with summative assessment practices (Pinto & Santos, 2013). The nature of the difficulties that teachers identify when seeking to change their assessment practices allows us to state that a large investment in teacher training in AfL is necessary.

What is common in the assessment practices displayed by Norwegian and Portuguese teachers? In both countries, practices that embed AfL are included in the steering documents that guide the practices of teachers. Still, teachers in both countries apparently find it difficult to adopt AfL because they strive to use criteria for judging the quality of students' mathematical learning. Although Norwegian teachers have somewhat more autonomy, similarities are found regarding the autonomy of teachers and local schools to interpret curricular goals and form their own practices within the frames provided by national policies and laws.

5. Discussion

The comparative analysis between Norway and Portugal presented here reveals that there are several similarities at the national and school levels. For instance, at the national level, the structure of the first stage of the educational system is very similar (only one path through compulsory education, the number of grades, inclusive schools, class teachers, etc.). At the school level, AfL practices are most likely not common in primary classrooms in either country. The national research developed in each country reveals that under the usual circumstances in primary classrooms, what might truly be termed AfL is scarce. Although this is a similarity between the two countries, the underlying factors that might help explain this result seem to be very different.

If we consider the national assessment cultures, we may find significant differences and thus implications. At the school level in Norway, teachers have lacked an assessment culture. For instance, no standards or benchmarks describing student competence have existed (Mortimer et al., 2004). This might explain why teachers have difficulties developing and using assessment criteria, whatever the importance that is given to such criteria whether formative or summative. Contrastingly, teachers in Portugal have an expressed assessment culture but it is a predominantly summative one (Santiago et al., 2012). At the national level,

while Norwegian students do not receive any grades or marks, students in Portugal are graded. Beginning in 2013, they will even take a formal exam as part of the approval system. We might even claim that the two countries are at two ends of a continuum. Norway, at one end, wants to focus strongly on AfL, but probably without success due to a lack of standards and clear criteria. In contrast, the strong focus on summative assessment in the approval system places Portugal close to the other end of the continuum. To adopt the principles of AfL—that is, to develop criteria that allow for the assessment of where students are in relation to curriculum goals (Wiliam, 2007)—teachers must be able to draw on both formative and summative assessment practices while recognising the relevance of both. However, in both countries, teachers do not practise assessment that can be termed AfL, and this is likely because they do not know how to do it or they might not value it. In conclusion, different factors lead to the same situation in both countries.

In both countries, teachers are seen to change their assessment practices after taking part in professional development with an assessment focus. However, how should such training be scaled up to the national level? In Norway, the AfL project attempts to do this. However, large differences are observed between schools (Hopfenbeck et al., 2013). One important factor is support—headteachers' support for teachers, as well as the school administration's support for headteachers in the local municipality. It can be concluded that changing teachers' practices related to the assessment for mathematical learning should build on national professional development projects that include support at both the national and local levels, enabling teachers to formulate clear criteria and recognise different levels of mathematical competence in their students. This necessitates forming national policies that are also rooted at the local level, as was experienced in Portugal (Santos, 2008). Moreover, a focus on specific aspects of teaching and learning that are related to mathematics seems to be a fruitful option (see, for instance, Black & Wiliam, 2003). Although general education

factors are important in supporting the development of AfL practices, teachers require principled mathematical and didactical knowledge to interpret students' misconceptions and errors, provide them with adequate feedback, and possibly even develop this capacity in students (Sadler, 2010).

The existence or non-existence of teacher autonomy may be a third reason why AfL practices are so rare. In Norway, primary school teachers in Grades 1–4 seem to have a high level of autonomy. There is no pressure from national test standards. In Portugal, the level of autonomy tends to be much smaller due to the appearance of a national examination. Nevertheless, in both countries, the mathematics curriculum pays special attention to learning outcomes, although this is more significant in Portugal. This type of curriculum implies a risk of overvaluing what is easily observed and measured, necessarily reducing teaching practices to only a small component of the mathematics curriculum. This will favor a few learning aspects, impoverish the mathematics tasks, and consequently reduce the potential of AfL practices (NCTM, 2000; Shepard, 2001).

We must take into account that AfL practice in mathematics in primary school education is complex. First, this complexity is reflected by the diversity of variables that may influence and explain whether or not such a practice truly exists. Thus, all attempts to study this phenomenon reduce the richness of reality. Second, the educational situation in the two countries under analysis has seen several recent and ongoing changes, the implications of which are impossible to predict. Third, the methodological approach used in this study has some limits. The data were collected exclusively via document analysis. Certainly, by using a greater diversity of data collection techniques, other dimensions could have enlightened the understanding of what is in focus, permitting us to achieve deeper knowledge, and new issues might have emerged. Therefore, we can say that by bringing new elements to the forefront, this study contributes to a better understanding of the area of interest. It might benefit from

being continued by others seeking to answer the questions that this study leaves unanswered because of its limitations.

