

Beneath the surface of accountability: Answerability, responsibility and capacity-building in recent education reforms in Norway

Thomas Hatch

Published online: 17 January 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Recent educational reforms in Norway include national tests and monitoring mechanisms to see if key outcomes are being achieved. At the same time, Norway has not established the follow-up mechanisms like high-stakes incentives and rewards that are characteristic of accountability policies in some other countries. As a consequence, one could argue that Norway has only moved “half-way” toward accountability. In contrast, this paper suggests that these developments in Norwegian policies demonstrate the difficulties of navigating the tensions between promoting two key aspects of accountability—answerability for the achievement of short-term goals and responsibility for the fulfillment of broader purposes—and the challenges of building capacity for both. Exploring developments in the Norwegian context highlights what it may take to develop policies that address both answerability and responsibility and reveals some of the cultural, geographic, political, and economic realities that make it difficult to do so.

Keywords Accountability · Capacity · Educational policy · School improvement

Accountability: The quality or state of being accountable; especially: an obligation or willingness to accept responsibility or to account for one's actions <public officials lacking accountability
— Merriam-Webster Online Dictionary (2011).

Like many countries around the world in recent years, Norwegian education authorities have embarked on an ambitious series of initiatives to increase accountability. Those initiatives include the establishment of specific educational outcomes and the development of a system of national tests and monitoring mechanisms that seek to improve the relatively low average performance level of

T. Hatch (✉)
Teachers College, Columbia University, 525 W. 125th St., New York, NY 10027, USA
e-mail: hatch@tc.edu

Norwegian students on international tests in reading, mathematics and science. At the same time, Norway has not yet chosen to put in place the specific follow-up mechanisms, such as high-stakes incentives and rewards, which are characteristic of new accountability systems in the United States and some other countries. As a consequence, one could argue that, in comparison to recent U.S. policies, the Norwegian educational system has only moved halfway toward accountability.

Rather than seeing this policy choice as a “failure” to fully implement an accountability system, however, I argue that these developments reflect the inevitable tensions between aspects of accountability that focus on *answerability* and those that focus on *responsibility* (Gregory 2003). Answerability emphasizes that individuals and groups should be accountable for meeting specified and agreed upon procedures and/or goals. Responsibility reflects the belief that individuals and groups should be held accountable for living up to and upholding norms of conduct and higher purposes that are often ambiguous and difficult to define. The tensions between answerability and responsibility underlie many of the debates over the growth of accountability policies around the world and the place of market-based control mechanisms and local autonomy within them (Ball et al. 1997; Olszen and Peters 2005; Suspitsyna 2010).

Part of the challenge policymakers face in addressing both answerability and responsibility lies in a well-known paradox: if members of an organization are answerable only for reaching certain outcomes, then logic suggests they should not be held responsible if their actions to reach those goals are inconsistent with broader, undefined responsibilities or purposes; conversely, if those individuals are behaving in ways that are consistent with the pursuit of larger purposes, it seems unreasonable to hold them accountable if they do not meet all of the specified targets along the way (Harmon 1995).¹

In addition to this moral dilemma, the paradox of accountability also reflects the more mundane tensions of organizational management. As Gregory (2003) puts it:

Rational control requires the specification of clear objectives, but organizational effectiveness is continually jeopardized by the realities of a rapidly changing, uncertain and politically charged, environment. Thus, a conundrum: if objectives must be clearly specified to ensure accountable and effective performance, then how often do they need to be revised in the face of pervasive societal change, lest they become obsolete and an impediment to (rather than a component of) effective administration? And if they need to be changed frequently, why try to specify them in the first place? (p. 563)

Despite these tensions, advocates for increasing accountability in many other countries focus attention on increasing answerability. For example, the No Child Left Behind Act (NCLB) in the United States suggests that specifying clear performance targets, developing more sophisticated monitoring systems, and

¹ This paradox underlies issues of central importance to both national governments and international relationships. For example, it was the focus of arguments in the Nuremberg trials over the responsibilities of Nazi officials and fuels contemporary debates about the responsibilities of government and military personnel for war crimes and violations of human rights that may have been ordered, sanctioned, or accepted by their superiors (Gregory 2003).

establishing significant incentives and rewards for meeting those targets will lead to greater motivation in the workforce and greater organizational efficiency. At the same time, researchers point out that such answerability-focused policies can undermine the achievement of other unspecified, yet highly-valued goals and outcomes (Horner et al. 2006; Kearns 2003; Koretz 2008; Supovitz 2009). Furthermore, critics of answerability-focused policies argue that these mechanisms are likely to impede, not enhance, the autonomy of teachers, school leaders and schools (Cochran-Smith and Lytle 2006; Darling-Hammond 2006; Mintrop and Sunderman 2009). Conversely, advocates of increasing answerability may see efforts to promote or preserve responsibility as contributing to education systems that are inefficient, hard to control, and prone to failure.

Whichever side one takes, increasing answerability inevitably impacts efforts to promote responsibility while focusing on responsibility has implications for establishing answerability. As a consequence, accountability systems that seek to promote both answerability and responsibility need to develop appropriate structures and mechanisms as well as the expertise, motivation, resources and other supports to pursue two very different aims at the same time. Unfortunately, education authorities consistently underestimate the extent of capacity building required to implement accountability policies of any kind, let alone those that attempt to promote both answerability and responsibility (Elmore et al. 1996; Elmore 2006). Given these realities, in this paper, I argue that accountability policies need to recognize the mutual relationships between *answerability*—*responsibility*—and *capacity*.² Accountability policies are unlikely to be effective unless all three are addressed and coordinated at least to some extent.

This paper aims to illuminate the relationships between answerability, responsibility, and capacity-building by examining changes made in the goal-setting, monitoring, and follow-up mechanisms in recent Norwegian educational policies. The analysis of these policies is part of a larger project that also looked at evolving school improvement and accountability policies in what can be considered “higher performing” countries on international tests (the Netherlands and Singapore) and “lower performing” countries (Norway and the United States).³ The study sought to describe the theories of action behind these new policies⁴; identify the capacities required to put these new policies in place and carry them out effectively; and document the challenges faced during policy implementation.

² Further illustrating the link between capacity, answerability and responsibility, people are not expected to be answerable or responsible if they have been coerced or if they do not have the competence to act responsibly (due to illness, inability to comprehend; lack of relevant understanding etc.) (see for example, Watson 2001).

³ These classifications were based on available data from the PISA and TIMMS tests at the time of the launch of the study in 2009. Norway and the United States are characterized as “lower performing” (as opposed to “low performing”) because their results are generally average or somewhat below average.

⁴ Theories of action are the beliefs and assumptions, often implicit and unarticulated, that lead people and groups to act in certain ways (Argyris and Schön 1978; Schön and McDonald 1998; Weiss 1995). The larger research study focuses primarily on the “espoused theories” in the policy documents and in the descriptions obtained in interviews with policymakers and administrators, though it also draws on evaluations which reflect what is happening in practice (for a related approach, see Janssens and de Wolf 2009).

In order to go beyond a focus on accountability-as-answerability, this paper examines Norwegian educational policy. Norway is a particularly interesting context in which to explore these issues because recent Norwegian educational reforms focus explicitly on accountability. Like many of their global counterparts, these reforms attempt to create some answerability for the performance of individual students, teachers and schools. Yet there are no equivalent words for “accountability” or “answerability” in Norwegian and responsibility takes center stage. Thus, Norway is a social-welfare state that has been characterized by high levels of trust (Christensen and Peters 1999) and many aspects of the Norwegian educational system reflect an assumption that individuals and organizations can be trusted to carry out their work. Longstanding Norwegian educational policies also reflect societal values and emphasize practices that focus on the development of a cohesive social group and an equitable and common educational for all that may have helped to blunt some of the impact of efforts to increase answerability in other contexts (Ahonen 2001; Volkmar 2008). Exploring developments in the Norwegian context highlights what it may take to develop policies that address both answerability and responsibility. At the same time, such an exploration reveals some of the cultural, geographic, political, and economic realities that influence efforts to balance answerability and responsibility and that raise questions about the values and viability of accountability-related reforms.

The first section of the article presents a brief review of literature on accountability and the conceptual framework related to capacity building that has informed the work. The second section describes the methods. The third section describes the key changes in educational policies in Norway between 1990 and 2000 and between 2000 and 2009, and the fourth section discusses the challenges and capacity issues that have been experienced in the efforts to implement new accountability-related policies between 2000 and 2009. The concluding section discusses the implications for the design and implementation of accountability policies and approaches to systemic school improvement in Norway and other contexts.

