

Test-based accountability in the Norwegian context: exploring drivers, expectations and strategies

Marjolein K. Camphuijsen, Jorunn Møller & Guri Skedsmo

To cite this article: Marjolein K. Camphuijsen, Jorunn Møller & Guri Skedsmo (2020): Test-based accountability in the Norwegian context: exploring drivers, expectations and strategies, Journal of Education Policy, DOI: [10.1080/02680939.2020.1739337](https://doi.org/10.1080/02680939.2020.1739337)

To link to this article: <https://doi.org/10.1080/02680939.2020.1739337>



Published online: 11 Mar 2020.



Submit your article to this journal [↗](#)





View related articles [↗](#)



View Crossmark data [↗](#)



Test-based accountability in the Norwegian context: exploring drivers, expectations and strategies

Marjolein K. Camphuijsen ^a, Jorunn Møller ^b and Guri Skedsmo ^{b,c}

^aDepartment of Sociology, Autonomous University of Barcelona, Barcelona, Spain; ^bDepartment of Teacher Education and School Research, University of Oslo, Oslo, Norway; ^cInstitute for Research on Professions and Professional Learning, Schwyz University of Teacher Education, Goldau, Switzerland

ABSTRACT

This paper investigates how and why test-based accountability (TBA), a global model for education reform, began to dominate educational debates in Norway in the early 2000s, and how this policy has been operationalised and institutionalised over time. In examining the adoption and retention of TBA in Norway, we build on the cultural political economy framework, in combination with a political sociology-driven approach to policy instruments. The analysis draws on two data sources: four White Papers and 37 in-depth interviews with top-level politicians, policy-makers and stakeholders, conducted between September 2017 and February 2018. The findings indicate that ‘scandalisation’ of Norway’s below-expected PISA results and promotion of standardised testing as a neutral device contributed to the relatively abrupt adoption of national testing in the early 2000s. The increasingly dominant policy discourse equalising education quality and learning outcomes led to the institutionalisation of TBA, developed to ensure equity and quality standards in a decentralised education system. Increased visibility, benchmarking and administrative control are identified as key mechanisms in putting pressure on local actors to re-orient their behaviour. The study provides original insights into the drivers, expectations and strategies underlying TBA in a social democratic institutional regime.

ARTICLE HISTORY

Received 16 May 2019
Accepted 2 March 2020

KEYWORDS

Globalisation; education policy; standardised testing; test-based accountability; New Public Management; policy instrument approach

Introduction

Rising concerns about the performance, equity and efficiency of education systems have policy-makers around the world focussed on education reform processes. Policy principles such as decentralisation, standards and accountability are central to education debates worldwide and feature increasingly in coherent reform initiatives based on standardised assessments (Ball, Junemann, and Santori 2017; Verger, Parcerisa, and Fontdevila 2019). This means that while lower government levels and schools are given greater authority for organisational and pedagogical decisions, these actors are simultaneously held accountable for achievement of centrally defined objectives measured by standardised tests. Commonly referred to as test-based accountability (TBA) (Hamilton, Stecher, and Klein 2002), this near-universal trend is seen even in countries previously

considered ‘immune’ to this globalising phenomenon, including Norway (Verger, Parcerisa, and Fontdevila 2019).

In the early 2000s, Norwegian authorities introduced standardised testing, teacher monitoring and evaluation and an outcome-based curriculum, while also promoting further devolution of responsibilities to local education authorities and schools. These policy initiatives represented a disruptive transformation within educational institutions and school governance (Hall et al. 2015). Standardised testing and teacher monitoring and evaluation were once considered controversial and out of step with Norwegian values and traditions. The radical shift in school governance and the adoption of once disputed policy measures has received significant research attention (e.g. see Hatch 2013; Hovdenak and Stray 2015; Langfeldt, Elstad, and Hopmann 2008; Tveit 2018). Studies have highlighted how major education reforms in Norway entailed a shift from input- to output governing, and how policy changes increasingly emphasise performance monitoring, accountability and data use to improve educational practices (Skedsmo 2009).

The present paper examines why and how TBA started to dominate educational debates in Norway, and how this policy has been operationalised by different policy tools. The following research questions are addressed: (a) What is the policy trajectory of TBA in Norway? (b) What are the main drivers and rationales for Norway’s adoption of TBA? (c) What are the main policy tools to adopt and develop TBA in Norway? To examine these research questions, this study analyses key policy documents and primary data from 37 interviews with key educational reformers, legislators and stakeholders, conducted between September 2017 and February 2018.

We aim to move beyond previous research by focussing not only on the adoption of TBA in Norway, but by tracing the institutionalisation and evolution of TBA over the past 15 years (2003–2018) and during three political cabinets. In comparison with the moment of policy adoption, policy evolution often remains under-researched in policy research, despite the importance of seemingly small adaptations for how policy plays out in practice. In a similar vein, we aim to contribute to existing evidence by complementing policy document analysis, commonly used in (Norwegian) policy research, with in-depth interviews with key actors and stakeholders. The interviews proved especially fruitful to identify the often implicit or unarticulated ‘world views’ informing policy tool selection (Maroy 2015), and to gain insight into the policy process leading to the selection of formal measures. By examining the selection of policy instruments and their 15-year evolution, the present analysis aims to provide a deeper understanding of the advance of this globalising policy trend. Considering that existing evidence concerning the enactment and effects of accountability is inconclusive and contradictory (e.g. Maroy 2015; Sahlberg 2016), it seemed particularly useful to analyse the drivers of this global reform approach and its operationalisation within specific settings.

In so doing, the analysis is guided by the cultural political economy (CPE) framework (Jessop 2010), in combination with a political sociology-driven approach to policy instruments (Lascoumes and Le Galès 2007). CPE examines the drivers of the spread and adoption of global education policies, while the political sociology-driven approach to policy instruments explores how and why particular global policy configurations are developed and retained by selected policy instruments and tools. Both theoretical perspectives help investigate the mediation of the adoption and re-contextualisation of

global policy ideas by societal and cultural factors and institutions, and actors operating at different scales.

In the sections below, we first briefly discuss the different nature of globalising TBA modalities, before outlining the study's guiding theoretical approach and analytical concepts. Thereafter, we describe the Norwegian educational context and school system to provide a context for the study's findings. Then follows a description of the data and methodology. In the section thereafter, we present our main results, structured according to the three evolutionary mechanisms identified by our theoretical model (Jessop 2010). We conclude this paper with arguing that while TBA formed a key policy instrument to ensure equity and quality standards in the highly decentralised Norwegian education system, to some degree both equity and quality have been rearticulated to performance indicators based on national and international tests.

