

**Title: Policy impact of PISA on mathematics education: The case of Norway**

**Running head: Policy impact of PISA**

**Abstract**

This article addresses the policy implications of participation in international large-scale assessments (ILSAs), particularly the Programme for International Student Assessment (PISA), and the ways in which such implications might influence mathematics education. Taking Norway as a special case, this discussion focuses on insights into teaching, learning and assessment practices that can be inferred from the PISA study, and how participation in ILSAs has contributed to educational policy and even changed policy makers' perspectives on schools, teachers and students. Following publication of the PISA 2000 results, Norway experienced a 'PISA shock', leading to the implementation of a national quality assessment system and national tests. In addition, changes were made to the mathematics curriculum for compulsory school and to mathematics teacher education. More recently, public debate has focused less on rank and league tables, shifting instead to the high number of low-achieving students and the low number of high achievers. Moreover, there has been little uptake of policy advice provided by the Organisation for Economic Co-operation and Development (OECD), which focuses on strengthening accountability measures. Furthermore, although the Norwegian educational system in the past decade has undergone a decentralisation process, the educational system still follows the Nordic model, which focuses on equity and 'education for all'. Analyses of the Norwegian case indicate that policy making takes place in highly cultural contexts, and that international studies might be used merely to validate existing policy directions.

**Keywords**

international large-scale assessment, PISA, policy making, curriculum reform, Norway, mathematics education

## **Introduction**

Over the past two decades, several international large-scale comparative assessments in mathematics have been conducted with primary and secondary students, mathematics teacher students, teachers and adults. Among these, studies such as the Programme for International Student Assessment (PISA), Trends in Mathematics and Science Study (TIMSS), Trends in Mathematics and Science Study Advanced (TIMSS Advanced), Programme for International Assessment of Adult Competencies (PIACC) and Teacher Education and Development Study in Mathematics (TEDS-M) have provided participating countries with substantial information about their education systems' outcomes that can be used by policy makers, researchers, teacher educators, principals and teachers in their decision making. An alternative view of international large-scale studies is to consider them as a source of political communication (de Lange, 2007). According to de Lange, discussions of the outcomes of international large-scale assessments (ILSA) are often about politics rather than performance. In addition, the views expressed about such outcomes may be political, and chains of argument 'may lead [to] . . . weak reasoning [that is] not . . . based on the actual data, or taking the data too seriously' (de Lange 2007, p. 1112).

In evaluating the outcome of an ILSA, policy makers, academics and the media may be primarily interested in country rankings and interpreting that outcome as an indication of the quality of a country's education system. However, such interpretations of the affordances and insights offered by comparative studies are naïve (Auld & Morris, 2016) because they reduce the complexity of the outcome field and miss opportunities to identify insights that could be used to learn valuable lessons about school effectiveness and to inform national educational policies. ILSAs, including PISA, may indeed 'provide benchmarks that help countries align themselves with a scale that indicates a country's position in an international context' (Sälzer & Prenzel, 2014, p. 57). Nevertheless, although Sälzer and Prenzel advocated comparison of international rankings, comparing a country to itself might prove more valid and fruitful. For example, a positive trend line might indicate that educational innovations have led to progress; a flat trend line might indicate little progress, and a negative trend line might suggest educational decline. In this way, measures of performance over time might prove to be as

informative as the average score or ranking. This is especially the case where published rankings are viewed as league tables that can be used to shame and blame a country's performance (Stobart, 2008).

The aim of this article is to reflect on affordances and implications of large-scale assessment studies and how they might shape education, using PISA and Norway as a special case, mainly restricting the discussion to mathematics education. Discussion of how PISA participation has shaped education policy in Norway in general, and mathematics education in particular, offers an useful illustration of the process of evidence-based policy making and might serve to inform similar discussions of policy in other education systems. Drawing on the Norwegian experience of PISA, the present discussion focuses on potential insights into teaching and learning practices provided by the ILSA, and on how participation in large-scale assessments may have altered policy makers' perspectives on schools, teachers and students.

### **International large-scale studies**

Middleton, Cai and Hwang (2015) stated that a large-scale study is large in comparison to 'something', and that certain kinds of studies fall into this category. Following Andersson and Postlethwaite (2007), Middleton and colleagues (2015) identified five criteria for large-scale studies: 1) sample size, 2) purpose of the research, 3) generalisability of the results, 4) type and complexity of the data analysis and 5) cost. Large-scale studies employ a large representative sample, and a complex sample design is often used to ensure a representative sample of diverse groups. The purpose of such studies is to describe systems as a whole rather than individual students, schools or school districts. A further goal of ILSAs is to compare and contrast the systems or groups within a system (Cai, Hwang & Middleton, 2015; Middleton et al., 2015); the results can then be generalised to the education system. A key aim of international studies is to model different education systems (Mullis, Martin, Foy & Hooper, 2016a; Mullis, Martin, Foy & Hooper, 2016b; OECD, 2016) in order to identify the affordances and strengths of each system. The complexity of the sample and of the target object is reflected in the complexity of the methods used to analyse the data. In international studies, large published volumes explain these complexities for those who wish to use outcomes or data for any of a

broad range of purposes, such as policy making or secondary analyses (e.g. Klieme, 2017; Martin, Mullis & Hooper, 2016; OECD, 2017a).

### ***Why participate in ILSAs such as PISA?***

Participating in large-scale international comparative studies is costly, requiring investment in instrument adaptation, translation, data collection and a team of researchers to conduct the study and analyse the data. In addition, students and teachers must allocate valuable instruction time to participate in such assessments. Why, then, should a country participate in a large-scale international study? Postlethwaite (1988) identified four reasons for participating: 1) to identify what is happening in different countries; 2) to identify similarities and differences between education systems; 3) to estimate the relative effects of important variables; and 4) to identify general principles concerning educational effects. Recently, Sälzer and Prenzel (2014) claimed that ILSAs provide a benchmark against which countries can measure themselves by ‘collecting and analysing empirical data on educational institutions, processes and outcomes [that] provides many institutional and political players with profound evidence that can help in different ways when decisions have to be made regard the educational system’ (p. 60).

While Postlethwaite (1988) was referring to comparative studies in general, similar objectives can be said to apply in ILSAs in mathematics education, where studies are a way of generating theories and identifying patterns in teaching practices that enable student learning. Lockheed (2015) noted that an increasing number of countries are participating in PISA; for many education systems, this participation builds capacity because it increases the likelihood of participating in other international studies (e.g., TIMSS). Such participation is likely to provide education authorities in the participating countries with high-quality and internationally comparable information in relation to their students and education system. Moreover, large-scale research has enabled researchers to ‘discover differential patterns in socioeconomic, gender, and ethnic groups and point out that, as a system, mathematics curriculum and instruction has hardly been equitable to all students’ (Middleton et al., 2015, p. 1). To policy makers and educationalists who are concerned about equity and mathematics education for all, large-scale studies may provide requisite insights for decision making.