Theoretical framework: answerability, responsibility and capacity

Accountability

In general, accountability systems refer to the mechanisms and instruments used to ensure that individuals, groups, organizations, and institutions meet their obligations. While the mechanisms used to hold national governments and education authorities accountable for their performance are also worthy of study, this paper, like many of the recent accountability policies themselves, focuses on the mechanisms and instruments by which the national education authorities (such as government ministers, policymakers, and other government leaders) assign duties and expectations to other members of the education system (such as teachers, school leaders, schools, and municipal and county governments) and attempt to ensure that those individuals and organizations meet those obligations (Leithwood and Earl

2000; Rothman 1995). In Norway, municipalities control primary (1st–7th grade) and lower secondary schooling (8–10th grade) and county governments oversee upper secondary education (11–13th grade), and many recent accountability related policies have focused considerable attention on strengthening the role of these “school owners”.

Following Gregory (2003), this paper highlights changes in the accountability-related mechanisms and instruments that support both answerability for short-term performance targets and responsibility for long-term education goals and purposes. Conventional definitions suggest a fair amount of overlap between answerability and responsibility that can lead to some confusion. Thus, answerability is often defined as the state of being responsible for something and responsibility is often defined as an obligation or expectation for which one is answerable or accountable. Instead, in this paper, I suggest that answerability and responsibility are two distinct aspects of accountability. In order to tease out the distinction between the two, this paper suggests that the extent to which policies promote answerability or responsibility depends on the mechanisms and instruments used to control three different aspects of activity—(a) planning and goal or expectation-setting; (b) monitoring; and (c) follow-up or evaluation (Verhoest 2005). Furthermore, recognizing the overlap between answerability and responsibility, the paper describes answerability and responsibility in terms of two dimensions. One dimension ranges from a situation in which plans, procedures, outcomes, monitoring, and follow-up are clearly specified in advance (characteristic of answerability) to one in which plans, procedures, and outcomes are broadly defined (characteristic of responsibility); the other dimension ranges from a situation in which the plans, procedures, outcomes, monitoring and follow-up are determined by some external authority (characteristic of answerability) to one in which determination of the plans, procedures, outcomes, monitoring and follow-up are left up to the actors (characteristic of responsibility) (see Table 1). Thus, those policies external authorities use to ensure that specific outcomes are identified, monitoring of those outcomes takes place, or follow-up happens are establishing answerability; those that put in place the means to enable the actors involved to participate in goal-setting, monitoring, or follow-up are emphasizing responsibility. Note that while some approaches to accountability may be associated with efforts to promote answerability (such as bureaucratic and legal accountability) and some approaches may emphasize responsibility (such as professional accountability), policies may

Table 1 Accountability and emphases for answerability and responsibility

Focus of accountability	Level of specificity	Locus of control
Plans and outcomes	Clearly specified or broadly defined?	Determined by external authority or left up to the actors?
Means of monitoring	Clearly specified or broadly defined?	Determined by external authority or left up to the actors?
Follow-up	Clearly specified or broadly defined?	Determined by external authority or left up to the actors?

draw on many different forms of accountability at the same time (Romzek and Ingraham 2000; Stone 1995). Similarly, this approach suggests that some aspects of a policy may promote answerability while others support responsibility.

Capacity

Accountability policies often reflect a trade-off between acceptance of short-term answerability for performance for autonomy in other aspects of operations and decision-making. However, creating outcomes, monitoring mechanisms and follow-up do not in and of themselves establish answerability; and simply leaving plans, procedures, outcomes, monitoring, and follow-up undefined does not constitute active promotion of responsibility. Active promotion of answerability and responsibility also depends on capacity-building—making sure that individuals and groups have the competence and resources needed to meet their obligations.

Policy analyses regard capacity-building as one of a number of instruments like mandates, inducements and system-changing mechanisms that policymakers can use to translate policy goals into actions (McDonnell and Elmore 1987). While these instruments are treated as conceptually distinct, in practice, they are all inter-related (McDonnell 1994). For example, in order for mandates and inducements to have some effect, systems have to have the capacity to fulfill the new requirements and to determine whether those requirements have been met. Similarly, policies that focus on building capacity may require some mandates and inducements to ensure that capacity-building takes place and that any newly-developed competence is used effectively.

In order to assess how and to what extent policies are building capacity for answerability and responsibility, the analysis presented here draws on an emerging conception of capacity that goes beyond a conventional focus on the acquisition of financial and physical resources (Cohen and Ball 1999; Ravitch 2000; Stein 2004). This emerging conception suggests that simply *having* resources does not mean that those resources will be *used* well (Corcoran and Goertz 1995; Malen and King Rice 2004). Thus, as Malen and Rice put it, capacity depends on a school's ability to "turn resources into expected outcomes" (p. 636). Examinations of what Malen and Rice call the "productivity dimension" of capacity emphasize that maximizing the use of resources depends on the relationship among a number of different elements. These elements include technical capital—money, physical materials, and resources; human capital—the skills, knowledge and dispositions of the personnel involved; and social capital—the networks, norms of trust and collaboration, and collective commitment among those involved (Cohen and Ball 1999; Corcoran and Goertz 1995; O'Day et al. 1995; Newmann et al. 2000; Putnam 2000; Spillane et al. 2003; Spillane and Thompson 1997). For example, from this perspective, implementing a new, more sophisticated curriculum program (an aspect of technical capital) is likely to be more effective in schools that have more competent teachers (an aspect of human capital) and a more collegial, collaborative culture (an aspect of social capital) (Bryk and Schneider 2002; Newmann et al. 2000). Thus, this analysis examines the ways in which the changes in Norwegian policies seek to support the

development of answerability and responsibility, and then explores the technical, human and social capital required to implement those policies.

Methods

The study was carried out by conducting literature reviews on the recent educational reforms in Norway, analyzing relevant policy documents from the websites of the Norwegian Ministry of Education and the Norwegian Directorate for Education and Training, and interviewing Norwegian policymakers at the national and municipal level who are involved in the design and implementation of new accountability policies. A yearlong stay in Norway also created opportunities to experience education as a parent of two children in a Norwegian elementary school and for many informal meetings and conversations with researchers, parents and teachers that helped to provide background and context for the study.

Collection and analysis of data proceeded in an iterative manner, beginning with the review of the policy documents that were publicly available in English on the websites of the Norwegian Ministry of Education and the Norwegian Directorate for Education and Training. While the initial analysis focused on the fifty-two documents produced after 2002, documents published earlier that described the history and evolution of the Norwegian educational system were also consulted for context and background information. The initial analysis focused on describing key developments in Norwegian educational policies since 2002 by focusing on sections of those documents that referred specifically to the Norwegian Education Act, to the Knowledge Promotion Reforms, or to quality assurance or assessment or monitoring of students, teachers or schools.

Formal interviews were then conducted with five members of the Directorate for Education and Training who worked on the key policy developments related to the Knowledge Promotion reforms, quality assurance, and/or assessment and monitoring;⁵ four members of local municipalities or representatives of municipalities responsible for implementing the Knowledge Promotion and related quality assurance reforms; and two researchers involved in developing or evaluating those reforms. Selection of interviewees was based on purposive and snowball sampling. Researchers and policymakers knowledgeable about the reforms were consulted to identify initial interviewees. These interviewees included a researcher who had previously been involved in the development of the Knowledge Promotion reforms and a senior official at the Directorate with overall responsibility for the Knowledge Promotion reforms. These initial interviewees then identified others at the Directorate who had played key roles in different aspects of the development of quality assurance aspects of the Knowledge Promotion reforms. In turn, representatives of the municipalities responsible for implementing the reforms were selected based on nominations from local researchers, policymakers, and the interviewees at the Directorate. Those selected included two individuals who worked specifically on

⁵ In order to maintain confidentiality the job titles and specific responsibilities of the interviewees are not described.

issues related to quality assurance at an organization representing all municipalities and two individuals who worked on issues related to quality assurance in a local municipality. The interviews focused on the specific roles and responsibilities of the interviewees in developing and/or implementing the reforms; the challenges they were experiencing; and the work on the reforms planned for the future. Analysis of interview transcripts focused on the discussions of the mechanisms for goal setting, monitoring and follow-up in the existing and new policies; the extent to which specific duties and expectations related to these activities were established; and the extent to which efforts were made to build technical, human, or social capital in order to carry out those tasks and duties.