6. Final Remarks

This study confirms what other studies have shown over the past decades: At the classroom level, adopting AfL practices is neither easy nor common (Black & Wiliam, 1998; Torrance & Prior, 2001; Wiliam, 2007). However, the results obtained in this study allow us to go further. In one sentence, we can say that similar situations often require different explanations, or, in the words of Wiliam, because different countries represent different systems, "there is no receipt that will work for everyone" (Wiliam, 2007, p. 1091).

The relationship between educational policies and the reality of the classroom emerged as non-linear in nature. In fact, it is multidimensional. If we consider educational policies for classroom practices, we may say that in both countries, there was clearly an investment made at the policy level. However, other variables, in combination with national policies, emerged as important in terms of policy effectiveness. The importance of the specificity of the subject matter for the uptake of AfL in professional development settings, as opposed to a more general approach to educational issues, is highlighted.

Regarding the implementation of educational policy at the classroom level, it could be claimed that collaborative work in heterogeneous teams promotes changes in assessment practices and improves mathematics learning in primary school students. We suggest that, at the national level, although AfL is a national policy, our findings indicate that these two nations have yet to recognise the indisputable impact of a national assessment culture, which is sometimes viewed as a significant obstacle to innovation regarding assessment practices. If the curriculum goals are difficult to interpret for the teachers, for instance, because they lack the necessary mathematical and didactical knowledge, this may also present an obstacle.

At the intermediate level, we want to highlight the relationship between curriculum and assessment. Curriculum determines the autonomy level of teachers, which prescribes which dimensions of learning should be assessed. However, if the idea of assessment is closely related to certain concepts—such as measurement, rigor, and objectivity, combined with a curriculum that is strongly related to learning outcomes—this may produce very restricted ideas about what it is to learn and know mathematics and, consequently, what to assess. The same outcome could also result from an assessment culture in which standards or descriptions of key (complex) competences, such as mathematical competence, are not given, because directions for further learning might then be conceived of as unfocused or even meaningless.

References

Abrantes, P. (2002). Introdução [Introduction]. In P. Abrantes & F. Araújo (Eds.), *Reorganização curricular do ensino básico* [*Curricular reorganization in basic education*] (pp. 7–15). Lisboa: Ministério de Educação, Departamento da Educação Básica [Minister of Education, Department of Basic Education].

ARG (2002). *Assessment for learning: 10 principles.* Assessment Reform Group. Retrieved from http://www.assessment-reform-group.org/.

Barreira, C., & Pinto, J. (2005). A investigação em Portugal sobre a avaliação das aprendizagens dos alunos (1990–2005) [Research about assessment of students in Portugal]. *Investigar em Educação, 4*, 21–105.

Bell, A. (1993). Principles for the design of teaching. *Educational studies in mathematics,* 24 (1), 5-34.

Bennett, R. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25.

Bergersen, H. O. (2005). *Kampen om kunnskapsskolen [The battle for the knowledge-based school]*. Oslo: Universitetsforlaget.

Black, P. (2013). Formative and summative aspects of assessment: Theoretical and research foundations in the context of pedagogy. In J. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 167–178). London: SAGE Publications.

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

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74.

Black, P. & Wiliam, D. (2005). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices. *The Curriculum Journal, 16*(2), 249 – 261.

Black, P., & Wiliam, D. (2012). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning: Second edition* (pp. 206–229). London: Sage.

Boaler, J. (1997). *Experiencing school mathematics: Teaching styles, sex and setting.* Buckingham; Philadelphia: Open University Press.

Boistrup, L. B. (2010). *Assessment discourses in mathematics classrooms: A multimodal semiotic study.* Stockholm: Universitetsservice.

Breiteig, T. (2013). Preparing teachers of mathematics in Norway. In J. Schwille, L. Ingvarson & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopedia. A guide to teacher education context, structure, and quality assurance in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)* (pp. 131–148): IEA.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively guided instruction.* Portsmouth, NH: NCTM.

CNE (Conselho Nacional de Educação). (2012). *Estado da educação 2012. Autonomias e descentralização*. Lisboa: Editorial do Ministério de Educação e Ciência.

Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston: Pearson.

Daughterly, R., Black, P., Ecclestone, K., James, M. & Newton, P. E. (2012). Alternative perspectives on learning outcomes: Challenges for assessment. In J. Gardner (Ed.), *Assessment and learning: Second edition* (pp. 72–87). London: Sage.

Decreto-Lei nº 137/2012 de 2 de julho de 2012. Agrupamentos de escolas (Grouping schools).

Decreto-Lei nº 22/2014 de 11 de fevereiro de 2014. Regime jurídico da avaliação contínua [Legal regime of in-service teacher education].

Despacho nº 8248/2013 de 25 de junho de 2013. Calendário escolar para 2013/2014 (Scholar calendar of 2013/2014).

Despacho normativo nº 98-A/92 de 20 de junho de 1992. Avaliação das aprendizagens dos alunos do ensino básico [Assessment of students in basic education].

DGEEC (Direção Geral de Estatísticas da Educação e Ciência – General Direcotorate for Education and Science Statistics). (2012). *Educação e números–Portugal 2012 [Education and numbers–Portugal 2012]*. Lisboa: Ministério de Educação [Ministry of Education].

Opplæringsloven [Educational act] (1998) https://lovdata.no/dokument/NL/lov/1998-07-17-61.

Elstad, E., Nortvedt, G. A., & Turmo, A. (2009). The Norwegian assessment system: An accountability perspective. *Cadmo, 17*(2), 89–103.

Fernandes, D. (2006). Vinte anos de avaliação das aprendizagens: Uma síntese interpretativa de artigos publicados em Portugal [Twenty years of assessment of learning: An interpretive summary of articles published in Portugal]. *Revista Portuguesa de Pedagogia [Portuguese Journal of Pedagogy], 40*(3), 289–348.

Gardner, J. (2012). Assessment for learning: Introduction. In J. Gardner (Ed.), *Assessment and learning: Second edition* (pp. 1–9). London: Sage.

Geary, D. C., & Hoard, M. C. (2003). Learning disabilities in basic mathematics: Difficulties in memory and cognition. In J. M. Royer (Ed.), *Mathematical Cognition* (pp. 93–116). USA: Information Age Publisher Inc.

GEP (Gabinete de Estudos e Planeamento–Office of Research and Planning). (2010). Educação e números–Portugal 2010*[Education and numbers–Portugal 2010]*. Lisboa: GEP, Ministério de Educação [Ministry of Education].

Hagen, A., & Nyen, T. (2009). Kompetanse – for hvem? Sluttrapport for evalueringen av "Kompetanse for utvikling. Strategi for kompetanseutvikling i grunnopplæringen 2005–2008." [Competence – for whom? Final report from the evaluation of "Competence for development. Strategy for developing competence in basic education 2005–2008"]. Fafo-Report 2009:21. Oslo: Fafo.

Harlen, W. (2005) Teachers' summative practices and assessment for learning: Tensions and synergies. *The Curriculum Journal*, *16*(2), 207–223.

Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning: Second edition* (pp. 171–185). London: Sage.

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Hopfenbeck, T. H., Throndsen, I. S., Lie, S., & Dale, E. L. (2012). Assessment with distinctly defined criteria: A research study of a national project. *Policy Futures in Education, 10*(4), 421–433.

Hopfenbeck, T. H., Tolo, A., Florez, T., & El Mastri, Y. (2013). Balancing trust and accountability? The assessment for learning programme in Norway. A governing complex educational systems case study: OECD.

INE (Instituto Nacional de Estatística–National Institute of Statistics) (2012). *Censos 2011 [Census 2011]*. Retrieved from http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos2011_apresentacao

INE - BP, DGO/MF (2012). Despesas do estado em educação [State expenditure on education]. PORDATA. Retrieved from http://www.pordata.pt/Portugal/Despesas+do+Estado+em+educacao+execucao+orcamental+em+percentagem+do+PIB-867

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]* , Oslo: Unipub, pp. 39-76.

Kuiper, W., Nieveen, N., & Berkvens, J. (2013). Curriculum regulation and freedom in the Netherlands. In W. Kuiper & J. Berkvens (Eds.), *Balancing curriculum regulation and freedom across Europe* (pp. 139–162). CIDREE Yearbook 2013. Enschede, the Netherlands: SLO.

Lagerstrøm, B. O. (2007). Kompetanse i grunnskolen. Hovedresultater 2005/2006 [Competence in compulsory education. Main findings in 2005/2006]. SBB 2007/21. Oslo: Statistics Norway.