Theoretical approach and analytical concepts

In recent decades, educational agendas have been influenced by globalising ideologies and policy paradigms, accompanied by reform packages based on similar discourses and rationales (Ball 1998; Sahlberg 2016). Promoted by international organisations such as the Organisation for Economic Co-operation and Development (OECD), TBA modalities are anticipated to raise the efficiency, academic excellence and equity of education systems. Although grounded in similar principles, TBA modalities around the world differ in a) degree of regulatory tool alignment; b) consequences of accountability measures; c) the conceptions of actors; and d) the nature of mediation underlying regulation (Maroy 2015). Therefore, policy trajectories and how policies are operationalised in context are important to examine.

The present analysis is informed by Jessop's (2010) CPE framework, which forms a useful heuristic device to study processes of profound institutional transformation. CPE examines economic, political and cultural factors that interact through three evolutionary mechanisms that characterise policy adoption and change: variation, selection and retention. *Variation* is triggered by a perceived need to revisit existing policies or practices, prompting *selection* of suitable and agreed policy instruments. *Retention* institutionalises new policy proposals and instruments in the context of existing regulations and practices. Both symbolic and material factors influence the process of policy change (Jessop 2010). These three evolutionary mechanisms provide useful analytical tools for systematic examination of the contingencies, events and actors involved in policy change, and identification of mechanisms inducing or restraining institutional change. Using this framework, we aim to contribute to a holistic explanation of why and how TBA has been adopted and developed in Norway, by identifying drivers of change in the political, economic and cultural domains.

Using a political sociology-driven account of policy instruments, we explore the operationalisation and institutionalisation of TBA over time, examining the main policy tools to adopt and develop TBA in Norway. This approach criticizes functionalist approaches that view public policy as pragmatic, policy instruments as 'natural' and at policy-makers' disposal and policy adoption as informed by instruments' proven effectiveness. In so doing, functionalist approaches oversimplify processes of instrumentation by neglecting economic and political factors. In contrast, the political sociology-driven

approach to policy instruments holds that the process of instrumentation must be further problematised (Lascoumes and Le Galès 2007). Here, public policy instruments are seen as 'bearers of values, fuelled by one interpretation of the social and by precise notions of the mode of regulation envisaged'. We seek the political and economic stakes of instrumentation by tracing 'power relations associated to instruments and issues of legitimacy, politicisation, or de-politicisation dynamics associated with different policy instruments' (Lascoumes and Le Galès 2007, 4).

Combining political sociology and CPE lenses, we investigate how the adoption, development and operationalisation of these policy instruments and tools are contingent on and mediated by institutional and contextual factors. As Verger, Fontdevila, and Zancajo (2016) showed, the economic, political, institutional and cultural factors that provoke, condition and legitimise domestic policy formation may differ widely. For example, factors such as a country's economic environment (Lenschow, Liefferink, and Veenman 2005) or an economic crisis or recession (Ball 1990) can create pressure for reform and shape the perception of policy instruments as legitimate and economically feasible. Political factors, motivations and interests can also influence the policy problems that education reforms will address, as will the policy instruments developed for that end. Cultural factors including 'the semiotic and meaning-making dimension of education policy processes', and social values or public opinions can similarly shape the adoption and retention of particular policy ideas and instruments (Verger, Fontdevila, and Zancajo 2016). Finally, our approach acknowledges that institutional legacies often mediate rather than provoke policy change. Public administration traditions, political institutions and regulatory frameworks shape policy-makers' views about new policies and define the institutional boundaries of policy formation.

The Norwegian educational context

The Norwegian social-democratic welfare state (Esping-Andersen 1990) has high levels of public social expenditure and direct provision of public services by state and local government. Governance responsibilities are divided between municipalities, counties and the state. Since the late 1980s, neoliberal thinking and new managerial ideas such as New Public Management (NPM) have influenced public sector reforms in Norway, promoting decentralisation, deregulation, horizontal specialisation and management by objectives (Christensen and Læg Reid 2011).

Education is predominantly publicly provided and financed and aims to promote values such as equity, solidarity, social justice and democracy. The comprehensive school model seeks equal opportunity for people of all genders, socio-economic or ethnic backgrounds, or geographical locations (Blossing, Imsen, and Moos 2014). While 82% of Norway's population lives in urban areas (Statistics Norway 2019a), many municipalities and schools are small. School choice is limited, especially for compulsory education, although local exceptions are found (Haugen 2019). The establishment of private schools is strictly regulated, with only 4% of the school-aged population enrolled in private primary or lower-secondary schools (Statistics Norway 2019b). Norway's national curriculum guidelines were traditionally formulated with broad aims, allowing significant interpretive leeway for teachers, whom were generally trusted to manage their own practices and enjoyed significant autonomy (Mausethagen 2013). Historically, investing in teacher education was a key strategy in guaranteeing education quality (Werler and Sivesind 2007).

In the 1990s, discussions first started regarding quality assessment measures, such as national testing. In 1988, the OECD had published a country review of the Norwegian education system, which questioned how central authorities could monitor and direct such a decentralised education sector, especially when lacking systematic data on education quality and outcomes. While several WPs in the 1990s addressed OECD concerns, and an evaluation system in a decentralised context was considered, there was little concrete action. There existed political controversy surrounding such a system, and lack of support from key political actors, parliamentarians and the main teacher's union (Møller and Skedsmo 2013). However, at the turn of the millennium, the public and political debate on education intensified, contributing to a strong push for political action.

Data and methodology

Our analysis draws on two principal data sources. First, we analysed four key White Papers (WPs) published by the Ministry of Education and Research (MER): WP 30 *Culture for Learning* ([2003] 2004), WP 31 *Quality in Schools* ([2007] 2008); WP 20 *On the Right Track* ([2012] 2013); and WP 21 *Eager to Learn* ([2016] 2017). We decided to use WPs as a key data source as these documents form the prime source of political decision-making, providing insight into the official justification behind the need to revisit existing policies or practices, and the selection and design of policy tools. Similarly, WPs form a key source of information in tracing the policy trajectory and gaining an understanding of the broader policy context. We selected the four above-named WPs after an initial screening of all WPs published by the MER over the past two decades. It was found that these four WPs represented key moments in TBA's trajectory. WP 30 ([2003] 2004), prepared under right-wing governance, proposed far-going policy changes that would lead to the formal introduction of TBA in Norway. WP 31 ([2007] 2008), prepared when a red-green coalition returned to office, proposed adjustments and additions to TBA, as did WP 20 ([2012] 2013). Finally, WP 21 ([2016] 2017) presents changes undertaken after a right-wing coalition again took office in 2013. Key policy documents published before 2003 (e.g. Moe Commission 1997; MER 2002), as well as legal documents, evaluation reports (e.g. Haug 2003), press releases, and the memoirs of Helge Ole Bergesen (former-State Secretary of Education, 2001–2005) provided background and contextual information for the analysis.