Interviewees were given an initial draft of a report on the findings with an opportunity to correct any errors and provide feedback. Follow-up interviews were conducted with the two initial interviewees and one of the representatives of the municipalities to probe them on their assessment of the accuracy of the report and to fill in details or to respond to questions that arose in the process of writing-up the initial report. Finally, the follow-up interviewees as well as two researchers involved in evaluating the reforms were given an opportunity to review this manuscript and provide further feedback.

Recent educational reforms in Norway: Shifting control and maintaining trust?

The establishment of national tests in 2004 (and a subsequent pause in national testing in 2006 and resumption in 2007) has been the most visible, and contentious move, towards establishing answerability in the Norwegian educational system. Nonetheless, this move is not a simple linear progression from an emphasis on responsibility in earlier years to an emphasis on answerability in later years. Policy debates reflecting different political philosophies, different approaches to student learning, and different approaches to centralization and de-centralization have been underway for some time; correspondingly, reforms in the Norwegian educational system over the past 40 years also reflect a shifting mix of emphases on answerability and responsibility and the capacity-building related to them both (Telhaug et. al. 2004). Whether to have formal assessments and grading procedures, for example, has been a constant and contentious issue in educational debates for decades (Lysne 2006). Similarly, policies over the past 40 years reflect a regular debate over the extent to which there should be a detailed national curriculum and specific regulations over the organisation of teaching and learning in local schools and classrooms. More recently, the rise of performance-based accountability and market-based reforms around the world have fueled these debates and have influenced the policy initiatives of both the conservative and more liberal political parties in Norway.

Even with these constant debates, Norwegian educational policies since the 1980s have focused less attention than many other countries on developing student assessment—a key instrument for monitoring and rewarding performance and outcomes. In fact, students in primary school in Norway (grades 1–7) receive no marks or grades and there is no “failure” and repeating of classes. While there is an

“exit” exam at the end of lower secondary school (10th grade) that leads students into different vocational and university-bound programs, it is perceived by many as relatively low-stakes compared to those in other countries since all students can proceed to the next level and since many can switch programs once they begin upper secondary education (Elstad et al. 2009). Long-standing exit exams mark the end of upper-secondary education and determine whether students get into universities or meet the qualifications for various technical and professional certificates. Therefore, performance on these exams is “high-stakes” for individual students, but generally has not had specific, formal, consequences for school staff, schools, or municipalities.

While these features of the Norwegian educational system endure, a comparison of Norwegian educational reforms in the 1990s and those between 2000 and 2009 highlights an important shift. The reforms of the 1990s reflect a system focused on broad general goals with relatively little attention to mechanisms that could ensure that those goals are fulfilled. The reforms between 2000 and 2009, however, suggest a system grappling with the challenges of establishing new mechanisms and instruments for promoting answerability while also building the capacity to support local and personal responsibility.

Norwegian education policy in the 1990s

A history of policy developments in the 1990s identified several broad public purposes and principles that governed the Norwegian educational system at the end of the twentieth century including:

- *Education for all*—free public education; equal access to education; education adapted to individual needs.
- *Integration*—mixed-ability teaching and integration of pupils with special needs into the ordinary schools.
- *Participation*—cooperation among pupils in school activities and close links between school and home and between school and local community (Norwegian Ministry of Education and Research 2001, p. 13).

In order to fulfill these purposes, at the time, the Norwegian educational system focused much more attention on expanding participation in education and specifying the content and procedures in schools than on specifying the outcomes, monitoring or follow-up activities. In particular, major reforms in 1994 gave all those between the ages of 16 and 19 the right to 3 years of upper secondary education, and reforms in 1997 extended compulsory schooling from 9 to 10 years by lowering the entry age to six. Key reforms during the 1990s also included the introduction of a “Core” curriculum that emphasized broad goals in moral outlook, creative abilities, work, general education, cooperation, and natural environment that linked together the different stages of education (Norwegian Ministry of Education and Research 2001). In addition, curricular reforms also emphasized the need for common subject matter through the development of a more detailed curriculum—“a heavy, authoritative text that makes it unnecessary to talk about the selection of content” (Karlgren and Klette 2008, p. 123). Newly produced guidelines also highlighted how students should

learn—“[Pupils] should be active, enterprising and independent. Pupils should learn by doing, exploring and experimenting” (Norwegian Ministry of Education and Research 2001, p. 14). The curricular reforms in 1997 also established a Saami Curriculum specifically designed for the indigenous Saami population in northern Norway and incorporated Saami culture, history, language and social life into the common Norwegian curriculum.

These curricular reforms in the 1990s were accompanied by “child-oriented”, “family-oriented,” and “culture-oriented” reforms that made the school partly responsible for the provision of a good physical and social learning environment for students, for supporting some organized after school activities, and for the integration of creative activities into the life of the school. Efforts to promote a more inclusive school also extended and reinforced reforms of special education that had abolished most special schools, required the integration of special education students into the regular classroom, and consolidated “the principle of an all-inclusive school, ‘one school for all’ (Norwegian Ministry of Education and Research 2001, p. 32)”. Thus, by the end of the 1990s, the Education Act emphasized regulations related to curriculum content, participation by pupils and parents, class size and mixed age groups, working environment and school premises, transport and lodging and a number of other topics that the central government expected local authorities to adhere to.

While these regulations established expectations that schools would meet certain conditions and use certain resources and procedures, consistent with decentralization efforts taking place in other sectors and parts of the world, the reforms in the 1990s were accompanied by the continuation of block grants initiated in the 1980s that covered all central government subsidies for education, culture, and health. Furthermore, the same reforms that focused attention on the development of the national curriculum also continued to emphasize the placing of “responsibility for educational interpretation and local adaptation of curriculum and syllabi within the individual school” (Norwegian Ministry of Education and Research 2001, p. 9).

Although the major reforms of the 1990s did not focus explicitly on accountability, concerns about the educational quality and performance of Norwegian schools and the lack of mechanisms for monitoring the educational system were also evident. As far back as 1988, a report from the OECD emphasized problems with Norway’s weak assessment and evaluation system and highlighted difficulties in getting data and information on education in Norway. (As one interviewee put it, the report seemed to be saying, “Hello, Mr. Minister, How do you know what’s going on in your schools?”). A series of commissions and related white papers in the 1990s⁶ picked up on these concerns and fueled discussion of the creation of assessment and quality assurance systems.

Despite these discussions, throughout the 1990s the educational system continued to function with a limited emphasis on student assessment, little attention to teacher

⁶ Concerning Student Assessment, School self-evaluation and the National Quality Assessment System in 1995–1996; Toward Richer Goals: Concerning the Comprehensive School, Equal Opportunity in Education and a National Strategy for Assessment and Quality Development in Compulsory and Upper Secondary Education in 1998–1999.

Table 2 Emphases for answerability and responsibility in 1990s reforms

Focus of accountability	Level of specificity	Locus of control
Plans and outcomes	Specified “inputs”—content of curriculum, work methods, programs Broadly defined outcomes focused on access, participation, equity	Determined by external authority
Means of Monitoring	Largely unspecified (no grading or marks before 8th grade; no national testing; only “high stakes” exit exam at end of upper secondary)	Largely left up to the actors
Follow-up	Unspecified	Left up to the actors

evaluation, no formal school inspections, no specific sanctions or rewards for performance, and, some have argued, little or no accountability overall (Elstad 2009) (See Table 2). Instead, the policies put in place in the 1990s used government directives and regulation as key instruments for promoting the development of an equitable, inclusive society through the delivery of a common educational experience and a good social and physical learning environment for all students. In the process, the central government exerted influence and control over the inputs of education—such as the curriculum content and the facilities and programs offered—but left it up to educators, schools, and local authorities to fulfill their responsibilities for carrying out those government directives and complying with those regulations (Klette 2002; Skedsmo 2011).⁷

Norwegian education reforms between 2000 and 2009

Concerns about issues of “quality assurance” in Norway came to the forefront with the publication of the first PISA test results in 2000. Those results showed Norwegian students performing at or below the average of other OECD countries (and far below Finland, another Nordic country that topped the charts). Now referred to simply as “the PISA shock”, the publication of the test results led to a burst of headlines in newspapers and speeches in Parliament about the suddenly discovered poor standing of the Norwegian schools.⁸ Adding to the concerns, results of the next rounds of both the TIMSS and PISA tests showed a decline in the performance of Norwegian students. These developments contributed to a broad consensus across both the conservative and liberal parties in the government that significant changes needed to be made. While the initial policy changes moved quickly in the direction of establishing answerability, efforts to maintain and,

⁷ In an analysis of different approaches to school restructuring in the 1990s in Scandinavia, Klette (2002) characterizes the Norwegian and Danish approaches as a “this you have to deliver” strategy. She contrasts this approach with a “this you have to achieve” approach more characteristic of Sweden and Finland during the same time period.