Lagerstrøm, B. O., Moafi, H., & Revold, M. K. (2014). *Kompetanseprofil i grunnskolen. Hovedresultater 2013/2014.* Oslo: Statistics Norway.

Looney, J. (2011). *Integrating formative and summative assessment: Progress toward a seamless system?* OECD Education Working Paper nº 58. Retrieved from http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/wkp(2011)4&doclanguage=en

Lovdata. (2006). *Regulations to the educational act.* Oslo: Lovdata.

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education, 10*(2), 209–220.

MEC (Ministério de Educação e Ciência–Ministry of Education and Science). (2007). *Programa de matemática do ensino básico [Mathematics program of basic education*]. Retrieved from http://www.dgidc.min-edu.pt/ensinobasico/index.php?s=directorio &pid=71

MEC (Ministério de Educação e Ciência–Ministry of Education and Science). (2013). *Programa e metas curriculares. Matemática. Ensino básico [Program and curriculum outcomes. Mathematics. Basic education*]. Retrieved from http://www.dgidc.minedu.pt/ensinobasico/ index.php?s=directorio&pid=29

MER (Ministry of Education and Research). (1996). *Læreplanverket for den 10-årige grunnskolen [Curriculum for the 10-year compulsory school in Norway]*. Oslo: Nasjonalt læremiddelsenter.

MER. (2014a). Press release. Innfører femåring lærerutdanning på masternivå [Teacher education becomes a 5-year-master programme]. Retrieved from http://www.regjeringen.no/nb/dep/kd/pressesenter/pressemeldinger/2014/Innforer-5-arig-grunnskolelarerutdanning-pa-masterniva.html?id=761439

MER. (2014b). En milliard til videreutdanning. Retrieved from http://www.regjeringen.no/nb/dep/kd/pressesenter/pressemeldinger/2014/En-milliard-til-videreutdanning-av-larere.html?id=770876

Mortimer, P., Field, S., & Pont, B. (2004). *Equity in education. Thematics review. Norway. Country note*. OECD: OECD.

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Boston and Amsterdam: TIMSS & PIRLS International Study Center,

Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).

NCTM. (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

NDET. (Norwegian Directorate for Education and Training). (2008). *The education mirror 2007*. Oslo: NDET.

NDET. (2012). *The education mirror 2012*. Oslo: NDET.

NDET. (2013). Bedre vurderingspraksis. Retrieved from http://www.udir.no/Vurdering-for-laring/Litteraturtips/Prosjekt-Bedre-vurderingspraksis/Prosjekt-Bedre-vurderingspraksis/

NDET. (2014a). Læreplan i matematikk fellesfag [Mathematics curriculum]. Retrieved from http://www.udir.no/Lareplaner/Finn-lareplan/#matematikk

NDET. (2014b). Assessment for læring [Assessment for learning]. Retrieved from http://www.udir.no/Vurdering-for-laring/VFL-skoler/

NDET. (2014c). Kartlegging grunnskole [mapping tests for compulsory education]. Retrieved from http://www.udir.no/Vurdering/Kartlegging-gs/

NDET. (2014d). Stipendordningen. Retrieved from http://www.udir.no/Utvikling/Etter-og-videreutdanning/Etter--og-videreutdanning/Retningslinjer-for-stipendordning-for-larere/

Newton, P. E. (2007). Clarifying the purpose of educational assessment. *Assessment in Education*, *14*(2), 149–170.

Newton, P. E. (2012). Validity, purpose and the recycling of results from educational assessments. In J. Gardner (Ed.), *Assessment and learning: Second edition* (pp. 264–277). London: Sage.

NMER. (2002). *First class from first grade. Official Norwegian report, 2002: 10*. Oslo.

NMER. (2007). *Summary of report No. 16 (2006–2007) to the Storting: Early intervention for lifelong learning*. Retrieved from

http://www.regjeringen.no/Rpub/STM/20062007/016EN/PDFS/STM200620070016000EN_PDFS.pdf.

NMER. (2009). *White paper no. 11, 2008–2009. Læreren, rollen og utdanningen [The teacher, the role and the training]*. Retrieved from http://www.regjeringen.no/nb/dep/kd/dok/regpubl/stmeld/2008-2009/stmeld-nr-11-2008-2009-/5/1.html?id=544961

NMER. (2010). *National curriculum regulations for differentiated primary and lower secondary teacher education programmes for years 1–7 and years 5–10.* Retrieved from http://www.regjeringen.no/upload/KD/Vedlegg/UH/forskrifter/National_Curriculum_Differentiated_Teacher_Education.pdf

NMER, & NDET. (2006). *Læreplanverket for kunnskapsløftet [National curriculum for knowledge promotion in primary and secondary education and training]*. Oslo.