Second, we drew on data from 37 interviews¹ with 40 educational stakeholders and key informants, conducted between September 2017 and February 2018. The interviews served to gain a deeper understanding of the policy process 'behind the scenes', amongst others to identify prominent policy networks and to explore power dynamics and issues related to legitimacy and (de-)politization (Lascoumes and Le Galès 2007). Moreover, the interviews allowed the refinement of regulation theories behind policy tool selection (Maroy 2015), which often remain implicit or unarticulated in formal policy documents.

Purposive and snowball sampling was used to select interview participants. Based on secondary sources and a-priori knowledge of the researchers, a first list of potential interviewees contained participants who were key players in education policy design, formation and/or implementation over the past three decades, experts on recent changes in education policy or influencers in the education debates. To avoid the risk of selection bias, the sample was expanded and refined during the fieldwork process by asking interviewees to identify other potentially relevant participants. Table 1 provides an overview of all participants.

Table 1. Overview of interview participants.

Actors	Number interviewed
Top-level politicians	4
Policy-makers from the Royal Ministry of Education and Research or the Norwegian Directorate for Education and Training	7
Members of government advisory committees	4
Union leaders; members of The Norwegian Association of Local and Regional Authorities (KS)	11
Academics, experts, evaluators	12
Active participants in the public debate on education	2

A semi-structured script guided the 60-minute interviews. Amongst other themes, the interviews addressed policy problem(s) intended to be addressed; perceived causes and relevance of these problem(s); promotion of particular policy solutions; rationales and expectations informing policy tool selection; the process of policy design, institutionalisation and policy evolution; administrative traditions and; social values and public opinion. Moreover, emphasis was placed on the role of key events and actors throughout the policy process, as well as the latter's motivations, strategies and ideational influences (see Fontdevila 2019). The interviews were audio-recorded and transcribed verbatim.

We combined inductive and deductive approaches during data analysis. Using ATLAS.ti software, three researchers independently performed a first reading of the raw data material, to identify frequent, dominant and significant themes and categories. Based on this first reading, as well as Jessop's (2010) CPE framework, we developed a codebook, which was used to code all data material structurally during a second reading. This way, we attempted to classify all data material according to the research questions (Saldaña 2009). Based on this structural coding exercise, six main groups of codes were identified: (1) Attributes of the interviewee; (2) Subjective perception of TBA; (3) Policy process: diagnosis; (4) Policy process: policy formation and evolution; (5) Policy process: balance of forces, actors; and (6) Knowledge mobilization. Subsequently, we performed a third, more in-depth reading of all segmented data material, both within and across the identified themes (e.g. MacQueen et al. 1998). During this third reading, we complemented and refined the six macro codes with a set of analytic codes, corresponding to key concepts, actors and mechanisms (Fontdevila 2019).

Findings

In the sections below, we have structured the presentation of our findings according to CPE's three evolutionary mechanisms (Jessop 2010). This division allows us to systematically examine our three research questions. By means of the three evolutionary mechanisms, the policy trajectory of TBA is outlined. During the sections 'variation' and 'selection', the main drivers of the perceived need for policy change and the promotion of TBA as a suitable and desirable instrument are discussed. During the section 'retention', the main policy tools by which TBA has been operationalised are presented. While all reviewed data material has informed the analysis, in this section we have used policy document citations and selected quotes by interviewees² to illustrate the ways in which policy changes have been typically described, explained and justified.

Variation: from ‘best school in the world’ to ‘school loser’

Following the results of Norwegian students in the inaugural Programme for International Student Assessment (PISA) 2000, student academic achievement began to dominate the education debate. Norway ended up in 13th place (out of 31) in the international rankings for reading and science, and 17th in mathematics. In reading, the gap between high- and low-performing pupils was greater than in most European countries. The Conservative Minister of Education and Research of the Bondevik II government, Kristin Clemet (2001–2005), who had taken office only six weeks before the PISA findings were published, made a statement reported by all the country’s major media outlets: ‘This is disappointing, almost like coming home from a Winter Olympics without a single Norwegian medal. And this time we cannot blame the Finns for being doped.’ (Dagbladet, 5 December 2001).

Norway’s self-image as ‘the best school in the world’³ was crushed, especially by the large disparities in pupils’ educational outcomes and the finding that a significant percentage of pupils left compulsory education without basic competencies such as reading (17%). For decades, the comprehensive school model had aimed to provide a school for all, but now it seemed that social and geographic inequalities were reproduced in schools. The subsequent release of other national and international studies reporting similar results (e.g. PIRLS 2001; PISA 2003; TIMSS 2003) contributed to the consensus that Norwegian education had ‘a problem’.

When PISA came, I remember that a lot of people said “this is wrong, this must be wrong, this is not . . . it does not fit with the Norwegian system, with our curricula”. When the next PISA study came three years later, everybody said “Yes, we have a problem, we have to do something”. The reason why this had changed during these three years was that there was a lot of other information and research, also national research (. . .), [which] all told us roughly the same. – Interview with Kristin Clemet, Minister of Education and Research (2001–2005), Conservative Party

Public and political education debates sought explanations for the below-expected results. In his memoir *The Battle for the Knowledge School*, Helge Ole Bergesen, who was State Secretary of Education during Clemet’s time in office, summarised the problems that afflicted Norwegian schools: ‘Norwegian schools seem to have entered a vicious circle where lack of clarity, lack of competence, low motivation, weak leadership, uncertainty about responsibility and lack of knowledge of results are mutually reinforcing’ (Bergesen 2006, 66; authors’ translation).

Criticising an era of school governance, it was argued that schools were subject to excessive input regulations, with little awareness of, or responsibility for, results due to the lack of systematic data on pupils’ learning outcomes.⁴ The PISA results shattered the belief that the desired results would be obtained through the rule of central authorities (Bergesen 2006, 86–87). Discourse emerged characterising Norwegian schools as ‘too soft’ and ‘playful’, with limited attention to basic competencies such as reading, writing and numeracy. An evaluation of Curriculum Reform (1997) (Haug 2003) cited another policy problem: education was geared towards the ‘average student’ and was not sufficiently adapted to the needs of the individual, an ideal of Norwegian schools since the 1930s. This finding was considered unacceptable, especially in an increasingly diverse society. As would later be argued, ‘All students are equal, but none of them are alike (. . .) If we treat everyone alike, we create greater inequality’ (MER [2003] 2004, 4; authors’ translation).