⁸ While “shocking” to many, the PISA results had been foreshadowed by the relatively poor performance of Norwegian students on the first administration of the Third International Mathematics and Science Study.

ultimately, promote responsibility, came to the forefront following the political and public controversies that arose over the initial implementation of the national tests.

Establishing answerability through new objectives and monitoring mechanisms

The broad consensus on the need for education reforms following the “PISA shock” launched a series of dramatic changes in the educational system after 2000. Many of these changes were associated with the Knowledge Promotion reforms launched in 2006. Reflecting the broad political consensus about the need for changes, the Knowledge Promotion reforms were designed after the conservative government took office in 2001 but were implemented by a more liberal coalition government elected in 2005. The Knowledge Promotion reforms introduced new curricula with clearly stated competence objectives and an emphasis in every subject on the development of basic skills including oral and written expression, reading, numeracy, and the use of digital tools, with reading and writing emphasized from first grade.⁹ The Knowledge Promotion reforms also introduced a quality assurance scheme that brought together a variety of instruments that could be used to determine what kind of progress students were making, how schools were doing in helping their students meet defined objectives, and the extent to which schools and local authorities were fulfilling their statutory obligations. These instruments included pupil and parent surveys, inspections of school owners’ compliance with legal requirements, and national tests.

Although embraced as part of the Knowledge Promotion reforms, the pupil and parent surveys actually arose out of discussions in the 1990s about the possible development of an assessment and inspection system in Norway that preceded the PISA shock of 2000. The first pupil surveys began in 2001 at the behest of the Education Minister at the time who argued that Norway did not need a system of formal school inspectors because the pupils themselves could play that role (consistent with that perspective, the surveys were referred to by some as the “pupil inspectors”).¹⁰ Subsequently, the government has required all schools to carry out the pupil surveys in the spring of the students’ 7th and 10th grade years and in the first year of upper secondary education. It provides information that is used to address whether or not key goals of Norwegian primary and secondary education—including those related to participation, academics, and the social and physical environment—are being achieved. For example, the survey covers such issues as whether students feel included in the school, whether they feel they can master the academic tasks given to them, whether their education is adapted for them, how challenged they are, what kind of feedback they get, as well as their perceptions of issues like bullying (Norwegian Directorate for Education and Training 2008). The Directorate for Education now produces similar surveys to collect comparable

⁹ This basic skills emphasis constituted a particularly significant change in the early years: up until 1997, six year olds did not even attend compulsory public school, much less focus on basic skills.

¹⁰ In addition to this, there is also an Apprentice Survey and an Instructor Survey. These are not discussed further here because vocational education and training (VET) based at companies are not a part of this study.

information from teachers and parents. The government does not require schools to use these surveys, but many do.

Like the surveys, a fledgling effort to monitor education authorities also began before the official launch of the Knowledge Promotion reforms. That effort involved the creation of a division within the newly established Directorate of Education¹¹ in 2004. That new division became responsible for coordinating what the Norwegians refer to as “Tylsyn” and loosely describe as the closest thing they have to some form of educational inspection. Essentially, the annual Tylsyn focuses on different themes or aspects of the Education Act (such as Special Education) and asks municipalities and counties for evidence that they have procedures, routines and reports that can determine whether or not their schools are in compliance with the relevant regulations.¹²

Although the national tests are closely linked to the Knowledge Promotion reforms of 2006, the Norwegian Parliament authorized the development of the national tests in 2003, and they were implemented for the first time barely 1 year later in 2004. The national tests are currently administered in the fall of 5th and 8th grade in reading, numeracy, and English. These tests were designed to focus attention on basic skills, to provide information on individual student performance, and to give local and national information on overall student progress. A series of “mapping tests” are also required in order to identify students who need special help in developing their basic skills (Norwegian Directorate for Education and Training 2011).

These three new instruments—surveys, legal inspections, and national tests—now help to specify exactly what outcomes students should be achieving and serve as the key means of monitoring the Norwegian education system and determining the extent to which the desired outcomes are being achieved. Beyond defining outcomes and establishing mechanisms to strengthen monitoring, Norwegian policymakers also used what could be considered “system-changing” (McDonnell and Elmore 1987) efforts to identify specific actors who could be held accountable for school performance. Thus, as Norway put in place the structures to support the new scheme for quality assurance, policymakers sought to clear up ambiguous regulations and specify the lines of authority in the Norwegian educational system. Considerable discussion also involved who should be held responsible for quality assurance, and reports and policy documents began to use the term “school owner” to refer to the county and municipal governments and to highlight their specific responsibilities in the process. In 2005, amendments to the Education Act also named “the school owner” as responsible for ensuring correct and necessary competence at the school level, and required school owners to have a system for providing necessary competence enhancement.

¹¹ The Directorate of Education was itself a product of the “agentification” and de-evolution of government functions, and was designed to increase the efficiency and effectiveness of the central government’s work in education.

¹² Note that the Tylsyn focuses on whether local education authorities have systems in place to determine the extent to which schools are in compliance; it does not assess whether schools actually are in compliance.

Promoting responsibility with flexibility, limited follow-up and capacity-building

The establishment of the new instruments for answerability were controversial, but some of the most vociferous debates focused on what to do to follow-up and ensure the achievement of the new competence aims. Initially, the conservative government in power when the Knowledge Promotion reforms were being designed sought to use the national test results as a means of putting public pressure on schools and motivating schools to make sure their students met the new outcomes. In conjunction with the conservative government's efforts to introduce greater competition and choice in the education system, publication of this outcome information was expected to lead to positive or negative publicity. In addition, the information was to be used by parents who would advocate for better schools and, if necessary, demand to switch their children to higher-performing schools. Correspondingly, the first tests were carried out in the spring of 2004; school-by-school results were released; and newspapers and media quickly produced rankings and "league-tables" of schools.

The subsequent outcry along with the election of a new government composed of a coalition of more liberal parties in 2005, however, led to what could be considered a second phase of the new reforms (Nusche et al. 2011). In this phase, key changes in these policies shifted from a focus on establishing answerability towards more emphasis on capacity-building and the active promotion of the responsibilities of schools and school staff. In particular, the Directorate of Education moved the national tests to the fall of 5th and 8th grade. That change was consistent with the arguments of those who felt that the tests should not be used as outcome measures of what had been learned in 4th and 7th grade; moving the tests suggested instead that the tests should be used in a monitoring capacity to provide teachers and schools with information to guide their instruction of students. Furthermore, regulations were passed forbidding education authorities from using the results of the national tests for school rankings. As a result, while the Norwegian Directorate of Education has developed a website for making overall test results and information from surveys available to school personnel and parents, the results are only publicly reported by municipality and region and not school-by-school or by student.¹³ In addition to this shift designed to limit the public pressure put on schools, the government has so far chosen not to attach any specific stakes—rewards or sanctions—to the outcomes of the national tests.¹⁴

¹³ Despite the focus on over-all results and the prohibition on using information from the national tests for rankings, the well-established right to public information in Norway means that newspapers can still get the data on the performance of individual schools, and it is not illegal for them to produce their own rankings using this information. As a consequence, newspapers and other news media often publish school results on the same day that the Directorate releases the information on performance at the municipal and county level. Denmark, which has a similar right to public information, addressed this issue by specifically exempting the school testing data from the right to public information.

¹⁴ While these policies describe what is happening in Norway overall, there are exceptions. For example, the municipality of Oslo, which continues to have a more conservative government, has also made a significant commitment to and investment in using tests to improve performance. That commitment includes the development and implementation of a wide range of tests in many different subjects and grade levels; the publication of school-by-school results; and the use of the results to inform decision-making by Oslo school administrators.