According to Cai and colleagues (2016), four types of lessons can be learned from ILSAs in mathematics: 1) lessons that help us to understand students' mathematical thinking; 2) lessons about students' experiences with mathematics teaching and their disposition toward mathematics; 3) lessons on classroom instruction; and 4) lessons on how to make global research locally meaningful (e.g., the TIMSS in South Africa). Although information regarding mathematics teaching and learning might be interpreted as feedback about the educational system, such as the success of mathematics education in a country, Cai, Mok, Reddy and Stacey (2016) focused on the teaching and learning of mathematics. Although both levels of analysis might constitute valid and important reasons for participating in large-scale studies of mathematics education, the two are targeting different publics. While policy makers might be interested primarily in information about the education system's success as compared to other countries, researchers in mathematics education, as well as principals and school administrators, might regard insights that can improve the teaching and learning of mathematics as the primary reason for participating in ILSAs.

### ***Reactions to ILSA outcomes***

When ILSA outcomes are published, policy makers, the media and others often react strongly to the results. While some countries may experience a 'shock' (as in the 'PISA shock' experienced in Germany in 2001, please see, de Lange, 2007, and in Norway, please see, Bergersen, 2005), others display pride or remain indifferent (like the US, please see, de Lange, 2007). According to Bergersen (2005), prior to 2001, Norwegians in general and Norwegian policy makers in particular believed that Norway had one of the best education systems in the world. However, the PISA 2000 report published just before Christmas in 2001 showed that Norwegian students scored at the international average, which was far below expectations. The Norwegian Minister of Education at the time stated that this outcome resembled the failure to bring home medals from the Winter Olympics (Bergersen, 2005). At the same time, however, the high average achievement of Finnish students received little attention in Finland. Indeed, it can be argued that while many countries experienced a PISA shock in 2001, Finland was more or less indifferent (e.g. Välijärvi, Linnakylä, Kupari, Reinikainen & Arffman, 2002). According to Pons (2016), France was also initially indifferent to PISA. However, the climate

changed from 2004 onwards, and more attention has been paid to the national results. Responses to ILSA outcomes, and to PISA in particular, are likely to differ in both cycles and countries, depending on factors such as how central the survey topic is seen to be for the national education system and ongoing educational debate (Steiner-Khamsi, 2003; Baird, Johnson, Hopfenbeck, Isaacs, Sprague & Stobart, 2016).

### ***Potential use of ILSA outcomes at national level***

The knowledge and insights gained from participating in ILSAs should be used to improve participants' education systems. The main goal of large-scale studies is to improve learning (Cai et al., 2016) as highlighted by the Organisation for Economic Co-operation and Development (OECD) (e.g. 2013b). The goal of the OECD and the IEA is to provide participating countries with evidence of educational outcomes that can be used to monitor progress and help to highlight possibilities and develop new policies (Cai et al., 2016; Baird et al., 2016). As proposed by Lockheed (2015), the information provided by ILSAs might also be used to analyse an education system's success. For instance, Hong Kong has used PISA 'for understanding the quality and equality of the Hong Kong basic education systems' (Ho, 2016, p. 518). Participation in international studies may also contribute to capacity building for the design and implementation of national assessments, where standards or measurement approaches are inspired by or adopted from ILSAs (Lockheed, 2015).

Educational reforms may be standards-based or result from ideal governance. Many countries use PISA outcomes as a rationale for changing curricula and revising educational standards (Baird et al., 2016; Breakspear, 2012). Sälzer and Prenzel (2014) discussed the ways in which systematic analyses and interventions linked to PISA led to substantial improvements in the German educational system and to a significant increase in the national PISA score. Although these outcomes can be interpreted as the result of 'teaching to the test' (e.g. Berliner, 2011), Sälzer and Prenzel emphasised systemic changes in national standards, national assessments, teachers' professional development and teaching and learning innovations rather than surface-level changes and teaching to the test. Although participation led to the 'PISA shock' of 2001, it also provided some unexpected information about the education systems in many 'Länder' (states).

Regarding mathematics education, Burkhardt (2014) claimed that, in revising their mathematics curricula, many countries have been inspired by the strong focus on modeling and problem solving in PISA's mathematics assessment framework and in international mathematics education research. Curriculum reforms implemented at national level were informed by the fundamental goals of mathematics education described by Niss (1996); among these, the primary goal is to educate citizens who can apply mathematics in a variety of contexts, and contribute to democracy and to societal, financial and technological growth of the society. At the same time, it is important to note that problem solving and modeling have been major research topics in the mathematics education literature over the past 50 years (Lesh & Doerr, 2003; Lesh & Zawojewski, 2007; Niss, Bruder, Planas, Turner & Villa-Ochoa, 2016). In this context, the PISA framework and much of the international research has been inspired by the work of the Danish professor Mogens Niss (Kilpatrick, 2014).

PISA has also been used to monitor the outcomes of mathematics education at national level (Breakspear, 2012) or to argue for implementation of national assessments to monitor educational standards (Baird et al., 2016; Elstad, Nortvedt & Turmo, 2009). It follows that PISA and other ILSAs might influence which mathematical content is valued, even when changes are not made to the curriculum—or perhaps not. There is some evidence that policy makers use PISA outcomes merely to validate decisions that have already been made; for example, in a study of policies allegedly informed by PISA data in six educational economies, Baird and colleagues (2016) found that although references to scandal and shock were used to argue for reforms, the corresponding innovations were mainly standards-based. Policy makers advanced different reasons to justify the educational reforms initiated in each of the six countries. Baird and colleagues (2016) also found that although the policy reaction to PISA outcomes seemed to lead to educational reform, countries with similar PISA results initiated very different policies. In the six countries included in their case study, the proposed reforms grounded on ILSA outcomes were mainly changes that had already been decided, but where the ILSA outcome could be used to argue in favour of the reform. These findings indicate that policy making

takes place in highly cultural contexts, and that international studies like PISA are often used merely to validate existing policy directions.

Both the IEA and the OECD share their data with external researchers, and datasets are publicly available through project websites shortly after the corresponding international reports are published. However, the instruments used to measure student achievement are not released; instead, only a small sample of test items is published so that future trend studies are not ‘destroyed’. This lack of information can make it difficult for external researchers to understand in depth what these international studies are measuring. There is also a risk of oversimplification because a country’s average achievement is likely to be interpreted as indicating the level of mathematical knowledge held by students in that country, with little discussion of what it comprises (Gorur & Wu, 2015). Similarly, although background questionnaires are published, ‘large-scale studies are prone to errors due to “fishing”’. Because, particularly for secondary data analysis, researchers have access to so many variables at once, the tendency to run analyses without clear hypotheses or theoretical justification is almost too easy’ (Middleton et al., 2015, p. 11). Finally, these international studies are sometimes used to analyse something they do not measure; for instance, Zhao and Meyer (2013) used PISA ranking in conjunction with Global Entrepreneurship Monitor ranking to suggest that PISA is not a valid measure of students’ preparedness for adult life.