The agenda-setting power of Clemet and her administration was significant. Since the early 1990s, student achievement was a topic of political discussion. However, in the new millennium, learning outcomes dominated debates as the main indicator of education quality. The perceived learning crisis, to which the media contributed significantly, provided a strategic opportunity for the liberal-conservative government coalition to advance a long-desired policy reform:

For us, who had just taken over political leadership of the Ministry of Education and Research, the PISA results were a “flying start”. Admittedly, the Conservative Party had complained for a long time about quality problems in Norwegian schools. (...). With the PISA survey, the climate of debate changed abruptly, radically and irrevocably. (Bergesen 2006, 42; authors’ translation)

Selection: ‘the most research-based reform’

Policy design process

In PISA’s aftermath, a variety of contributors proposed policy solutions. Concluding that schools needed modernisation, Kristin Clemet initiated ‘The school knows best’ project to determine ways to improve Norwegian schools (MER 2002). In this project, experts from outside the MER examined international policy research and policy experiences. The project report reviews, amongst other measures, competition and freedom of choice as tools for promoting student achievement and maintaining low government costs. Since the project report was criticised both within and outside the MER, the members published a second draft that set a milder tone. The following was their overarching conclusion:

We must decentralise responsibility, improve quality control and increase users’ empowerment. The school should be controlled from below, not from above, within nationally targeted goals ... We will mobilise for greater creativity and dedication by giving the freedom to take responsibility. (MER 2002, 1; authors’ translation)

The Quality Commission, established in December 2001 by the Stoltenberg I government, played an important role in selecting policy instruments to modernise and enhance Norway’s education system.⁵ Drawing on national and international policy research and expertise (e.g. Granheim and Lundgren 1990; OECD [1988] 1989; UNESCO 1990, 1995) and earlier policy proposals (e.g. the Moe Commission 1997), the Commission delivered its first deliberation in June 2002, outlining the framework of a National Quality Assessment System (NQAS). As part of this system, they proposed the establishment of annual national standardised tests and a web portal publishing the results. The publication of the results, so increasing schools’ ‘visibility’ and facilitating comparison and benchmarking, was seen as a significant pressure mechanism for motivating actors and eliciting improvement:

I think many [of us] believed that it is incentive enough in itself that everybody knows that you are doing badly. As to performance-based pay, I cannot remember that we were talking about that. – *Interview with member of Quality Commission*

After the ‘PISA shock’, the introduction of national tests was considered necessary by an increasing number of key actors, but controversy persisted regarding the web portal that published school results. While some felt strongly that visibility, comparison, benchmarking and competition would improve education in the long run, this was far from the

general view. Nevertheless, the framing of the policy proposal, and the decision-making process, enabled the adoption of both measures in the early 2000s. With regards to the national tests, they had been promoted and framed as an information-gathering tool, allowing for local and individual adaptations following data-based decision-making. This implied that the tests were largely described as a neutral device (cf. Lascoumes and Le Galès 2007), benefiting teachers, pupils, and parents, while the underlying agenda of monitoring, evaluation and control remained masked and under-communicated. Scepticism of teachers about the need for these data and criticisms from test developers about the challenges of the double-purpose test were ignored or dismissed as ideological rather than objective. With regards to the web-portal, it was argued that once data were collected, the results could not be kept secret, pointing to the existing Act on Public Information.

Despite that several actors expressed concern about ‘the possibility that the portal may contribute to ranking’ and how the introduction of tests may guide teaching (MER [2002] 2003), the decision-making process was characterised by a strong sense of urgency. Interviewees contended that government officials’ reaction reflected what Steiner-Khamsi (2003) has described as ‘scandalisation’ in pushing for the contested policy instrument’s adoption. This reaction contributed to the agreement to adopt national tests in reading, writing, numeracy and English and the online web portal during deliberations on the 2003 state budget:

Kristin Clemet used this presentation of the [PISA] test results to create a wave of “this is exceptional, we must act!” So, the decision to implement national tests was never discussed in principle; it was never discussed as a case – it just came as an amendment to the budget proposition [because] the need for action was perceived to be so great. – *Interview with former MER policy-maker*

The modernisation project’s conclusions, the 2003 Quality Commission’s final proposal, and the evaluation reports of Curriculum Reform 94 (upper-secondary education) and 97 (compulsory education) laid the foundation for the 2004 WP 30 (*Culture for Learning*) issued by the MER. Reflecting many of the principles dominating educational reforms worldwide, this WP proposed systematic changes in school governance to improve the Norwegian school system. The proposed measures built on five key principles: a) clear national goals; b) knowledge of results; c) clear responsibility placement; d) greater local freedom of action; and e) a good support and guidance system (MER [2003] 2004, 25). The NQAS, containing both traditional and new assessment tools, became a key element in the proposed reform.

More concretely, local government and school authorities would be given more decision-making power, including greater local freedom to decide on educational content and working methods, more flexible rules around class and time distribution, and the transfer of negotiations on teachers’ salaries, working hours and employment conditions from state to regional and local authorities. The national tests were presented as an information-gathering tool for the schools, but would also be a key control mechanism for central and local authorities, enabling them to monitor schools’ (and municipalities’) efforts to achieve centrally defined goals. Finally, municipalities would become obliged to establish quality assessment systems to measure and follow up on school performance⁶: teachers, principals and municipal superintendents would be held formally accountable for students’ performance in standardised tests and other quality measures.

Key drivers behind institutional change

Various drivers contributed to policymakers' perceptions of TBA as a necessary and suitable policy instrument. First, interpretations of global economic trends and societal changes paved the way for TBA. In a knowledge economy, a country's educational achievement is considered fundamental to its economic potential and competitiveness, thus competencies such as literacy and numeracy are perceived as essential for individual and nation-state success (MER [2003] 2004). Accordingly, school reforms were sought to ensure that basic competencies would be central to educational content, and acquired by all pupils by the end of compulsory education. Moreover, in an increasingly diverse society, individual and local adaptation was considered ever more important (MER [2003] 2004). TBA became a key instrument for central authorities to grant greater freedom to local actors while ensuring compliance with government priorities and goals, e.g. regarding pupils' development of basic competencies.

At the same time, the promotion of TBA cannot be seen in isolation from broader changes in public sector governance. TBA is compatible with the NPM policy paradigm, which was formally introduced in Norway in the late 1980s to restructure and modernise public administration. A second wave of NPM reforms subsequently addressed issues of coordination across administrative levels, which were partly caused by public sector fragmentation accompanying the first wave of NPM reforms. NPM allowed government officials to regain oversight over and responsibility for public services, and hence to ensure equity in educational outcomes across social groups. Moreover, both reform waves followed a belief in the need to steer the public sector by means of a performance-oriented culture, in order to raise its efficiency and effectiveness.

This belief was reinforced by an emerging scientific base, to which international organisations such as the OECD contributed, which argued in favour of outcome-based management, accountability and assessment as key measures to modernise and raise the performance of education systems. As such, in the eyes of policymakers, TBA had empirical credibility. Interviewees who were more directly engaged with the adoption and design of reform measures often referred to the OECD reports as an important source of evidence. Nonetheless, the policy document analysis reveals that in addition to international policy research, also national research formed a key source of evidence in reform proposals (e.g. Granheim and Lundgren 1990; Haug 2003).

With the push for evidence-based policymaking, such research documents, as well as advice from external experts, researchers and consultants play an increasingly important role during policy design processes. With regards to the promotion of TBA as a policy solution, expert advice and research documents seem to have served as both an important source of inspiration, as well as justification. In turn, ideological concerns and critique were largely silenced. Nonetheless, despite a tendency towards scientification of policy processes, Norway's political institutions and social-democratic welfare ideologies mediated the selection and retention of TBA tools.