Paralleling these developments, an overall report on the results of the Tylsyn is now produced and those school owners who are out of compliance are informed of their status and ordered to make changes. Beginning in 2009, school owners were also asked to produce an annual report that describes the state of their educational system (Nusche et al. 2011). However, no specific rewards or sanctions have been established nor legal actions pursued to follow-up on the results of the Tylsyn or to increase compliance.

In addition to trying to limit the use of the national tests as an instrument for the follow-up necessary to establish answerability, the Knowledge Promotion reforms also emphasized the responsibility of local actors for achieving the new outcomes in a number of ways. First, at the same time that the Knowledge Promotion reforms established specific outcome targets, those same reforms also gave schools and school owners substantial freedom to adapt and develop local curricula. The reforms between 2000 and 2009 also loosened the control on the “inputs” to the educational system by relaxing some of the regulations on the content and the organization of classroom instruction that were a part of the 1997 reforms. For example, in 2003, changes to the Education Act included an amendment that repealed regulations covering class size and groupings of students (Nusche et al. 2011).

Second, the post- 2000 reforms also involved a series of initiatives after 2005 that explicitly recognized that the local actors need considerable support in order to take advantage of the information of the new quality assurance system and the new flexibility that came along with it. Several of these initiatives efforts focused specifically on building competence for assessment and evaluation. The “Competence for Development” strategy, for example, (carried out from 2005 to 2008) emphasized local capacity-building and “a boost in competence” for local actors as a key means for meeting the challenges “related to the changes in content and structure which the reform involves” (Norwegian Ministry of Education and Research 2004, p. 4). From 2006 to 2010, the Directorate also led a project—Knowledge Promotion—From Word to Deed—that included the development of eleven regional groups to provide support for school evaluations. The Better Assessment Project (2007–2009) and the Assessment for Learning Programme (2010–2014) have also involved a number of schools in pilot projects to try out different approaches to the development of assessment criteria and the use of formative assessments.

Several projects have also focused specifically on building the capacity of school leaders and teachers and raising the status of the teaching profession. These projects include a program for school leaders established in 2009. A “Competence for Quality” initiative, launched in 2008 was also designed to create a more robust system for teachers’ professional development by enabling some teachers to take courses to ensure that they have the competence to teach their subjects. The “SPARK” initiative, also begun in 2008, is an ongoing partnership between teacher training institutions, school owners, unions, and the education authorities designed to raise school quality by improving teachers’ skills and improving the recruitment of teachers. While not technically a part of the Knowledge Promotion reforms, efforts to strengthen teacher education and improve the skills and knowledge of teachers have been underway for some time. Minimum competence requirements were put in place in 2005; in 2010, reforms sought to emphasize the development of

subject matter knowledge and pedagogical skills by introducing new competence aims for teacher education and by establishing distinct programs that qualify candidates to teach in either years 1–7 or 5–10.¹⁵

For the most part, however, these initiatives continue to place much of the responsibility for meeting these development obligations with the local actors by suggesting that competence development “can only happen through school owners and schools themselves accepting responsibility” and that “the responsibility for determining priorities and approving measures for the development of competence lies with school owners” (Norwegian Ministry of Education and Research 2004, p. 3). Even in cases where outcomes have been identified, in general, means for following-up remain unspecified. Thus, the funds for Competence Development from the Ministry of Education were to be awarded only after meeting several conditions including that school owners “fulfill their responsibility for competence development through plans approved by the local council or regional assembly” which specifies the concrete competence development measures proposed (Norwegian Ministry of Education and Research 2004, p. 14). Despite the establishment of these conditions, it is not clear how the plans of the local authorities were assessed or whether any funds were actually withheld. Similarly, the partners involved in the SPARK initiative have agreed on a set of twenty-three different indicators to use for monitoring and evaluating their progress, but the consequences for failing to meet the key goals remains unclear (Nusche et al. 2011).

The challenges of building the capacity for answerability and responsibility

The reforms from 2000 to 2009 mark a significant departure from previous policies that imply that simply providing appropriate inputs could lead to the fulfillment of the broad purposes of the educational system. Instead, the new reforms suggest that in order to assure that all students receive a quality education, local actors need to be answerable for making sure that students meet specific competence aims, but that, with adequate support, those local actors can be trusted to monitor their performance and make adjustments as needed (See Table 3). At the same time, the competence for development and competence for quality strategies and related initiatives suggest a recognition that improving the performance of Norwegian schools is not solely a matter of motivation—of encouraging educators and schools to do things they had been unwilling to do; instead, these investments suggest that local actors can and will make improvements if they have the competence to do so, but that work needs to be done to build that competence and enable local actors to carry out their responsibilities.

Thus, the 2000 reforms suggest an explicit effort to shift a system focused previously on the development, delivery, and implementation of curriculum and other government sponsored programs to a system that produces useful information on system performance and enables people and organizations to use that information to make appropriate adjustments in instructional and organizational practices at the local level. At the same time, the efforts to implement the post-2000 reforms also

¹⁵ Previously, qualified graduates of teacher training could teach at any level from years 1–10.

Table 3 Emphases for answerability and responsibility in reforms from 2000 to 2009

Focus of accountability	Level of specificity	Locus of control
Plans and outcomes	Specified outcomes in all subjects, clearly indicating what pupils and apprentices are expected to learn	Determined by external authority
Means of monitoring	Established specific Quality Assurance Scheme including pupil surveys; Tylysyn (inspection of compliance with regulations); and national tests; clarified regulations regarding responsibilities of “school owners” for quality assurance	Determined by external authority
Follow-up	Developed a series of initiatives to help build local capacity for quality assurance; for using assessment and evaluation data to improve performance; and for strengthening school leadership, and teaching and learning	Left up to the actors

demonstrate the many, often unanticipated, capacity demands that come with promoting both answerability and responsibility. These include the technical demands of building an assessment and testing infrastructure almost from scratch; the human capital demands of building expertise on testing, assessment, and the use of data to improve student and school performance; and the social capital demands of connecting and sharing information and expertise across widely distributed and largely independent local actors.

Establishing systems and structures: Technical capital in the post-2000 reforms

Given the limited use of tests in Norway, the demand to develop and implement new national tests in barely a year's time produced considerable challenges. These challenges included the fact that there were no existing large-scale tests matched to the Norwegian curriculum. International tests and related resources could have been imported (and in some cases they were used as a basis for test design), but those did not match well with the content of the Norwegian curriculum. Furthermore, the Norwegian curriculum itself did not provide a detailed framework that could be used to guide test construction.

Under these conditions, the first generation of national tests implemented in 2004 were judged by experts to be inadequate and unreliable (Elstad et al. 2009; Lie et al. 2004, 2005). Concerns about the quality of the tests were so serious that results of some of the initial test results were never released, and the advisory panel reviewing the tests concluded one of their reports on the tests by saying that in their best judgment “a continuation of the national tests would not be good for the Norwegian schools”.¹⁶ Resistance to the tests was further fueled by the fact the tests included many open-ended items and subscales that placed considerable demands on the teachers who were given responsibility for scoring them. As a consequence of these

¹⁶ This was a translation provided by an interviewee who was reading from the report in Norwegian. The report also concluded that the Directorate did not have sufficient oversight over the administration of the tests and that the poor quality of many of the tests and a high boycott rate at the upper secondary level meant the results from the 2005 tests should not have been made public.

problems, the government declared a “pause” in the national testing and no tests were administered in 2006. During the pause, the test makers revised the tests substantially and significantly upgraded their quality according to the expert reviewers. While some argued for a continuation of the “pause”, the national tests were re-launched again in 2007; however, the initial plans for some tests, such as tests in writing were dropped because they still were judged as inadequate.