Nordahl, T., & Hausstätter, R. S. (2009). Spesialundervisningens forutsetninger, innsats og resultater [The conditions, efforts and results of special needs education]. Hamar University College. Hamar.

Nortvedt, G. A., Elvebakk, G., & Lindstrøm, T. L. (2010). *Norsk matematikkråds forkunnskapstest 2009 [Norwegian Mathematics Council's entrance test 2009]*. Oslo: University of Oslo.

Nushe, D. Earl, L. Maxwell, W., & Shewbridge, C. (2011). *OECD reviews of evaluation and assessment in education, Norway.* OECD: OECD.

OECD. (2013). *PISA 2012 results. Student performance in mathematics, reading, science. Volume I*: OECD.

Peddar, D., & James, M. (2012). Professional learning as a condition for assessment for learning. In J. Gardner (Ed.), *Assessment and learning: Second edition* (pp. 33–49). London: Sage.

Pinto, J., & Santos, L. (2013). *Contributions pour un dialogue entre l'évaluation formative et sommative [Contributions for a dialogue between formative and summative assessment*]. *Actes du 25ème colloque de l'ADMEE–Europe: Evaluation et autoévaluation, quels espaces de formation [Proceedings of ADMEE–Europe: Assessment and self-assessment, what areas of training]*. Fribourg. Retrieved from http://www.admee2013.ch/ADMEE-2013/7.html

Robson, C. (2002). *Real world research: Second edition.* Malden, MA: Blackwell Publishing.

Sadler, D. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education, 35*(5), 535–550.

Santiago, P., Donaldson, G., Looney, A., & Nusche, D. (2012). *OECD reviews of evaluation and assessment in education: Portugal*. OECD. Retrieved from http://www.oecd.org/edu/evaluationpolicy).

Santos, L. (2008). Um olhar sobre o Plano da Matemática [A look at the Plan of Mathematics]. *Educação e Matemática [Education & Mathematics], 97,* 3–6.

Santos, L., Pinto, J., Rio, F., Pinto, F., Varandas, J., Moreirinha, O., ... Bondoso, T. (2010). *Avaliar para aprender: Relatos de experiências de sala de aula do pré-escolar ao ensino secundário [Assessing for learning: Reports of experiences of the classroom from preschool to secondary school]*. Porto: Porto Editora e Instituto de Educação, Universidade de Lisboa.

Serrazina, M. (2013). O programa de formação contínua em matemática para professores do 1.º ciclo e a melhoria do ensino da matemática [The lifelong learning programme in mathematics for primary teachers and the improvement of mathematics teaching]. *Da Investigação às Práticas [From Research to Practice]*, *3*(2), 75–97.

Shepard, L. (2001). The role of classroom assessment in teaching and learning. V. Richardson (Ed.), *Handbook of research on teaching* (pp. 1066–1101). Washington: American Educational Research Association.

Statistics Norway. (2014). Fakta om utdanning 2014: Nøkkeltall fra 2012 [Facts about education 2014: Key numbers from 2012.] Oslo: Statistics Norway.

Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Oxon: Routledge.

Suurtamm, C., Koch, M., & Arden, A. (2010). Teachers' assessment practices in mathematics: Classrooms in the context of reform. *Assessment in Education: Principles, Policy & Practice, 17*(4), 399–417.

Throndsen, I. S., & Alseth, B. (2013). Kartleggingsprøver på barnetrinnet - prøvenes formål og prinsipper for prøveutvikling. In T. H. Hopfenbeck, M. Kjærnsli & R. V. Olsen (Eds.), *Kvalitet i norsk skole. Internasjonale og nasjonale undersøkelser av læringsutbytte og undervisning.* (pp. 187–199). Oslo: Universitetsforlaget**.**

Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal, 27*(5), 615–631.

Yin, R. K. (2003). *Case study research: Design and methods*. London: SAGE Publications.

van Daal, V., Solheim, R. G., Gabrielsen, N. N., & Begnum, A. C. (2007). *Norske elevers leseinnsats og leseferdigheter. Resultater fra fjerde og femte trinn i den internasjonale studien PIRLS 2006*. Stavanger: Lesesenteret.

Wiliam, D. (2007). Keeping learning on track. In F. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1053–1098). Charlotte: Information Age Publishing.