In this light, a final driver in the adoption and institutionalisation of TBA relates to accountability as an 'empty vessel' policy that can be adopted to serve a diverse set of goals (Steiner-Khamisi 2016). Unlike early adopters of TBA, where this policy measure often served to promote market-based reforms (Verger, Parcerisa, and Fontdevila 2019), Norway's TBA was adopted and developed to ensure equity and quality standards in a decentralised education system. The presumed ability of TBA to ensure a basic standard

for all, thus equalising disadvantage, contributed to support beyond party political lines. Similarly, adjustments made soon after the adoption of TBA, as explained below, led to broader acceptance of the formerly contested measure, and its eventual institutionalisation in the context of existing regulations and practices.

Retention: 'inclusion as a basic ethical warrant'

WP 30 (*Culture for Learning*) received unanimous approval by the Norwegian Parliament, laying the foundation for the Knowledge Promotion reform. However, before the Knowledge Promotion could be implemented, there was a shift in government. The liberal-conservative coalition of Bondevik's Second Cabinet was replaced by Stoltenberg's Second Cabinet, a red-green coalition comprised of the Labour Party, the Socialist Left Party, and the Center Party. Some expected the new government coalition to halt the reform, given the significant outcries from teachers, students and even municipalities to postpone it. However, as the new Minister of Education and Research (Øystein Djupedal, member of the Socialist Left Party) explained, they decided to adhere to the principles and the timetable. They shared the view that there was a quality crisis in Norwegian schools and had accepted the logic behind the reform, but also feared the political consequences of abandoning or postponing it.

I do not think the government would have done it anyway. I do not think Jens Stoltenberg [the Prime Minister] would have said "Yes" if I went in and said "the Knowledge Promotion reform is not sufficiently prepared". He would say that we would look soft on knowledge and skills and that we do not appreciate quality. – *Interview with Øystein Djupedal, Minister of Education and Research (2005–2007), Socialist Left Party*

One change in the Knowledge Promotion reform implementation related to the administration of national tests and the 'School Portal', where results were published. The rushed implementation of the 2004 national tests had contributed to validity problems and destructive evaluation reports. This prompted criticism, even from test proponents, and country-wide boycotts of the test by upper secondary students and the School Student Union.

The School Student Union arranged the boycott against the national test in the spring of 2004 and in January 2005. The basis for the boycott was, on the one hand, opposition to publication of test results and the ideology of competition that this expressed, and secondly, the lack of opportunity to share our views during the planning of the test system. When the authorities neglect students' opinions and do not allow us to advance our criticisms, warnings and recommendations, it is legitimate to use other forms of influence. (Hølleland 2007, 37; authors' translation)

The State Secretary to the Minister of Education and Research (2005–2013) admitted that 'national tests were one of the most difficult issues' during the negotiations between the three parties on the 2005 political platform. Eventually, they agreed to pause the tests in order to improve their quality and negotiate the conditions of administration. The national tests were reintroduced in 2007, no longer administered during the spring of the 4th, 7th, 10th and 11th grades but instead during the autumn of the 5th, 8th and 9th grades.⁷ This decision changed how TBA played out:

So, while the previous tests could be used to test and follow up actual teachers, the new tests [were] intended more to discuss the quality of the school rather than the individual teacher. (...). The most important thing with these tests was to measure school quality on a municipal level, and to provide the municipalities with a good tool to measure quality, and to be able to talk about quality differences between schools, not so much focus on individual teachers but on the school as system. This meant [it] took a lot of pressure away from the actual teachers. – *Interview with State Secretary to Minister of Education and Research (2005–2013)*

Moreover, the new government coalition saw publication of school results on the ‘School Portal’ as a means of scapegoating teachers, school leaders and local authorities rather than empowering them to improve results. This shaped the decision to no longer publish national test scores at the school level, but only at municipal, county and national levels.⁸ At the same time, belief persisted that national test data was necessary for school development, shifting emphasis to the data’s utility in internal discussion, learning and quality improvement. As such, the decision to proceed with national testing signals acceptance of the policy tool while rejecting its initial use for school comparison and competition.

Nonetheless, while political rhetoric highlighted a belief in the professional responsibility of local actors and educators to use test results to improve schools, control mechanisms have been layered onto one another over time. During eight years of a red-green government coalition (2005–2013), TBA steered towards administrative or bureaucratic forms of accountability, where the consequences of national test scores are determined by national and local authorities rather than by market and competitive pressures. This attempt to weaken the role of external actors, such as the media, in holding schools accountable for results instead caused even greater pressure for local authorities to be liable in educational outcomes.

For example, in WP 31 (*Quality in Schools*) (MER [\[2007\]](#) 2008) a new administrative policy tool requires municipal councils to produce yearly assessments of school academic levels and learning environments and, where necessary, to develop improvement plans. The rationale for intensifying the administrative consequences of school performance was largely, but not solely, based on the perceived detrimental effects of market accountability. According to two policy-makers with key roles in policy formation over the last two decades, the additional pressure on municipalities and schools also reflected parents’ failure to hold schools accountable for results:

All these various [assessments], national tests, pupil surveys, value-added indicators require a professional school administration to make use of them. We thought that if parents saw the results of the national tests and of the pupil survey, they might perhaps ... come to the school and say “OK, look here, we have a problem, we are not performing as well as we could or should in the national tests and the important things that they measure; we do not perform as well as we should in terms of the student learning environment”, and so on. But my impression is that parents seldom do this, so the system is very dependent on a good, professional and forward-looking school administration. – *Interview with MER policy-maker*

The Conservative Party returned to office in 2013, this time with the right-wing Progress Party. The new coalition reintroduced the publication of schools’ national test scores on the government’s ‘School Portal’. In addition, they also introduced a new accountability tool as another external device to put pressure on local actors: a quality standard in key areas of education, using the number of students achieving a minimum level of

performance as a key indicator (MER [2016] 2017). Municipalities and counties averaging below a basic level would receive mandatory support and follow-up. This measure was justified as a means of giving direction and conveying the expectations for the role of local and regional authorities.

In addition, in 2016, value-added indicators for compulsory education were published for the first time at national level. The valuation of a school, municipality or county's contribution to student learning was considered a valuable supplement to information about results from national tests and examinations and a key tool for school development. According to interviewees, municipal superintendents have already welcomed the value-added model as a new accountability tool in their annual talk with the leadership teams of each school.

Finally, adoption of the NQAS and the expectation that municipal superintendents take responsibility for monitoring and following up assessment results afforded opportunities to operationalise and expand the specific governing model. While central policy rhetoric, to a certain extent, expresses trust in the professional responsibility of local and school actors to engage in data-based quality improvement without incentives or sanctions, interpretations differ at the local level. Various municipal superintendents have decided to enact quality assessment systems based on detailed and in advance specified performance indicators, risk assessment, publication of school test scores and performance contracts. In addition, in some large cities, municipal superintendents use merit-based pay, which is part of local salary negotiations, to reward principals who can demonstrate good performance.