Complicating matters further, the Tylsyn’s review of local education authorities made it clear that in addition to the technical problems in producing relevant assessment data, there were very few systems, routines and procedures in place for communicating data or sharing information on performance between schools, school leaders, school owners and the central government. As one interviewee put it, since 2006, the annual Tylsyn has shown that there’s “a lack of knowledge” from the local authorities about how their schools are working and “there’s no control of whether every student in those schools are getting their education in the right way”. Correspondingly, the first report of the Tylsyn suggested that as many as 70 % of municipalities surveyed did not fulfill the requirements for the evaluation and follow-up of schools (Elstad et al. 2009; Norwegian Directorate for Education and Training 2006).¹⁷

Illustrating the complex relationship between system-changing policies and capacity building, the very decentralization efforts before 2000 that set the stage for later accountability-related reforms may have also undermined the capacity of the school owners to use data on school performance effectively. Thus, consistent with the “de-evolution” rhetoric in the reforms before the turn of the century, the local authorities like the municipal city council (who are responsible for all local services not just education) delegated many tasks and responsibilities to the school directors and schools themselves. In the process, many local authorities reduced their administrative staff. Yet the post-2000 reforms specifically sought to establish the local authorities (such as the municipal city council) as the “school owners”. As a result, the same “owners” who were urged to delegate and reduce their administrative structures before 2000 were now being asked to take responsibility for a whole set of new tasks and ensure the achievement of specific outcomes. With an extremely thin educational leadership structure, in many cases, there was no one available to take on these new responsibilities. Given these realities, many responsibilities of school owners continue to be delegated to school administrators and recent evaluations suggest a continued lack of clarity around who the “school owners” really are and who should be responsible for providing the “assurance” in the new Quality Assurance scheme (Møller et al. 2009; Price Waterhouse Coopers and The Norwegian Association of Local Authorities (KS) 2009).

Finding and building expertise: Human capital in the post-2000 reforms

The technical problems in test design and the development of an infrastructure for monitoring performance and using assessment data productively went hand-in-hand with a host of human capital issues. The scale and time-line for the design of the

¹⁷ It should be noted that this is not a random sample of municipalities; the Tylsyn focused on municipalities where some problems were suspected.

national tests, in particular, created a demand for expertise that would have taxed even experienced psychometricians. However, with little or no tradition of test construction in Norway, there were few experienced psychometricians to turn to. Instead, the responsibility for designing the tests was turned over to discipline-based centers at Norwegian universities that focused primarily on curriculum and instruction not assessment or testing.

Similarly, with legal regulations prohibiting the use of formal marks and with limited assessment before eighth grade, many teachers and school leaders also had limited experience using test data to improve performance. Furthermore, neither school owners nor the local and central authorities were accustomed to getting and using performance data on the schools. This meant that people throughout the education system had to figure out how to integrate the collection and use of performance data, the planning for competence development, and the relevant professional development into their existing schedules and responsibilities. At the same time, those that needed help developing their assessment expertise also had difficulty finding local experts who could offer workshops and other support needed to respond to the new demands. Even the Directorate of Education has sometimes had difficulty finding researchers who had the relevant expertise and interest in collecting and analyzing the data needed to carry out systematic formative and summative evaluations of key aspects of the reforms and related efforts to support them.

While many members of the education system were expected to be using data to improve performance, the school owners in the counties and municipalities were also being asked to perform an entirely new role. While their work before the post-2000 reforms had often focused on providing resources and support, school owners were now asked to monitor progress toward the new outcomes. As one evaluator reported, many of the interviewees she talked to recognized that the school owners were given a new role: “no more development. Control”.¹⁸

Developing networks: Social capital in the post-2000 reforms

The technical and human capital problems were compounded by the demands of a widely-dispersed population with many geographically isolated regions and small communities. Thus, smaller communities, particularly those with only one school are unlikely to have much if any administrative structure beyond the school leader, and those municipal leaders are unlikely to have much, if any, expertise in education (Norwegian Directorate for Education and Training 2011; Nusche et al. 2011). Furthermore, members of those small schools and communities face more significant hurdles in developing their human and social capital—they are likely to have to travel much greater distances to get access to and to get connected with experts and others who can help them to meet their new assessment responsibilities.

In contrast, larger municipalities like Oslo have a much larger administrative structure for their schools. These larger structures mean more opportunities to hire

¹⁸ In some ways, this shift in focus for Norwegian educational administrators appears to move in the opposite direction from shifts underway in the United States where many district administrators accustomed to monitoring and control are now asked to take on support roles for schools (Honig 2008; Honig et al. 2010).

individuals with knowledge and competence in assessment and more flexibility in reorganizing responsibilities to take on the new tasks of monitoring performance and using data to improve performance. In addition, municipalities in and around larger population centers are within easier reach of major Universities and other education related organizations and institutions. For example, the Oslo schools could take advantage of connections to and courses provided by researchers at the University of Oslo, including some of the foremost experts in the country on testing and assessment. The Oslo school administration also includes staff who have worked at the Directorate of Education (also housed in Oslo), and the Oslo School Director (or Superintendent) herself, is the former chair of one the government committees that authored the white paper that led to the development of the national tests and many of the Knowledge Promotion reforms. All of these connections help to build a strong network of relationships that can facilitate the sharing of resources, information, and expertise in one municipality. However, the wide distribution of schools and weak connections among many parts of the educational system suggest that this social capital is inequitably distributed. The result is a situation in which better schools in larger municipalities may have an easier time in building the technical and human capital needed to carry out the post-2000 reforms.

To address the challenges of building this social capital, the Directorate for Education and Training and organizations like the Norwegian Association of Local Authorities (KS) have focused specifically on fostering the development of local and regional networks. The Norwegian Association of Local Authorities supports the governing boards and administration of the counties and municipalities and has long-pursued a network-based strategy to help local authorities share information, strategies and expertise related to governance in a variety of sectors. This same network strategy has been used explicitly in recent years to help the school owners develop a common understanding of the new role they are expected to play in education as a result of the post-2000 reforms. In addition, the Association has also sponsored and supported the development of networks of schools within and across regions and held meetings designed to bring together the unions, school administrators, local authorities and universities and colleges in dialogues around issues like assessment and quality assurance.

Several of the capacity-building initiatives launched by the Directorate of Education are also specifically designed to create opportunities for the development of cross-school and cross-region exchanges and relationships. For example, the Better Assessment Project provided support for schools to work together with other schools, teacher training institutions, and the local and county authorities to develop their assessment competence. Similarly, the SPARK partnership builds on the network strategy of the Norwegian Association of Local Authorities to foster collaborations at the regional and local level among the many different organizations involved in education.

Conclusion and implications

On one level, the Norwegian educational reforms since the turn of the twenty-first century reflect a shift in the focus of control from the curriculum and the inputs of

education to the outcomes. In the process, Norway has put many of the mechanisms for defining specific outcomes and determining who is answerable in place, but it has not established the specific means for following-up and ensuring that those who are answerable meet their objectives. However, Norway has also explicitly chosen to encourage local actors to take the responsibility for monitoring performance and making improvements when necessary. As a result, while one could argue that Norway has only gone part-way to establishing answerability, Norway also seems to be trying to promote responsibility in ways that to balance some of the “side-effects” that come with concentrating on one or the other.

In order to make such a system work, however, Norway has to confront the key challenges that come with a massive capacity-development project in a system that up until the past 10 years has neither provided much support for answerability nor actively promoted the capacity development needed for individuals and organizations throughout the system to carry out their responsibilities. In addition, the Norwegian reforms may need to confront significant conflicts and contradictions between the theory of action of the new accountability system and the basic values and assumptions of Norwegian society underlying the traditional education system. Any efforts to address and change some of those traditional assumptions and values also have to confront the fact that, even with relatively poor international test scores, other indicators of educational, economic, and social well-being suggest that the society is doing relatively well.

Issues of capacity-building

The efforts to build assessment competence have only just begun, and they reflect the need to work at every level of the system. Teachers need the tools, resources and expertise to assess students' strengths and needs and respond appropriately; school leaders and school owners need systems that provide accurate data on student and school performance, and they need to learn how to use it to guide their planning, resource-allocation, and decision-making; researchers need to be able carry out projects that evaluate the new developments and help to build theoretical and practical knowledge about assessment and quality assurance and how the assessment-related reforms are working in Norway; higher education institutions need to build on growing expertise about what is working and what is not to provide courses and professional development experiences that focus on assessment, data-based decision-making, and organizational development; government officials and politicians need to learn how to equip and motivate people to work in a system focused on the achievement of competence aims and to understand the inappropriate and appropriate uses for testing in that system; and parents need to understand information on the quality of their children's schooling and to base their decisions about how to educate their children on it.