### **What PISA measures**

Conducted by the OECD, PISA is a triannual international large-scale study that aims to evaluate education systems worldwide by assessing the competencies of 15-year-old students. Specifically, students’ reading literacy, mathematical literacy and scientific literacy are assessed in each cycle; one of the three assessed areas is the major domain while the others form are treated as minor domains (OECD, 2017a). However, the scope of PISA is widening. In 2015, the PISA assessment encompassed cooperative problem solving and financial literacy as well as reading, scientific and mathematical literacy (OECD, 2016). In total, half a million 15-year-olds in 72 countries and economies participated in the 2015 PISA assessment (OECD, 2017a). In 2018, students’ global competence will also be measured.



The OECD aims to provide participating countries and economies with data and interpretations (such as policy briefs, country notes and reviews) that may shape educational policy and inspire change and improvement (e.g. Breakspear, 2012; Hopfenbeck, Tolo, Florez & El Masri, 2013; Nusche, Earl, Maxwell & Shewbridge, 2011; OECD, 2017a). The OECD also aims to support secondary analyses conducted by other institutions and researchers. After each PISA cycle, the assessment data are made available to the public upon release of the international reports.

The PISA study has been criticised by many researchers, lobbyists, teachers and policy makers worldwide. These critics have argued that PISA is an invalid assessment, that the measurement model underlying the analysis is invalid and that the OECD is globalising education (Baird et al., 2016; Kreiner & Christensen, 2014; Goldstein, 2017; Hopfenbeck, Lenkeit, El Masri, Cantrell, Ryan & Baird, 2017). Because all research and assessments can be criticised, care should be taken to avoid invalid extrapolations of the measured trait when interpreting assessment outcomes. However, as the main aim here is to discuss current use of PISA outcomes to inform education policy, this article prioritises a balanced discussion of this practice rather than reviewing the critical literature.

### ***Defining mathematical competence as modelling and problem solving***

The PISA study assesses mathematical literacy, which is defined as

an individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals in recognising the role that mathematics plays in the world and to make the well-founded judgements and decisions needed by constructive, engaged and reflective citizens. (OECD, 2013a, p. 25)

This definition focuses on preparing individuals to participate in society, to fulfil individual goals and to contribute to both democracy and growth, which might be seen as the 'fundamental goals of mathematics education' (see Niss, 1996, for a discussion of the purpose of mathematics education). According to the PISA definition of mathematical literacy, mathematical competence is linked to the application of mathematics (i.e. being able to reason mathematically, knowing procedures and knowing how to apply tools) in problem solving and modeling—that is, formulating real-world situations mathematically, applying mathematical procedures and strategies to solve problems, and interpreting and evaluating the outcomes of such modeling.

In 2003 and 2012, mathematics was the major domain of the PISA assessment. In those years, students' mathematical achievement was measured along with their beliefs about and attitudes to mathematics. For instance, in 2012, the PISA assessment included questionnaires that addressed the attitudes of principals and students to teaching and learning mathematics, as well as paper- and computer-based tests assessing students' mathematical competence (OECD, 2013b, 2013c). A separate questionnaire was developed for parents. In the other PISA cycles (2000, 2006, 2009 and 2015), only mathematical achievement was measured. Several trend items were used in each cycle, enabling comparison over time. Until 2012, the assessment was exclusively paper-based; in that year, a computer-based assessment was also administered for the first time. In 2015, the assessment was completely computer-based but included only trend items from prior paper-based tests, as science was the major domain in that year (OECD, 2016).

### **The case of Norway**

The Norwegian education system can be characterised as following the Nordic model (Imsen, Blossing & Moos, 2016). More than 96 per cent of Norwegian students attend public schools (NDET, 2017), and the education system is regarded as an important means of ensuring welfare for all, emphasising qualities such as social justice, equity, equal opportunities, inclusion, education for all and nation building (Imsen et al., 2016). The national mathematics curriculum has traditionally focused on 'mathematics for all', and topics such as problem solving, the mathematics of everyday life and mathematical thinking and reasoning for students at all achievement levels were emphasised in the last curriculum reforms. All schools, even private ones, must follow this curriculum; this strong focus on offering the same content and competencies to all students can be considered a core value in the Nordic model.

Norway has participated in several international comparative studies, including all six PISA cycles to date. Along with other ILSAs, PISA has been identified as a source of data for monitoring progression over time in the national Quality Assessment System (QAS)<sup>1</sup> (Elstad et al. 2009). For that

---

<sup>1</sup> Please see [https://webgate.ec.europa.eu/fpfi/mwikis/eurydice/index.php?title=Norway:Quality\\_Assuranceandprintable=yes](https://webgate.ec.europa.eu/fpfi/mwikis/eurydice/index.php?title=Norway:Quality_Assuranceandprintable=yes) for further information about QAS.

reason, PISA is seen as an important source of knowledge about the Norwegian education system and especially about the success of lower secondary education.

### *Norwegian PISA results: Stable outcome with equity challenges*

Table 1 shows the main mathematical literacy achievement outcomes in all cycles based on international and national PISA reports. Average scores with standard errors are shown for Norway and OECD, allowing comparisons of average achievement between Norway and participating OECD countries. Although first conducted in 2000, PISA 2003 was the first to focus on mathematics, and the mathematics framework was fully implemented in this cycle (OECD, 2013a). For that reason, data from 2000 are not included in the table or discussion. Nevertheless, it may be useful to know that Norwegian students on average scored 499 (SD = 92, SE = 2.8) in PISA 2000 while the OECD average was 500, with a standard deviation of 100. This means that Norwegian students scored at the international average in mathematical literacy in 2000.

Table 1.

*Achievement scores and proportion of high and low achieving students in Mathematical Literacy for PISA Cycles 2003–2005 for Norway and OECD.*

<b>Cycle</b>	<b>2003</b>	<b>2006</b>	<b>2009</b>	<b>2012</b>	<b>2015</b>
Norway Average (SE)	495 (2.4)	490 (2.6)	498 (2.4)	489 (2.7)	502 (2.2)
OECD Average (SE)	499 (0.6)	497 (0.5)	494 (0.5)	494 (0.6)	490 (0.4)
Norway Gender differences (SE)	6 (3.2)	-	-	2 (3.0)	-2 (2.8)
OECD Gender differences (SE)	<b>10</b> (0.8)	-	-	<b>11</b> (0.6)	<b>8</b> (0.6)
Norway (OECD) Proportion low achievers (scoring below level 2)	21 (22)	22 (23)	18 (22)	22 (23)	17 (23)
Norway (OECD) Proportion high achievers (scoring at level 5 and 6)	11 (14)	10 (13)	10 (13)	9 (13)	10 (11)

*Note.* Numbers are from OECD (2013); OECD (2016); Kjærnsli and Olsen (2013); significant differences are in bold.

Table 1 shows that Norway's average mathematics achievement score was close to the OECD average in all cycles, and that in 2015, Norway scored significantly higher than the OECD average for the first

time (OECD, 2016). Although the average scores of Norwegian students seem to differ over time, the differences are small when compared to the standard error. Moreover, the Norwegian trend line (showing progress) is flat (OECD, 2016), confirming that over the past 15 years, innovations in mathematics education have yielded little change in student outcomes. Interpretation of the Norwegian results as stable might be challenged by the outcomes of other large-scale studies, especially the TIMSS study, which revealed that although Norwegian students scored close to the international average in PISA in 2003–2004, student achievement decreased in other studies only to rise again in subsequent years (Olsen, Hopfenbeck & Lillejord, 2013). However, Table 1 confirms that overall PISA outcomes were more or less stable. This stable pattern of achievement might be perceived as a positive result, as many countries that Norway often looks to when seeking to possible education policies or curriculum improvement (such as Sweden and Finland) showed a steep decline in the same period (OECD, 2016).