Discussion and conclusion

This paper has addressed the following research questions: (a) What is the policy trajectory of TBA in Norway? (b) What are the main drivers and rationales for Norway's adoption of TBA? (c) What are the main policy tools to adopt and develop TBA in Norway? Our analysis portrays that TBA formed a key policy instrument to modernise and raise the performance and equity of the Norwegian education system. TBA replaced a steering tradition based on prescription and intervention, by allowing government officials to steer a highly decentralised education system from a distance, by means of outcome measures, visibility, comparison and accountability.

With regards to the policy trajectory, the analysis reveals that a window of opportunity for major educational reform opened when the arrival of a right-wing government coincided with the publication of the first PISA results. The below-expected results served as an 'external authority' for already existing ideas and policy initiatives (Steiner-Khamsi 2003), and contributed to justify the need to revisit existing policies and practices, and to legitimise the advance of formerly controversial reform measures. In addition to the 'scandalisation' of Norway's PISA results, to which also the media contributed (see Elstad 2012), the promotion of standardised testing as a neutral device played a key role in the abrupt adoption of national testing in the early 2000s.

Meanwhile, the increasingly dominant policy discourse equalising education quality with academic learning outcomes measured by standardised tests meant that national testing was here to stay. Multiple managerial devices have been introduced to address achievement gaps across different social and cultural groups and, to some degree, the

public debate about equity and quality has been rearticulated to performance indicators based on national and international tests. Politicians across the political spectrum have referred to the PISA, as well as national test results, as key measures to judge the education system's quality. As such, while WP 30 (MER [2003] 2004) introduced a transformation in school governance, policy evolution ever since has largely remained characterised by continuity. While adjustments and adaptations have been made during different political cabinets, in particular to downplay competition dynamics and to promote learning, these changes remain characterised by a belief in national testing as a valid measure to assess education quality, promote data-based decision-making and hold key actors accountable.

Beyond the general acceptance of a 'quality crisis', which proved a strong catalyst for change, our analysis highlights key drivers and rationales behind Norway's adoption and retention of TBA. In particular the interpretation of global economic trends and societal changes, TBA's compatibility with the NPM policy paradigm, and the assumed empirical credibility of TBA form key explanatory factors in this regard. Regarding the latter, the analysis reveals the importance of research documents and expert advice, as both a source of inspiration and legitimisation of reform proposals. Interestingly, while providing policy documents with 'scientific rationality', Steiner-Khamsi, Karseth, and Baek (2019) show that 'expert advice' is referred to in a selective manner in policy documents, highlighting the importance of critically examining the use of research in policymaking.

Nonetheless, despite that national and international research inspired and legitimised reform proposals, Norway's political institutions and social-democratic welfare ideologies played a key role during the adaptation and evolution of TBA, contributing to its broad acceptance over time. Analysing this semiotically, the fact that global education policy principles such as accountability commonly operate as 'empty vessels' is important in this regard. In contrast to early adopters' use of TBA to promote market-based reforms (Verger, Parcerisa, and Fontdevila 2019), Norway has developed TBA to ensure equity and quality standards in a decentralised education setting.

In this light, our analysis illuminates how TBA has been operationalised by a range of policy tools. Relying on the conception of an actor guided less by self-interest than by social obligations, the Norwegian way of TBA seeks to orient the behaviour of local actors by a combination of external devices that pressure actors, with internal measures designed to mobilise internal feelings of responsibility (Maroy 2015). This finding aligns to the conclusion drawn by Hatch (2013) that the Norwegian accountability system is characterised by a tension between answerability for short-term goals and responsibility for broader purposes. The institutionalisation of TBA furthermore has a clear 'Norwegian touch' in the limited 'hard' consequences attached to school performance, which conflicts with professional values and administrative traditions (Skedsmo 2011). At the same time, despite the political rhetoric implying trust in educators' professional judgement and responsibility, school performance is increasingly monitored, controlled and made visible by administrative mechanisms layered on top of one another over time.

Finally, Norway's TBA is a good illustration of how policy instruments exist 'independent of the decisions that created them' (Kassim and Le Galès 2010, 11), taking forms and generating outcomes that may contradict or extend beyond initial policy goals (Le Galès 2010). In contrast to national policy rhetoric, some municipal superintendents have attempted to re-orient the behaviour of local actors by increasing controls and raising the stakes for school performance. These efforts build on the assumption that

many aspects of teaching and learning can be controlled and documented, and that holding school leaders and teachers accountable for students' results will make them more efficient and effective. These local differences highlight the importance of examining the ways in which municipal superintendents act as 'brokers' of predefined goals, in particular when studying school responses to data-use and accountability demands (Prøitz, Mausethagen, and Skedsmo 2019)

To sum up, by tracing the adoption, retention and adaptation of TBA in Norway, this study provides interesting insights into the complex, diverse and hidden drivers of TBA reforms in a social-democratic welfare state. Beyond the study's immediate aims, we identified signs of contradiction and paradox in the enactment of the described reforms. The combined effect of delegating responsibility and decision-making power downward while raising pressure has created tensions over time. As some interviewees argued, although the original aim was to encourage individual and local adaptation and creativity, stricter local authority control and supervision has sometimes constrained teacher autonomy and promoted standardised teaching methods. The ongoing struggle of teachers and school leaders to find meaningful ways of integrating national test data (Mausethagen, Prøitz, and Skedsmo 2018), moreover, confirms that the challenges associated with assigning a double purpose to a single test have not been resolved (see also Tveit 2018).

Paradoxically, although the Norwegian curriculum guidelines provide a broad framework allowing autonomy of local schools, the scope is narrowed by the municipal use of national standardised tests to hold schools accountable. While policy tools such as national tests may appear neutral, they carry values and meaning, thereby foregrounding certain aspects of teaching and learning while constraining others. For greater understanding of the circumstances and mechanisms contributing to the operation and policy outcomes of TBA modalities, further research is needed to examine how the reforms are interpreted and put into practice on the ground.

Notes

1. There were 34 individual interviews and 3 group interviews with two participants in each.
2. Indirect identification of interviewees is prevented by not referring to a participant's specific role. Interviewees for whom indirect identification is unavoidable, such as previous Ministers, gave explicit consent to be quoted and named.
3. This conclusion, drawn in 1975 by Bjartmar Gjerde, Minister of Church and Education Affairs from 1973 to 1976, reflected a widely shared belief.
4. That is, while much was known concerning what was going into schools in terms of resources, teachers, pupils, buildings, computer equipment etc., barely anything, except final exam scores, was known about what was coming out.
5. Upon establishment, the Commission had been mandated to 'evaluate the content, quality, and organisation of basic education'. Under the Bondevik II government, 8 members were added to the Commission, while the mandate was extended and now included the request to develop a quality assessment system for basic education.
6. As a result of NPM reforms introduced during the late 1980s, many municipalities had already established quality assessment systems.
7. This meant that tests were not reintroduced at the upper-secondary level but only in compulsory education.
8. Nevertheless, results could still be requested by anyone under the Freedom of Public Information Act.