While these capacity development efforts may seek to improve performance of all students and schools, those efforts also have to confront the fact that it takes capacity to build capacity. If larger, better connected school systems, and schools with strong self-assessment structures already in place are most likely to be effective; they are also those that are most likely to benefit from positive publicity

and other rewards, and most likely to improve. In contrast, those that lack the capacity to meet the new competence aims are also those least likely to be interested in or able to take responsibility for the planning and self-assessment needed to become more successful. Overall, initial results are at least moving in the right as direction as Norway's scores on the most recent PISA tests in 2010 in reading, mathematics, and science all rose (Norwegian Directorate for Education and Training 2011). Nonetheless, there are also signs that the capacity to meet the demands of the new reforms are already inequitably distributed in ways that advantage larger and more central municipalities (Norwegian Directorate for Education and Training 2011; Nusche et al. 2011). Correspondingly, schools in larger municipalities are likely to benefit from the social, technical and human capital concentrated in those areas; while schools in smaller, more geographically-dispersed municipalities will have to reach beyond their borders to get access to the expertise and support they need.

Conflicts of values and assumptions

In addition to the challenges of building capacity, the future of the Norwegian education reform efforts depends on dealing with some basic tensions including: recognizing individual differences or treating all students equally, focusing on the individual or the group, and concentrating on developing discrete academic skills or healthy development overall. While it is possible rhetorically to support many of these goals, decisions over how to allocate scarce resources, time and attention often bring out conflicting values and assumptions about how to achieve equity, what kind of society to have, and how children develop. Thus, investing in assessing individual skills and increasing answerability makes differences in performances visible; but those efforts fly in the face of aspects of Norwegian society that discourage individuals from distinguishing themselves from others and touting their achievements. Similarly, reserving rewards for selected individuals may not make sense to many in a society that emphasizes equal treatment for all.

One could also argue that the new reforms conflict with some of the basic assumptions underlying the way that the central government treats other aspects of Norwegian society. For example, extensive parental leave policies and the limited emphasis on early education in comparison to many other countries together with the low levels of homework, instructional time, and the lack of any marks or formal feedback before the end of 7th grade suggest that childhood and the primary school years in Norway should be protected space. In contrast, government policies place much more attention on adult education than in other countries and relegate the development of academic skills, and preparation for college and careers to the upper secondary school.

Furthermore, the maintenance of many small schools (and many small local teacher education programs) serves as a way for the government to provide some of the funding, jobs and support needed to sustain widely dispersed, small communities. Yet, maintaining such a geographically dispersed system makes it difficult to consolidate schools, tighten networks, increase oversight and ensure the achievement of common objectives. A system that prizes the allocation of funding based on

equity and political and economic realities (rather than expertise or performance) may also have undermined the very efforts to develop the national tests and an assessment infrastructure as that work was farmed out around the country. As long as these political and economic emphases are maintained, concerns about educational performance and capacity-building may not drive decision-making and resource allocation.

Issues of effectiveness

Ironically, one of the biggest impediments to addressing these tensions in basic values and assumptions and to putting a greater emphasis on accountability in the Norwegian education system may be the fact that the Norwegian system can also be viewed as working. While Norway has not been a top performer on the PISA and TIMSS tests, there are numerous other indicators that suggest that Norway is doing quite well educationally in both absolute and relative terms. For example, adult literacy rates are quite high, and rates of participation in adult learning are also significantly higher than many other countries (Boarini 2009; Statistics Norway 2011). Student engagement and student–teacher relations are also much better than they are in many other countries; and Norway, like other Nordic countries, consistently ranks high on many measures of health and well-being (Legatum Institute 2010; United Nations Human Development Program 2010).

While the dangers of a poorly performing educational system for a nation's economy are touted around the globe (and in Norway) (see for example, Hanushek 2010), Norway has emerged as one of the strongest economies in the world—despite relatively low international test scores. Furthermore, the emphasis in Norwegian society on social cohesion and equity seem to be accompanied by notable accomplishments. Thus, Norway's schools are much less segregated and more inclusive than those in many other parts of the world, and results across schools and across socio-economic groups are much more equitable than they are in almost all other OECD countries (Boarini 2009; OECD 2006). And Norway has managed to do all of this with a school system that has not relied on tests, grading, or other summative assessment in primary school; has had a wide range of shifting policy demands over the past 20 years; little or no means of determining whether or not the education system is working; and few if any rewards or consequences for poor performance. The high-functioning society and relatively powerful economy provide reinforcement to resist the very changes so many argue that a successful economy requires.

Looking toward the future

The specific challenges for educational reforms in Norway now include figuring out whether and how some improvements in key skills and competencies can be made without sacrificing the other educational, social and equity outcomes that have been achieved, and within a context where the emphasis is on quality of life overall rather than the achievement of individual objectives. The challenge for Norway and other countries in general lies in how to confront the paradoxes of accountability and how

to address the tensions of answerability and responsibility and the related needs for capacity-building. This analysis suggests that accountability policies need to recognize the differences between promoting answerability and promoting responsibility. In particular, there is no guarantee that meeting the specific goals used to ensure answerability will also ensure the fulfillment of the broader purposes of education that are the focus of efforts to promote responsibility (and vice versa). “Side effects” will come with every system. Accountability systems have to recognize the tensions between answerability and responsibility and address the predictable problems that result while taking into account the specific demands of the cultural, economic, and political context that shape their policies.

Norway’s approach to accountability reflects one kind of compromise where mechanisms for planning, goal setting, and monitoring emphasize answerability while follow-up depends on responsibility. Such a system that moves part-way toward answerability may never produce the kind of tight alignment and efficiency that could result in chart-topping performance on international tests; but it may also allow for more attention to the capacity-building that supports a focus on the fulfillment of broader purposes—an emphasis lacking in many other accountability systems. It remains to be seen, however, whether another approach could more effectively promote answerability and responsibility and facilitate the capacity-building central to the achievement of both.

References

- Accountability. (2011). In *Merriam-Webster Online Dictionary*. Retrieved March 20, 2011, from <http://www.merriam-webster.com/dictionary/accountability>.
- Ahonen, S. (2001). The end of common school? Change in the ethos and politics of education in Finland towards the end of the 1990s. In S. Ahonen & L. Rantala (Eds.), *Nordic lights: Education for nations and civic society in the Nordic countries, 1850–2000* (pp. 175–203). Helsinki: Tammer-Paino Oy.
- Argyris, C., & Schön, D. (1978). *Organizational learning: A theory of action perspective*. Reading, MA: Addison-Wesley.
- Ball, S. J., Vincent, C., & Radnor, H. (1997). Into confusion: LEAs, accountability and democracy. *Journal of Education Policy*, 12(3), 147–164.
- Boarin, R. (2009). Making the most of Norwegian schools. *Economics Department working papers* no. 661. Paris, France: OECD.
- Bryk, A., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.
- Christensen, T., & Peters, B. G. (1999). *Structure, culture, and governance: A comparison of Norway and the United States*. Lanham, MD: Rowman & Littlefield.
- Cochran-Smith, M., & Lytle, S. (2006). Troubling images of teaching in no child left behind. *Harvard Educational Review*, 26(4), 668–697.
- Cohen, D., & Ball, D. (1999). *Instruction, capacity, and improvement*. Pennsylvania, PA: Consortium for Policy Research in Education.
- Corcoran, T., & Goertz, M. (1995). Instructional capacity and high performance schools. *Educational Researcher*, 24(9), 27–31.
- Darling-Hammond, L. (2006). No child left behind and high school reform. *Harvard Educational Review*, 26(4), 642–667.
- Elmore, R. (2006). The problem of capacity in the (Re)Design of educational accountability systems. *Paper presented at a symposium on NCLB and its alternatives*. Campaign for Educational Equity, New York, NY.