In the Nordic model, equity and equal opportunities to learn mathematics are two of the driving forces. Over the years, gender differences have been small and mainly non-significant. However, the more or less stable differences between majority and minority students indicate an achievement gap between these groups. In 2012, the difference between majority and minority students was 40 points, an increase of 3 points from 2003 (Olsen, 2013). As a difference of this size is comparable to one year of mathematics teaching, it might be concluded that although Norwegian schools achieve gender equity in mathematics outcomes, the education system is less successful in terms of equity of achievement between minority and majority students. On a positive note, the influence of socioeconomic status (SES) on achievement is smaller in Norway than in countries outside the Nordic area, and only small differences have been observed between schools (OECD, 2013b, 2016), indicating that Norway's education system is equitable. However, the large within-school differences are a key concern, as these may indicate sizeable differences in teaching quality within the same school. Because Norwegian schools are inclusive, and students are not streamed or tracked, these within-school differences may indicate that some teachers have more success in teaching diverse classrooms. This interpretation is supported by the large number of low-achieving

students and the low number of high achievers, which is a more or less stable pattern (Nortvedt & Pettersen, 2016). Similar patterns have been observed in Norwegian students' level of achievement in several international comparative studies (PISA, TIMSS, TIMSS Advanced) (Olsen, Hopfenbeck & Lillejord, 2013). These studies have shown that, with the exception of statistical knowledge, Norwegian students have been relatively successful in solving applied problems that do not require specific mathematical knowledge. Previous studies have shown that any shortcomings often reflect a lack of knowledge of the algebraic language of mathematics (Grønmo & Bergem, 2009; Grønmo, Ostad, Nilsen, Hole, Aslaksen & Borge, 2012; Nilsen, Angell & Grønmo, 2013; Nortvedt, 2013b), even in TIMSS Advanced Physics (Nilsen et al., 2013; Ræder, 2017).

The PISA study also collects data on students' views of themselves and their experiences of mathematics teaching. So, what does PISA tell us about mathematics teaching in Norway? Student responses to survey questions about instructional quality suggested that Norwegian teachers probably focus less on structuring activities in mathematics lessons that might consolidate student learning. In addition, students reported taking part in cognitively stimulating activities less frequently than their peers in other OECD countries (Olsen, 2013). However, it is important to keep in mind that these findings are based on students' reports on classroom activities and their perceived understanding of instructional quality.

Students' attitudes and beliefs about mathematics, mathematics teaching and themselves demonstrated that Norwegian students are less motivated, have lower self-belief, show less perseverance and experience more anxiety than students on average in OECD countries (OECD, 2013c). The evidence also suggests that attitude and achievement are more strongly related among Norwegian students than in many other countries, as the effect of more positive attitudes and beliefs on achievement was much higher in Norway.

### **Use of PISA results for policy making in Norway**

In Norway, the amount of media coverage indicates that the media, policy makers and researchers pay significant attention to PISA (Bergersen, 2005; Hopfenbeck & Gørgen, 2017; Sjøberg, 2014). The outcomes of PISA and other comparative studies have been extensively used to provide

information for national white papers on Norway's education system (Elstad et al. 2009) and in policy making (Baird et al. 2016; Elstad et al. 2009; Tveit 2014). Subsequent policy recommendations have been based—at least to some extent—on discussion and argumentation on international comparative studies in national and international reports.

When the PISA 2000 outcomes were published in December 2001, Norway experienced a 'PISA shock' (Bergersen, 2005). The average results of Norwegian students contradicted the longstanding national belief that the Norwegian education system was highly efficient and that its students' knowledge and competency levels were high. Following that sobering revelation, there was much discussion of the quality of Norwegian curricula, teacher education and schools. A national committee was appointed to formulate initiatives to improve Norwegian primary and secondary education. Based on the committee's recommendations, proposals included national tests in reading (in Norwegian and English) and mathematics, and these were implemented in 2004 (Elstad et al. 2009). Most policy makers and the media reacted in similar fashion to the 'PISA shock' that followed subsequent PISA studies (Baird et al., 2016), while others claimed that the PISA outcomes told a positive story about the Norwegian education system. For instance, the Minister of Education took a positive view of the PISA 2015 outcomes that could even be interpreted as expressing pride in Norwegian students and teachers (Ministry of Education and Research, 2016b). PISA 2015 focused on science, but the press release emphasised good reading outcomes, indicating that Norwegian policy makers used the PISA outcomes to support rhetorical arguments about the Norwegian education system as indicated by Baird and colleagues (2016).

When the Norwegian national quality assessment system (NQAS, later QAS) was implemented in 2004, international comparative studies were incorporated to monitor the Norwegian education system at the national level (Baird et al., 2016; Elstad et al., 2009). In addition, national tests were used to assess the outcomes of teaching, informing the education system at several levels, from school and local municipality to aggregated national level. Prior to 2004, Norway did not administer national tests; in fact, there were no national assessments at that time that could be used to monitor the Norwegian education system. National exams were not (and are not) piloted, linked or equated, and

national tests were just about to be launched. Not until 2014 was an anchor design introduced that enabled national tests to be linked. For that reason, international comparative studies were the only available assessments offering reliable trend data that might indicate change and that could be used to monitor the Norwegian education system's achievement outcomes (Elstad et al. 2009). When compared to countries such as the United Kingdom, the Netherlands and Germany, Norway is a special case, as there were no national assessments designed for monitoring purposes prior to 2014.

The PISA framework has influenced both the national test frameworks and the national curriculum framework for basic skills, which strongly resemble the PISA framework (Frønes, Roe & Vagle, 2012; Nortvedt, 2013a; Tveit, 2014). Two years after they were implemented, the national tests were revised. The Ministry of Education and the Directorate for Education and Training jointly decided that, instead of a mathematics assessment, a numeracy assessment should be implemented, and the test framework developed for this assessment was also closely related to the PISA mathematical literacy assessment framework (Nortvedt, 2013a). However, this may be a chicken-and-egg situation, as the PISA mathematics framework draws heavily on the mathematics framework previously developed in the Danish KOM report on competences in mathematics education (Niss, Blum & Galbraith, 2007; Niss et al., 2016; Niss & Højgaard, 2011; OECD, 2013a; Turner, Dossey, Blum & Niss, 2013) and is also influenced by the recent focus on modeling in mathematics education (Niss, 2007; Niss et al., 2007; Niss & Jablonka, 2014). This reciprocal situation seems to reflect the influence of international research on the PISA study and the influence of PISA outcomes on national policy making. Like the PISA framework (Burkhardt, 2014), the KOM report influenced curricula in several countries (Kilpatrick, 2014), and Norway's development and adoption of the basic skill numeracy framework most probably followed these international influences.