Acknowledgments

We would like to thank members of the research project ‘Reforming Schools Globally: A Multi-Scalar Analysis of Autonomy and Accountability in the Education Sector (REFORMED)’ at the Autonomous University of Barcelona and members of the research group ‘Curriculum Studies, Leadership and Educational Governance (CLEG)’ at the University of Oslo for their comments on earlier drafts of the manuscript. We also express our gratitude to the anonymous reviewers for valuable feedback during the review process.

Disclosure Statement

No potential conflict of interest was reported by the authors

Funding

This work was supported by the European Research Council under Grant [680172].

Notes on contributors

Marjolein K. Camphuijsen is a doctoral candidate in Sociology at the Autonomous University of Barcelona, Spain. Her research focus is on the political sociology of education, global education policy, policy enactment, and the effects of public policy on social and educational inequalities.

Jorunn Møller is an emeritus professor at University of Oslo, the Department of Teacher Education and School Research. Her professional interests are in the areas of educational leadership and governance, reform policies and school accountability.

Guri Skedsmo is a professor at Schwyz University of Teacher Education in Switzerland and associate professor at University of Oslo in Norway. Her professional interests are in areas of educational governance and leadership, quality assessment and accountability.

ORCID

Marjolein K. Camphuijsen  <http://orcid.org/0000-0002-7824-7674>

Jorunn Møller  <http://orcid.org/0000-0002-4159-9131>

Guri Skedsmo  <http://orcid.org/0000-0001-7151-8630>

References

- Ball, S. J. 1990. *Politics and Policy-making in Education: Explorations in Policy Sociology*. London: Routledge.
- Ball, S. J. 1998. “Big Policies/Small World: An Introduction to International Perspectives in Education Policy.” *Comparative Education* 34 (2): 119–130. doi:10.1080/03050069828225.
- Ball, S. J., C. Junemann, and D. Santori. 2017. *Edu-net: Globalisation and Education Policy Mobility*. London: Routledge.
- Bergesen, H. O. 2006. *Kampen om kunnskapsskolen* [The Battle for the Knowledge School]. Oslo: Universitetsforlaget.
- Blossing, U., G. Imsen, and L. Moos. 2014. *The Nordic Model. A ‘School for All’ Encounters Neo-Liberal Policy, Policy Implications of Research in Education*. Vol. 1. Dordrecht: Springer Series.

- Christensen, T., and P. Lægrend. 2011. "Democracy and Administrative Policy: Contrasting Elements of New Public Management (NPM) and post-NPM." *European Political Science Review* 3 (1): 125–146. doi:[10.1017/S1755773910000299](https://doi.org/10.1017/S1755773910000299).
- Elstad, E. 2012. "PISA Debates and Blame Management among the Norwegian Educational Authorities: Press Coverage and Debate Intensity in the Newspapers." *Problems of Education in the 21st Century* 48: 10–22.
- Esping-Andersen, G. 1990. *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.
- Fontdevila, C. 2019. "Researching the Adoption of School Autonomy with Accountability Reforms: A Methodological Note on Country Case Studies." *REFORMED Methodological Papers* 1. doi:[10.5281/zenodo.3557302](https://doi.org/10.5281/zenodo.3557302).
- Granheim, and S. Lundgren 1990. "Goal Management and Evaluation in the Norwegian School." *Final report EMIL project*. doi:[10.1099/00221287-136-2-327](https://doi.org/10.1099/00221287-136-2-327).
- Hall, D., E. Grimaldi, H. Gunter, J. Møller, R. Serpieri, and G. Skedsmo. 2015. "Educational Reform and Modernisation in Europe: The Role of National Contexts in Mediating the New Public Management." *European Educational Research Journal* 14 (6): 487–507. doi:[10.1177/1474904115615357](https://doi.org/10.1177/1474904115615357).
- Hamilton, L. S., B. M. Stecher, and S. P. Klein. 2002. *Making Sense of TBA in Education*. Santa Monica, CA: Rand Corporation.
- Hatch, T. 2013. "Beneath the Surface of Accountability: Answerability, Responsibility and Capacity-building in Recent Educational Reforms in Norway." *Journal of Educational Change* 14 (1): 113–138. doi:[10.1007/s10833-012-9206-1](https://doi.org/10.1007/s10833-012-9206-1).
- Haug, P. 2003. *Evaluering av Reform 97: Sluttrapport frå styret for Program for evaluering av Reform 97* [Evaluation of Reform 97: Final Report from the Board of the Program for Evaluation of Reform 97]. Oslo: Norges forskningsråd.
- Haugen, C. R. 2019. "Teachers' Experiences of School Choice from 'Marginalised' and 'Privileged' Public Schools in Oslo." *Journal of Education Policy*. doi:[10.1080/02680939.2019.1587519](https://doi.org/10.1080/02680939.2019.1587519).
- Hølleland, H. 2007. "Nasjonale Prøver Og Kvalitetsutvikling I Skolen" [National Tests and Quality Development in the School]. In *Elevvurdering I Skolen* [Student Assessment in Education], edited by S. Tveit, 29–43. Oslo: Universitetsforlaget.
- Hovdenak, S. S., and J. H. Stray. 2015. *Hva skjer med skolen? En kunnskapssosiologisk analyse av norsk utdanningspolitikk fra 1990-tallet og frem til i dag* [What Happens to the Schools? A Sociological Analysis of Norwegian Education Policy from the 1990s until Today]. Bergen: Fagbokforlaget.
- Jessop, B. 2010. "Cultural Political Economy and Critical Policy Studies." *Critical Policy Studies* 3 (3–4): 336–356. doi:[10.1080/19460171003619741](https://doi.org/10.1080/19460171003619741).
- Kassim, H., and P. Le Galès. 2010. "Exploring Governance in A Multi-Level Polity: A Policy Instruments Approach." *West European Politics* 33 (1): 1–21. doi:[10.1080/01402380903354031](https://doi.org/10.1080/01402380903354031).
- Langfeldt, G., E. Elstad, and S. T. Hopmann. 2008. *Ansvarlighet i skolen* [Accountability in Schools]. Oslo: Cappelen Forlag.
- Lascoumes, P., and P. Le Galès. 2007. "Introduction: Understanding Public Policy through Its Instruments—From the Nature of Instruments to the Sociology of Public Policy Instrumentation." *Governance: an International Journal of Policy, Administration, and Institutions* 20 (1): 1–21. doi:[10.1111/j.1468-0491.2007.00342.x](https://doi.org/10.1111/j.1468-0491.2007.00342.x).
- Le Galès, P. 2010. "Policy Instruments and Governance." In *The SAGE Handbook of Governance*, edited by M. Bevir, 142–160. London: SAGE Publications.
- Lenschow, A., D. Liefierink, and S. Veenman. 2005. "When the Birds Sing: A Framework for Analysing Domestic Factors behind Policy Convergence." *Journal of European Public Policy* 12 (5): 797–816. doi:[10.1080/13501760500161373](https://doi.org/10.1080/13501760500161373).
- MacQueen, K. M., E. McLellan, K. Kay, and B. Milstein. 1998. "Codebook Development for Team-based Qualitative Analysis." *Cultural Anthropology Methods* 10 (2): 31–36. doi:[10.1177/1525822X980100020301](https://doi.org/10.1177/1525822X980100020301).