- Elmore, R. F., Ablemann, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. In H. F. Ladd (Ed.), *Holding schools accountable: performance-based reform in education* (pp. 65–98). Washington, DC: Brookings Institution.
- Elstad, E. (2009). Schools which are named, shamed and blamed by the media: School accountability in Norway. *Educational Assessment, Evaluation and Accountability*, 21, 173–189.
- Elstad, E., Nordtvedt, G., & Turmo, A. (2009). The Norwegian assessment system. An accountability perspective. *CADMO*, 17(1), 89–103.
- Gregory, R. (2003). Accountability in modern government. In B. G. Peters & J. Pierre (Eds.), *Handbook of public administration* (pp. 557–568). London: Sage.
- Hanushek, E. (2010). *The high cost of low educational performance*. Paris: OECD.
- Harmon, M. (1995). *Responsibility as paradox: A critique of rational discourse on government*. London: Sage.
- Honig, M. I. (2008). District central offices as learning organizations: How sociocultural and organizational learning theories elaborate district central office administrators' participation in teaching and learning improvement efforts. *American Journal of Education*, 114, 627–664.
- Honig, M. I., Copland, M. A., Rainey, L., Lorton, J. A., & Newton, M. (2010). School district central office transformation for teaching and learning improvement. A report to the Wallace Foundation. Seattle, WA: The Center for the Study of Teaching and Policy.
- Horner, L., Lekhi, R., & Blaug, R. (2006). Deliberative democracy and the role of public managers: Final report of The Work Foundation's public value consortium. London: Work Foundation. Retrieved on November 1, 2012 from http://www.theworkfoundation.com/DownloadPublication/Report/107_107_Deliberative%20democracy%20and%20the%20role%20of%20public%20managers.pdf.
- Janssens, F. J., & de Wolf, I. F. (2009). Analyzing the assumptions of a policy program: An ex-ante evaluation of "educational governance" in the Netherlands. *American Journal of Evaluation*, 30(3), 411–425.
- Karlgren, I., & Klette, K. (2008). Reconstructions of Nordic teachers: Reform policies and teachers' work during the 1990s. *Scandinavian Journal of Educational Research*, 52(2), 117–133.
- Kearns, K. P. (2003). Accountability in a seamless economics. In G. Peters & J. Pierre (Eds.), *Handbook of public administration* (pp. 581–589). London: Sage Publications.
- Klette, K. (2002). Reform policy and teacher professionalism in four Nordic countries. *Journal of Educational Change*, 3, 265–282.
- Koretz, D. (2008). The pending reauthorization of NCLB: An opportunity to rethink the basic strategy. In G. L. Sunderman (Ed.), *Holding NCLB accountable: Achieving accountability, equity, and school reform* (pp. 9–26). Thousand Oaks, CA: Corwin Press.
- Legatum Institute (2010). *The 2010 Legatum Prosperity Index*. Retrieved on March 20, 2011 from <http://www.prosperity.com/country.aspx?id=NO>.
- Leithwood, K., & Earl, L. (2000). Educational accountability effects: An international perspective. *Peabody Journal of Education*, 75(4), 1–18.
- Lie, S., Caspersen, M., & Björnsson, J. K. (2004). *Nasjonale prøver på prøve [National tests on test]*. Oslo: University of Oslo.
- Lie, S., Hopfenbeck, T. N., Ibsen, E., & Turmo, A. (2005). *Nasjonale prøver på ny prøve [National tests on test again]*. Oslo: University of Oslo.
- Lysne, A. (2006). Assessment theory and practice of students' outcomes in the Nordic countries. *Scandinavian Journal of Educational Research*, 50(3), 327–359.
- Malen, B., & King Rice, J. (2004). A framework for assessing the impact of education reforms on school capacity: Insights from studies of high-stakes accountability initiatives. *Educational Policy*, 18(5), 631–660.
- McDonnell, L. (1994). Assessment policy as persuasion and regulation. *American Journal of Education*, 102(4), 394–420.
- McDonnell, L., & Elmore, R. (1987). Getting the job done: Alternative policy instruments. *Education Evaluation and Policy Analysis*, 9, 133–178.
- Mintrop, H., & Sunderman, G. (2009). Predictable failure of federal sanctions-driven accountability for school improvement and why we may retain it anyway. *Educational Researcher*, 38, 353–364.
- Møller, J., Prøitz T. S., & Aasen, P. (Ed.) (2009). Kunnskapsløftet—tung bør å bære? Underveisanalyse av styringsreformen i skjæringspunktet mellom politikk, administrasjon og profesjon. [Education and promotion—a heavy burden to bear? Analysis of the governance reforms in the intersection between politics, administration and professional]. Oslo, Norway: NIFU STEP.

- Newmann, F., King, M., & Youngs, P. (2000). Professional development that addresses school capacity: Lessons from urban elementary schools. *American Journal of Education*, 108(4), 259–299.
- Norwegian Directorate for Education and Training. (2006). *The education mirror 2005*. Oslo: Author.
- Norwegian Directorate for Education and Training. (2008). *The education mirror 2007*. Oslo: Author.
- Norwegian Directorate for Education and Training. (2011). *OECD review on evaluation and assessment frameworks for improving school outcomes: Country background report for Norway*. Oslo: Author.
- Norwegian Ministry of Education and Research. (2001). The Development of Education 1991 to 2000: National Report from Norway. Retrieved on March 20, 2011 from <http://www.regjeringen.no/en/dep/kd/Documents/Reports-and-actionplans/Reports/2001/The-Development-of-Education-1991-to-2001.html?id=277458>.
- Norwegian Ministry of Education and Research. (2004). *Competence for development: Competence development strategy in basic education 2005–2008*. Oslo: Author.
- Nusche, D., Earl, L., Maxwell, W., & Shewbridge, C. (2011). *OECD reviews of evaluation and assessment in education: Norway*. Paris: OECD.
- O'Day, J., Goertz, M., & Floden, R. (1995). *Building capacity for education reform*. New Brunswick, NJ: Consortium for Policy Research in Education.
- OECD. (2006). *Education at a glance*. Paris: Author.
- Olssen, M., & Peters, M. (2005). Neoliberalism, higher education and the knowledge economy: From the free market to knowledge capitalism. *Journal of Education Policy*, 20(3), 313–345.
- Price Waterhouse Coopers & The Norwegian Association of Local Authorities (KS). (2009). *Come closer Oslo*. Norway: The Norwegian Association of Local Authorities.
- Putnam, R. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon and Schuster.
- Ravitch, D. (2000). *Left back: A century of failed school reforms*. New York: Simon & Schuster.
- Romzek, B. S., & Ingraham, P. W. (2000). Cross pressures of accountability: Initiative, command, and failure in the Ron Brown plane crash. *Public Administration Review*, 60(3), 240–253.
- Rothman, R. (1995). *Measuring up: Standards, assessment, and school reform*. San Francisco: Jossey-Bass.
- Schön, D., & McDonald, J. (1998). *Doing what you mean to do in school reform: Theory of action in the Annenberg challenge*. Providence, RI: Annenberg Institute for School Reform.
- Skedsmo, G. (2011). Formulation and realization of evaluation policy: Inconsistencies and problematic issues. *Educational Assessment Evaluation & Accountability*, 23, 5–20.
- Spillane, J., Hallett, T., & Diamond, J. (2003). Forms of capital and the construction of leadership: Instructional leadership in urban elementary schools. *Sociology of Education*, 76(1), 1–17.
- Spillane, J., & Thompson, C. (1997). Reconstructing conceptions of local capacity: The local education agency's capacity for ambitious instructional reform. *Educational Evaluation and Policy Analysis*, 19(2), 185–203.
- Statistics Norway. (2011). *Facts about norway education 2011: Key figures 2009*. Oslo: Author.
- Stein, S. (2004). *The culture of education policy*. New York: Teachers College Press.
- Stone, B. (1995). Administrative accountability in the “Westminster” democracies: Towards a new conceptual framework. *Governance*, 8(4), 505–526.
- Supovitz, J. A. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, 10(2), 211–227.
- Suspitsyna, T. (2010). Accountability in American education as a rhetoric and a technology of governmentality. *Journal of Education Policy*, 25(5), 567–586.
- Telhaug, A. O., Aasen, P., & Mediaas, O. A. (2004). From collectivism to individualism? Education as nation building in a Scandinavian perspective. *Scandinavian Journal of Educational Research*, 48, 141–158.
- United Nations Human Development Program. (2010). *Human Development Index 2010*. Retrieved on March 20, 2011 from <http://hdr.undp.org/en/statistics/>.
- Verhoest, K. (2005). The impact of contractualization on control and accountability in government-agency relations: The case of Flanders (Belgium). In: G. Drewry, C. Grieve, & T. Tanquerel (Eds.) *Contracts, performance and accountability* (pp. 135–156). EGPA/IOS.
- Volkmar, N. (2008). Knowledge and Solidarity: The Norwegian social-democratic school project in a period of change, 1945–2000. *Scandinavian Journal of Educational Research*, 52(1), 1–15.
- Watson, G. (2001). Reason and responsibility. *Ethics*, 111, 374–394.
- Weiss, C. (1995). Nothing as practical as a good theory. In J. Connell, A. Kubisch, L. Schorr, & C. Weiss (Eds.), *New approaches to evaluating community initiatives* (pp. 65–92). Aspen, CO: The Aspen Institute.