The status of PISA and TIMSS as part of the QAS was strengthened when the latest national strategy to enhance learning in mathematics and the sciences (*Realfagstrategien*) was implemented in 2015. Combined with the national numeracy tests, these two ILSAs were identified as indicators of the national strategy's success in raising the number of high-achieving students in mathematics and science in compulsory education while lowering the number of low achievers in the same subjects.

The distribution of Norwegian students' PISA and TIMSS scores were to be used as indicators of the Norwegian education system's ability to improve student knowledge and competences (Ministry of Education and Research, 2015). It might be argued that the use of ILSAs in formulating policies and reforming the national quality assessment systems reflects the emerging focus on performance measurement, accountability, decentralisation and local autonomy after the turn of the millennium (Imsen et al., 2016). However, Baird and colleagues (2016) drew very different conclusions, indicating that PISA was not the driver.

In the Norwegian context, the PISA results have been used not only to advocate for the implementation of the QAS but also to argue for the new national curriculum, The Knowledge Promotion, which was implemented in 2006 (Baird et al., 2016). According to Bergersen (2005), policy makers saw PISA as a gift because it could be used as a tool for justifying policies. In addition to the average PISA results, other international studies reported a decrease in Norwegian students' achievement in mathematics, science and reading in 2003 (Olsen, Hopfenbeck & Lillejord, 2013), which national white papers used in conjunction with PISA outcomes to argue for reform. In another clear example of the direct use of results of international comparative studies, the Ministry of Education and Research (2012) asked for 'more algebra' in the latest adjustment of the national mathematics curricula, based on the outcomes of TIMSS study.

Further changes in Norway's education system may reflect a stronger focus on formative assessment, which many see as a response to OECD country reviews claiming that Norway lacked an assessment culture (Nortvedt, Santos & Pinto, 2016). Assessment for learning has been identified as a general guiding principle of the Education Act and is the focus of a large national professional development project initiated by the Directorate for Education and Training (Hopfenbeck et al., 2013). Baird and colleagues (2016) claimed that the implementation of assessment for learning as a national strategy and a national portal for displaying outcomes of national tests and exams was consequences of the Norwegian 'PISA shock'. In addition, insights from PISA, TIMSS and TEDS-M have been used to advocate change in mathematics teacher education (Birkeland & Breiteig, 2012; Breiteig, 2013; Helgøy & Homme, 2006). Until 2011, Norway had a general teacher education programme that



allowed qualified teachers to teach all subjects in grades 1–10 in compulsory school—that is, to function as classroom teachers. Starting in 2011, two distinct programmes were established: educating general teachers to teach grades 1–7 and subject-oriented teachers to teach grades 5–10. Since 2017, all teacher education programmes in Norway are master’s degree programmes (Ministry of Education and Research, 2016a). Many of Norway’s education reforms follow from 2001, when the first PISA study was published, and PISA results have contributed to the rhetoric underlying these changes.

## **Discussion**

As PISA outcomes inform policy making in several countries (Baird et al. 2016), PISA’s possible influence on Norway’s education policies may provide insights that are useful in other contexts. The present discussion sought to examine policies implemented in light of Norwegian traditions such as the Nordic model.

Norway was one of five case studies in an external evaluation of the policy implications of PISA conducted on behalf of the OECD PISA Governing Board. Breakspear (2012) rated the influence of PISA data on policy making in Norway at 10 on a scale of 1 to 14, indicating that Norway has taken home many messages. These have mainly been used to revise curriculum standards and to develop and implement PISA-like competencies in the curriculum. On that basis, PISA can be seen as a major influence on the Norwegian education system. Breakspear claimed that Norway looks to Finland to understand driving forces behind higher student achievement, and policy makers may consider this appropriate, as both Norway and Finland follow the Nordic model. For instance, the newly implemented master’s degree studies in teacher education can be seen as a consequence of the attempt to understand the significant differences among the Nordic countries. In PISA, Finland has traditionally outscored Norway in mathematical literacy (OECD, 2013b), and Finnish teachers are required to have master’s degrees. However, Norway’s implementation of national assessments to monitor student learning and of the QAS seems very different from the Finnish system.

The Norwegian school system has been based on the principles of ‘education for all’ and trust in teachers and school leaders (Imsen et al., 2016). The restructuring of the education system that began after the turn of the millennium was perhaps a response to the PISA shock. Amendments to the

Education Act and the implementation of a new national curriculum (the Knowledge Promotion Reform) focused on individualised education and achievement goals. At the same time, decentralisation was seen to transfer responsibilities from the national level to the municipalities (school ‘owners’) and principals. In this regard, the introduction of national tests and other accountability measures after the 2001 PISA shock indicates a level of inconsistency (Helgøy & Homme, 2006) and might be viewed as ‘recentralisation’ (Imsen et al., 2016).

As the Norwegian language has no word for ‘accountability’ (Elstad et al. 2009), perspectives on monitoring and assessment in Norway may differ from those in more accountability-oriented countries. In a decentralised system, accountability should be implemented at different administrative levels; for instance, Imsen and colleagues (2016) claimed that as national tests are used by school owners (municipalities) and national governments to monitor educational outcomes, national tests form part of both decentralisation and recentralisation processes. However, the QAS continues to refer to the results of international studies as the main national-level assessments. Indeed, the monitoring function of international studies and national tests has recently been strengthened by the implementation of a new national strategy for science and mathematics teaching (*Realfagstrategien*). Nevertheless, despite the implementation of PISA-driven policies, Norway has maintained its social democratic ethos (Helgøy & Homme, 2006), and Norwegian educational policy still prioritises equity, ‘education for all’ and inclusive schooling. The fact that these values are no less important today than in the progressive pedagogy of earlier years indicates that Norway still embraces the Nordic model (Imsen et al., 2016). Unlike the US and UK, Nordic legislation focuses on comprehensive schooling and education for democracy, participation, *Bildung* and equality (Imsen et al., 2016). These values align very well with the fundamental reasons for mathematics education as expressed by Niss (1996). For instance, the goal of the Norwegian mathematics curriculum for compulsory and secondary education is to help students develop a positive attitude to mathematics, as well as skills in problem solving and modeling, enabling them to become engaged and responsible citizens, to be successful and to contribute to the financial and technical growth of society (NDET, 2015).

Baird and colleagues (2016) indicated that although PISA outcomes are used in policy making, they do not necessarily increase the uniformity of educational systems. As PISA and TIMSS outcomes in mathematics are highly correlated, the PISA test is probably a valid assessment of students' mathematical competence (Jerrim, 2013). What lessons has Norway learned, then, from its participation in PISA? The latest curriculum reforms draw on international research with strong roots in the Nordic research communities (e.g. Niss, 1996) and stress the values represented in the Nordic model. These reforms can therefore be regarded as culturally responsive to Norwegian traditions. In combination with outcomes of the TIMSS and national tests, PISA outcomes are used to collect data addressing the four lessons identified by Cai and colleagues (2016). The fact that Norwegian students' average scores are close to the OECD average, public discussion of the 'flat' trend line and the recent shift in focus from rankings to the number of high- and low-achieving Norwegian students (too few and too many, respectively) has resulted in a new national strategy for mathematics and science education (Ministry of Education and Research, 2015). In line with this initiative, many changes have also been implemented to improve teacher education. Among the many policies implemented in recent years, it remains to be seen whether these can improve mathematics teaching and teacher education in Norway, or whether policies must address the content of mathematics education, taking account of classroom activities to improve student learning.