- Maroy, C. 2015. "Comparing Accountability Policy Tools and Rationales: Various Ways, Various Effects?" In *In Governing Educational Spaces: Knowledge, Teaching, and Learning in Transition*, edited by H. G. Kotthoff and L. Klerides, 35–56. Rotterdam: Sense Publication.
- Mausethagen, S. 2013. "A Research Review of the Impact of Accountability Policies on Teachers' Workplace Relations." *Educational Research Review* 9: 16–33. doi:10.1016/j.edurev.2012.12.001.
- Mausethagen, S., T. S. Prøitz, and G. Skedsmo. 2018. "Teachers' Use of Knowledge Sources in 'Result Meetings': Thin Data and Thick Data Use." *Teachers and Teaching: Theory and Practice* 43 (6): 37–49. doi:10.1080/13540602.2017.1379986.
- Ministry of Education and Research (MER). [2002] 2003. "Appendix No. 3." <https://www.regjerin.no/no/dokumenter/stprp-nr-1-tillegg-nr-3-2002-2003-/id435850/?ch=2>
- Ministry of Education and Research (MER). [2003] 2004. "Kultur for læring [Culture for Learning]." *White Paper No. 30*.
- Ministry of Education and Research (MER). [2007] 2008. "Kvalitet i skolen [Educational Quality]." *White Paper No. 31*.
- Ministry of Education and Research (MER). [2012] 2013. "På rett vei [On the Right Track]." *White Paper No. 20*.
- Ministry of Education and Research (MER). [2016] 2017. "Lærelyst [Eager to Learn]." *White Paper No. 21*.
- Ministry of Education and Research (MER). 2002. "Skolen vet best! [The School Knows Best!]." Oslo: Ministry of Education and Research.
- Møller, J., and G. Skedsmo. 2013. "Modernising Education: New Public Management Reform in the Norwegian Education System." *Journal of Educational Administration and History* 45 (4): 336–353. doi:10.1080/00220620.2013.822353.
- Statistics Norway. 2019a. "Population and Land Area in Urban Settlements". Statistics Norway. <https://www.ssb.no/en/befolkning/statistikker/befteft>
- Statistics Norway, . 2019b. "Pupils in Primary and Lower Secondary School." Statistics Norway. <https://www.ssb.no/en/utdanning/statistikker/utgrs>
- OECD. [1988] 1989. "Ekspertvurdering fra OECD. OECD-vurdering av norsk utdanningspolitikk [Expert Review from the OECD: OECD Review of Norwegian Education Policy]." Oslo: Aschehoug.
- Prøitz, T., S. Mausethagen, and G. Skedsmo. 2019. "District Administrators' Governing Styles in the Enactment of Data-use Practices." *International Journal of Leadership in Education* 1–22. doi:10.1080/13603124.2018.1562097.
- Quality Commission. 2002. "Førsteklasses fra første klasse. Forslag til rammeverk for et nasjonal kvalitetsvurderingssystem av norsk grunnopplæring [Proposal for Framework for the National System of Quality Assessment of the Norwegian Comprehensive Education]." *Green Paper No. 10*.
- Quality Commission. 2003. "I første rekke. Forsterket kvalitet i en grunnopplæring for alle [Improving Quality in Comprehensive Education for All Students]." *Green Paper No. 16*.
- Sahlberg, P. 2016. "The Global Education Reform Movements and Its Impact on Schooling." In *The Handbook of Global Education Policy*, edited by K. Mundy, A. Green, B. Lingard, and A. Verger, 128–144. Oxford: Wiley Blackwell.
- Saldaña, J. 2009. *The Coding Manual for Qualitative Researchers*. London: Sage.
- Skedsmo, G. 2009. "School Governing in Transition? Perspectives, Purposes and Perceptions of Evaluation Policy." PhD diss., University of Oslo.
- Skedsmo, G. 2011. "Formulation and Realisation of Evaluation Policy: Inconsistencies and Problematic Issues." *Educational Assessment, Evaluation and Accountability* 23 (1): 5–20. doi:10.1007/s11092-010-9110-2.
- Steiner-Khamsi, G. 2003. "The Politics of League Tables." *Journal of Social Science Education*, 2003-1 <http://www.jsse.org/index.php/jsse/article/view/470>
- Steiner-Khamsi, G. 2016. "New Directions in Policy Borrowing Research." *Asia Pacific Education Review* 17 (3): 381–390. doi:10.1007/s12564-016-9442-9.

- Steiner-Khamsi, G., B. Karseth, and C. Baek. 2019. "From Science to Politics: Commissioned Reports and Their Political Translation into White Papers." *Journal of Education Policy*. doi:[10.1080/02680939.2019.1656289](https://doi.org/10.1080/02680939.2019.1656289).
- Tveit, S. 2018. "Ambitious and Ambiguous: Shifting Purposes of National Testing in the Legitimation of Assessment Policies in Norway and Sweden (2000–2017)." *Assessment in Education: Principles, Policy & Practice* 25 (3): 327–350. doi:[10.1080/0969594X.2017.1421522](https://doi.org/10.1080/0969594X.2017.1421522).
- UNESCO. 1990. "Planning for Quality." Paris: UNESCO.
- UNESCO. 1995. "Schools under Scrutiny." Paris: UNESCO.
- Verger, A., C. Fontdevila, and A. Zancajo. 2016. *The Privatisation of Education. A Political Economy of Global Education Reform*. New York: Teachers College Press.
- Verger, A., L. Parcerisa, and C. Fontdevila. 2019. "The Growth and Spread of National Assessments and Test-based Accountabilities: A Political Sociology of Global Education Reforms." *Educational Review* 71 (1): 5–30. doi:[10.1080/00131911.2019.1522045](https://doi.org/10.1080/00131911.2019.1522045).
- Werler, T., and K. Sivesind. 2007. "Norway." In *The Education System of Europe*, edited by H. Döbert, W. Hörner, B. von Kopp, and W. Mitter, 573–589. Dordrecht: Kluwer.