Rather than simply scaling down international comparative research, Cai and colleagues (2016) proposed that large-scale research should be complemented by targeted small-scale comparative research to provide in-depth and culturally sensitive information. This kind of research already exists, as in the TIMSS video study and PISA+ (Hiebert, 2003; Klette, Bergen & Roe, 2016). Cai and colleagues (2016) further argued that education policy should derive from careful attention to local situations as classroom activities can best be interpreted from a cultural perspective (Clarke, 2013). Similarly, policy must be culturally responsive, taking account of historical development to be consistent (Helgøy & Homme, 2006; Imsen et al., 2016).

The OECD recommends aligning policy with governance to ensure efficient policy implementation (Nusche et al., 2011); otherwise, municipalities and local school leaders may

experience difficulties in interpreting and implementing policies locally (Helgøy & Homme, 2006). Nusche and colleagues (2011) proposed that communication of key strategies for policy implementation would reinforce the role and capacities of policy makers at different levels. In the Nordic model, however, schools have a great deal of autonomy (Imsen et al., 2016); although principals are responsible for enhancing student learning, neither the national curriculum nor the *Realfagstrategien* provides guidelines or regulations for reaching the identified goals. The Norwegian government may see the implementation of accountability measures (e.g. national tests) and other indicators included in the latest national strategy as building a ‘culture of evidence’—that is, using data strategically to achieve national goals. However, there is evidence that Norwegian teachers, principals and even school administrators at municipal level find it challenging to implement assessment for learning as a national policy, as they face the difficult task of balancing accountability and trust (e.g. Hopfenbeck et al., 2013).

### **Concluding Remarks**

The aim of this article has been to discuss possible policy impacts of PISA on mathematics education, applying the case of Norway to exemplify how ILSAs can influence national policies. Following the PISA ‘shock’ and subsequent implementation of the quality assessment system (QAS), national tests and the strategy for mathematics and science teaching, Norwegian policy makers seem to have shaped policies to create national means to steer mathematics education, drawing in part on PISA outcomes to inform their decision making. For instance, Breakspear (2012) concluded that Norway has taken home many lessons from PISA, which has exerted a substantial influence on policy making. This conclusion is supported by Imsen and colleagues (2016), although Baird and colleagues (2016) reach a different conclusion, asserting that policies were already decided upon and were not dependent on PISA outcomes. Indeed, the reciprocal relationship between the PISA framework and the framework for basic numeracy skills and the mathematics curriculum suggests that questions of whether PISA influenced policy or whether ‘trends’ in international research influenced PISA may be a chicken-and-egg debate. One possible interpretation of this relationship is that educational policies are discussed and shaped in cultural contexts that serve as lenses through which policies are shaped. In

the case of Norway, the Nordic model's strong emphasis on education for all and inclusion may have had some impact, for instance, on issues related to the proportion of low-achieving students. Additionally, as the Norwegian language lacks a word for 'accountability', and as Norway has traditionally supported the strong autonomy of schools and municipalities, this may contribute to how data is used to shape a national education system in which national tests can be seen to serve both decentralisation and recentralisation.

It is costly to participate in an ILSA such as PISA in terms of both financial and human resources, and for that reason, any investment in an international study should lead to a worthwhile outcome. Postlethwaite (1988) identified four reasons for participating in comparative studies, one of which is 'identifying what happens'. This is apparently the underlying reason for Norwegian participation, as the country's most fundamental educational policy is the QAS, which identifies ILSAs as the main instrument for monitoring the success of the Norwegian education system. In addition, PISA and TIMSS were recently identified as the measure against which the success within mathematics and science education is judged—that is, by raising the number of high-achieving students while lowering the number of low achievers. This can be seen as an instance of subscribing to naïve interpretations as described by Auld and Morris (2016), who called for elaborated readings of assessment outcomes while taking account of the complexity of assessment data.

Today, as the present discussion suggests, implemented policies are viewed as tools to uphold the Nordic model. The current use of ILSA data in policy making supports this interpretation, as Norway mainly uses average scores and the distribution of scores at different levels of achievement to inform policy making. In so doing, opportunities are lost to capture insights that might help in addressing more complex questions—for instance, why Norway has so few high-achieving students, or why majority students score consistently higher than minority students in mathematical literacy. Policy analyses addressing the four reasons established by Cai and colleagues (2016) are less apparent. Of the four reasons, understanding students' mathematical thinking and experiences with mathematics teaching and learning would surely contribute to education for all and to equity and equality in schools, including equal opportunities for majority and minority students. In an education system

where equity is the gold standard, the large and consistent achievement gap between majority and minority students must be addressed, and more emphasis must be placed on understanding this gap and the factors that contribute to it. Perhaps Norway should take home more elaborated lessons from ILSAs such as TIMSS and PISA, keeping in mind that any lessons should also support values embedded in the Nordic model as these are culturally appropriate.

## References

- Anderson, L. W., & Postlethwaite, T. N. (2007). *Program evaluation: Large-scale and small-scale studies*. International Academy of Education: International Institute for Education Planning. [www.unesco.org/iiep/PDF/Edpol8.pdf](http://www.unesco.org/iiep/PDF/Edpol8.pdf)
- Auld, E., & Morris, P. (2016). PISA, policy and persuasion: Translating complex conditions into education 'best practice'. *Comparative Education*, 52(2), 202–229.
- Baird, J.-A., Johnson, S., Hopfenbeck, T. H., Isaacs, T., Sprague, T., Stobart, G., & Yu, G. (2016). On the supranational spell of PISA in policy. *Educational Research*, 58(2), 121–138.
- Bergersen, H. O. (2005). *Kampen om kunnskapsskolen* [The battle for the knowledge-based school]. Oslo, Norway: Universitetsforlaget.
- Berliner, D.C. (2011). Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*. 41(3), 287–302.
- Birkeland, P. A., & Breiteig, T. (2012). Norsk lærerutdanning i et internasjonalt perspektiv [Norwegian teacher education in an international perspective]. In L. S. Grønmo & T. Onstad (Eds.), *Mange og store utfordringer. Et nasjonalt og internasjonalt perspektiv på utdanning av lærere i matematikk basert på data fra TEDS-M 2008* (pp. 35–60). Oslo, Norway: Unipub.
- Breakspear, S. (2012). *The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance* (OECD Education Working Papers No. 71). <http://dx.doi.org/10.1787/5k9fdfqffr28-en>
- Breiteig, T. (2013). Preparing teachers of mathematics in Norway. In J. Schwille, L. Ingvarson, and 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). Hamburg, Germany: IEA.
- Burkhardt, H. (2014). Curriculum design and curriculum change. In Y. Li & G. Lappan (Eds.), *Mathematics curriculum in school education* (pp. 13–33). Dordrecht, Netherlands: Springer.
- Cai, J., Hwang, S., & Middleton, J. A. (2015). The role of large-scale studies in mathematics education. In J. A. Middleton, S. Hwang, and J. Cai (Eds.), *Large-scale studies in mathematics education* (pp. 405–414). Dordrecht, the Netherlands: Springer.
- Cai, J., Mok, I. A. C., Reddy, V., & Stacey, K. (2016). International comparative studies in mathematics: Lessons for improving students' learning. In *ICME-13 topical surveys* (pp. 1–36). Hamburg, Germany: ZDM.
- Clarke, D. (2013). The validity-comparability compromise in crosscultural studies in mathematics education. In B. Ubuz, Ç. Haser, & M. A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 1855–1864). Antalya, Turkey: ERME.
- De Lange, J. (2007). Large-scale assessment and mathematics education. In F. K. J. Lester (Ed.), *Second handbook of research on mathematics teaching* (Vol. 2, pp. 1111–1142). Charlotte, NC: Information Age Publishing.
- Elstad, E., Nortvedt, G. A., & Turmo, A. (2009). The Norwegian assessment system: An accountability perspective. *Cadmo*, 17(2), 89–103.
- Frønes, T. S., Roe, A., & Vagle, W. (2012). Nasjonale prøver i lesing - utvikling, resultater og bruk [National tests in reading, development, results and use]. In T. H. Hopfenbeck, M. Kjærnsli,

- and R. V. Olsen (Eds.), *Kvalitet i norsk skole: Internasjonale og nasjonale undersøkelser av læringsutbytte og undervisning* (pp. 135–153). Oslo, Norway: Universitetsforlaget.
- Goldstein, H. (2017). Measurement and evaluation issues with PISA. In L. Volante (Ed.), *The PISA effect on global educational governance* (pp. 49–58). London; New York: Routledge.
- Gorur, R., & Wu, M. (2015). Leaning too far? PISA, policy and Australia's 'top five' ambitions. *Discourse: Studies in the Cultural Politics of Education*, 36(5), 647–664.
- Grønmo, L. S., & Bergem, O. K. (2009). Prestasjoner i matematikk [Achievement in mathematics]. In L. S. Grønmo and T. Onstad (Eds.), *Tegn til bedring. Norske elevers prestasjoner i matematikk og naturfag i TIMSS 2007* (pp. 49–111). Oslo, Norway: Unipub.
- Grønmo, L. S., Onstad, T., Nilsen, T., Hole, A., Aslaksen, H., & Borge, I. C. (2012). *Framgang, men langt fram. Norske elevers prestasjoner i matematikk og naturfag i TIMSS 2011* [Progress but still a long way to go. Norwegian students' achievement in mathematics and science in TIMSS 2011]. Oslo, Norway: Akademika Forlag.
- Helgøy, I., & Homme, A. (2006). Policy tools and institutional change: Comparing education policies in Norway, Sweden and England. *Journal of Public Policy*, 26(2), 141–165.
- Hiebert, J. et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. Washington, DC: National Centre for Educational Statistics.
- Ho, E. C. H. (2016). The use of large-scale assessment (PISA): Insights for policy and practice in the case of Hong Kong. *Research Papers in Education*, 31(5516–5528).
- Hopfenbeck, T. H., Lenkeit, J., El Mastri, Y., Cantrell, K., Ryan, J., & Baird, J.-A. (2017). Lessons learned from PISA: A systemic review of peer-reviewed articles on the programme for international student achievement. *Scandinavian Journal of Educational Research*. [online first] 1–21.
- Hopfenbeck, T. H., & Gørgen, K. (2017). The politics of PISA: The media, policy and public responses in Norway and England. *European Journal of Education*, 52(2), 195–205.
- 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*. Retrieved from [http://www.oecd-ilibrary.org/education/balancing-trust-and-accountability-the-assessment-for-learning-programme-in-norway\\_5k3txnpqlsnn-en](http://www.oecd-ilibrary.org/education/balancing-trust-and-accountability-the-assessment-for-learning-programme-in-norway_5k3txnpqlsnn-en)
- Imsen, G., Blossing, U., & Moos, L. (2016). Reshaping the Nordic education model in an era of efficiency. Changes in the comprehensive school project in Denmark, Norway, and Sweden since the millennium. *Scandinavian Journal of Educational Research*, 1–16. Retrieved from <http://dx.doi.org/10.1080/00313831.2016.1172502>
- Jerrim, J. (2013). The reliability of trends over time in international education test scores: Is the performance of England's secondary school pupils really in relative decline? *Journal of Social Policy*, 42(2), 259–279.
- Kilpatrick, J. (2014). Competency frameworks in mathematics education. In F. K. J. Lester (Ed.), *Encyclopedia of mathematics education* (pp. 85–87). Amsterdam, Netherlands: Springer.
- Klette, K., Bergem, O.K., & Roe, A. (2016). *Teaching and learning in lower secondary schools in the era of PISA and TIMSS*. Dordrecht, Netherlands: Springer.
- Klieme, E. (2017). International large-scale assessments. Keynote presentation at the 18<sup>th</sup> annual conference for the Association for Educational Assessment Europe, Prague, November 9–11.
- Kreiner, S., & Christensen, K. B. (2014). Analyses of model fit and robustness. A new look at the PISA scaling model underlying ranking of countries according to reading literacy. *Psychometrika*, 79(2), 210–231.
- Lesh, R. A., & Doerr, H. M. (2003). *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, teaching, and learning*. Mahwah, NJ: Lawrence Erlbaum.
- Lesh, R. A., & Zawojewski, J. (2007). Problem solving and modeling. In F. K. J. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 763–804). Charlotte, NC: Information Age.
- Lockheed, M. E. (2015). *Why do countries participate in international large-scale assessments? The case of PISA* (Policy Research Working Paper No. 7447). Washington, DC, World Bank Group.
- Martin, M. O., Mullis, I. V. S., & Hooper, M. (2016). *Methods and procedures in TIMSS 2015*. Retrieved from <http://timssandpirls.bc.edu/publications/timss/2015-methods.html>

- Middleton, J. A., Cai, J., & Hwang, S. (2015). Why mathematics education needs large-scale research. In J. A. Middleton, J. Cai, & S. Hwang (Eds.), *Large-scale studies in mathematics education* (pp. 1–3). Switzerland: Springer.
- Ministry of Education and Research. (2012). *Oppdragsbrev 42-10*. [Description of task 42-10]. Oslo.
- Ministry of Education and Research. (2015). *Realfagstrategi - Tett på realfag. Nasjonal strategi for realfag i barnehagen og grunnsopplæringen (2015–2019)*. [Strategy for mathematics and sciences. National strategy for mathematics and sciences in kindergarden and compulsory education (2015–2019)]. Oslo, Norway: Author.
- Ministry of Education and Research. (2016a). *Slik blir den nye lærerutdanningen* [This is how the new teacher education will be] (Press release). Retrieved from <https://www.regjeringen.no/no/aktuelt/slik-blir-den-nye-larerutdanningen/id2503270/>
- Ministry of Education and Research. (2016b). *Norske 15-åringer i toppsjiktet i lesing* [Norwegian 15-year olds in the lead in reading]. (Press release) December 16, 2012. No: 107–116. Retrieved from <https://www.regjeringen.no/no/aktuelt/norske-15-aringer-i-toppsjiktet-i-lesing/id2522863/>
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016a). *TIMSS 2015 international results in mathematics*. Boston, MA: Boston College, TIMSS and PIRLS International Study Center.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016b). *TIMSS Advanced 2015 international results in advanced mathematics and physics*. Boston, MA: Boston College, TIMSS and PIRLS International Study Center.
- NDET (2017). *The education mirror*. Retrieved from <https://vimeo.com/239061112>
- Nilsen, T., Angell, C., & Grønmo, L. S. (2013). Mathematical competencies and the role of mathematics in physics education: A trend analysis of TIMSS Advanced 1995 and 2008. *Acta Didactica Norge*, 7(1), 1–21.
- Niss, M. (1996). Goals of mathematics teaching. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook on mathematics education* (pp. 11–47). Amsterdam, Netherlands: Springer.
- Niss, M. (2007). Reflections on the state of and trends in research on mathematics teaching and learning. In F. K. J. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1293–1312). Charlotte, NC: Information Age.
- Niss, M., Blum, W., & Galbraith, P. L. (2007). Introduction. In W. Blum, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education: The 14th ICMI study* (pp. 3–32). Boston, MA: Springer Science + Business Media.
- Niss, M., Bruder, R., Planas, N., Turner, R., & Villa-Ochoa, J. A. (2016). Survey team on: Conceptualisation of the role of competencies, knowing and knowledge in mathematics education research. *ZDM Mathematics Education*, 48(5), 611–632.
- Niss, M., & Højgaard, T. (2011). *Competencies and mathematical learning: Ideas and inspiration for the development of mathematics teaching and learning in Denmark* (English edition). Roskilde, Denmark: IMFUMFA.
- Niss, M., & Jablonka, E. (2014). Mathematical literacy. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 391–396). Dordrecht, Netherlands: Springer.
- Nortvedt, G. A. (2013a). Matematikk i PISA - matematikdidaktiske perspektiver [Mathematics in PISA]. In M. Kjærnsli & R. V. Olsen (Eds.), *Fortsatt en vei å gå. Norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012* [The Norwegian PISA Report] (pp. 43 - 66 ). Oslo: Universitetsforlaget.
- Nortvedt, G. A. (2013b). Resultater i matematikk [Results in mathematics]. In M. Kjærnsli & R. V. Olsen (Eds.), *Fortsatt en vei å gå. Norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012* [The Norwegian PISA Report] (pp. 67 - 95 ). Oslo: Universitetsforlaget.
- Nortvedt, G. A., Santos, L., & Pinto, J. (2016). Assessment for Learning in Norway and Portugal: the case of primary school mathematics teaching. *Assessment in education: Principles, Policy & Practice*, 23(3), 377-395.
- Nortvedt, G.A. & Pettersen, A. (2016). Matematikk [Mathematics]. In M.Kjærnsli & F. Jensen (Eds.), *Stø kurs. Norske elevers kompetanse i naturfag, matematikk og lesing i PISA 2015* [Norwegians students competence in reading, mathematics and science in PISA 2015], ( pp. 107–135). Oslo: Universitetsforlaget.



- Norwegian Directorate for Education and Training (NDET). (2015). *Curriculum for the common core subject of mathematics*. Retrieved from <http://www.udir.no/kl06/MAT1-04/Hele/?lplang=eng>
- Nusche, D., Earl, N., Maxwell, W., & Shewbridge, C. (2011). *OECD reviews of evaluation and assessment in education: Norway*. OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD). (2013a). *PISA2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. OECD Publishing.
- OECD. (2013b). *PISA 2012 results: Student performance in mathematics, reading, science* (Vol. 1). OECD Publishing.
- OECD. (2013c). *PISA 2012 results: Ready to learn. Students' engagement, drive and self-beliefs* (Vol. 3). OECD Publishing.
- OECD. (2016). *PISA 2015 results: Excellence and equity in education*. OECD Publishing.
- OECD. (2017a). *PISA: Programme for International Student Achievement*. Retrieved from <http://www.oecd.org/pisa/>
- OECD. (2017b). *PISA 2015 technical report*. Retrieved from <http://www.oecd.org/pisa/data/2015-technical-report/>
- Olsen, R. V. (2013). Undervisning i matematikk [Mathematics teaching]. In M. Kjærnsli and R. V. Olsen (Eds.), *Fortsatt en vei å gå. Norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012* (pp. 121–156). Oslo, Norway: Universitetsforlaget.
- Olsen, R. V., Hopfenbeck, T. N., & Lillejord, S. (2013). Elevenes lærings situasjon etter Kunnskapsløftet. [The conditions and situation for learning after the introduction of the Knowledge promotion]. *Norsk Pedagogisk Tidsskrift*, 97(6), 355–369.
- Pons, X. (2016). Tracing the French policy PISA debate: A policy configuration approach. *European Educational Research Journal*, 15(5), 580–597.
- Postlethwaite, T. N. (1988). *Encyclopedia of comparative education and national systems of education*. Oxford, UK: Pergamon.
- Ræder, H. G. (2017). *Matematikk i fysikkoppgaver* [Mathematics in physics tasks] (Master's thesis). Department for Teacher Education and School Research, University of Oslo, Norway.
- Sjøberg, S. (2014). PISA-syndromet. Hvordan norsk skolepolitikk blir styrt av OECD [The PISA-syndrome. How Norwegian education policy is steered by OECD]. *Nytt Norsk Tidsskrift*, 31(1), 30–43.
- Steiner-Khamsi, G. (2003). The politics of league tables. *Journal of Social Science Education*, 1, 1–6.
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Oxford, UK: Routledge.
- Sälzer, C., & Prenzel, M. (2014). Looking back at five rounds of PISA: Impacts on teaching and learning in Germany. *Solsko Polje*, 25(5/6), 53–72.
- Turner, R., Dossey, J., Blum, W., & Niss, M. (2013). Using mathematical competencies to predict item difficulty in PISA: A MEG study. In M. Prenzel, M. Koberg, K. Schöps, & S. Rönnebeck (Eds.), *Research on PISA: Research outcomes of the PISA Research Conference 2009* (pp. 23–37). Dordrecht, Netherlands: Springer.
- Tveit, S. (2014). Educational assessment in Norway. *Assessment in education: Principles, policy and practice*, 21(2), 221–237.
- Väljjarvi, J., Linnakylä, P., Kupari, P., Reinikainen, P., & Arffman, I. (2002). *The Finnish success in PISA—and some reasons behind it: PISA 2000*. Jyväskylä: Institute for Educational Research, University of Jyväskylä.
- Zhao, Y., & Meyer, H.-D. (2013). High on PISA, low on entrepreneurship? What PISA does not measure. In H.-D. Meyer and A. Benavot (Eds.), *PISA, power and policy: The emergence of global educational governance* (pp. 267–278). Oxford, UK: Symposium Books.