Intelligent accountability in schools: A study of how school leaders work with the implementation of assessment for learning

Astrid Tolo[1] · Sølvi Lillejord[2,3] · María Teresa Flórez Petour[4] ·
Therese N. Hopfenbeck[5]

Abstract

In response to accountability systems dominated by external inspections and achievement data, calls are being made for *intelligent* accountability or a new accountability paradigm that focuses on *meaningful* learning, enabled by professionally skilled and committed educators within the system. In such systems, the actors are encouraged to strive for continuous development in learning organisations based on teamwork, distributed leadership, and professional learning communities. School leaders are positioned between district level administrators and teachers in such processes and have the responsibility to secure professional development. Using the implementation of the national program 'Assessment for Learning' in Norway as a case, the article shows that leaders approach professional development differently. Analyses of interviews with leaders from 7 schools reveal three distinct approaches related to how school leaders perceive knowledge. Some school leaders assume that teachers have the necessary knowledge and skills and trust them to manage the implementation process without leader support. Other school leaders distrust teachers' knowledge and skills and assume that the proper knowledge exists outside the school. These leaders seek external support when they meet teacher resistance. In a third approach, school leaders assume that knowledge develops through collaboration and thereby engage with teachers in continuous judgment about the implementation procedure. In the discussion, questions of trust and distrust are analysed in relation to how professional knowledge is developed and how professional discretion can support the development of intelligent accountability in schools.

Keywords Intelligent accountability · Assessment for learning · Implementation · Teacher–leader-relation · School leaders · Professional development

✉ Astrid Tolo
astrid.tolo@uib.no

Extended author information available on the last page of the article

Introduction

Bovens (2010) states that accountability in all sectors is either about *virtues* or *mechanisms* of accountability, which are meant to secure quality of products or services offered by an organisation. In the education sector, practices of accountability appear in different ways. Accountability as securing teacher *virtues* can be seen in teacher appraisal (Flores 2012; Zhang and Ng 2011; UNESCO 2017), value added modelling (VAM) and teacher evaluations (Darling-Hammond 2013; Lillejord et al. 2014). Accountability *mechanisms* are seen in high-stakes testing (Russell et al. 2009) and external inspections (Ehren et al. 2013), but are also evident in more "soft touch" approaches such as networks (Ehren and Perryman 2018) and school self-evaluations (Hall 2017; Ozga 2009), which use a range of sources of evidence for the assessments (UNESCO 2017). To be able to develop schools according to new challenges, demands and knowledge, there is a strong emphasis to *learn* from accountability processes. Clear lines of responsibility in these processes are vital to securing good quality education for all children, as well as knowing when those lines are broken and what actions to subsequently take (UNESCO 2017). This opens for an emphasis on horizontal forms of accountability, social forms that are "suited to induce reflexivity and learning" (Bovens 2010, p. 956).

Over the last decades, vast amounts of summative data collected for accountability purposes have globally put the education sector under pressure. Schools are expected to use this summative information formatively, for improvement purposes (Schildkamp et al. 2017). Research finds, however, that data use is a complex enterprise (Mandinach and Jimerson 2016) and since many teachers and school leaders lack sufficient data literacy, they struggle to analyse and use data to improve practice (Hubbard et al. 2014). Therefore, Datnow and Hubbard (2016) suggest decoupling data use from external accountability demands and, instead, involve a variety of information on student learning. In line with this, Wilson (2017) has argued that classroom assessment is at least as important as large-scale assessments for education. Brown et al. (2017) also suggest a merger between data-based-decision-making and research-informed-teaching-practice.

There is also a growing sense of urgency amongst researchers on the detrimental effects of current accountability reforms' inherent distrust of the teaching profession. Calls are, therefore, being made for a new accountability paradigm that focuses on *meaningful* learning, enabled by professionally skilled and committed educators (Darling-Hammond et al. 2014). Alternative accountability strategies are suggested, for example, to strengthen teachers' professional responsibility (Cochran-Smith et al. 2017). Schools should strive for continuous development as learning organisations (Kools and Stoll 2016) based on teamwork (Abrahamsen et al. 2015), distributed leadership (Nguyen and Hunter 2018), and professional learning communities (Philpott and Oates 2017). The notion of *intelligent accountability* could serve as an umbrella term for such a new accountability paradigm.

The concept *intelligent accountability* came to attention after Onora O'Neills BBC Reith Lectures in 2002 (O'Neill 2002; Cowie and Croxford 2007) and has been taken up in educational literature since then (Andreasen et al. 2013; Baldwin

2004; Cowie and Croxford 2007; Ellison 2012; Hodgson 2011; O'Neill 2013a, b; Ozga 2009; Sahlberg 2010). This has come about as a response to the concerns about an education sector under increasing accountability pressure and a growing distrust in the teaching profession. In a brief theoretical description of the accountability process, Schillemans et al. (2013) present three distinct phases: The first is the process of information gathering, where professionals retrospectively make an account of their performance. The second phase is the reporting and assessment of this account, and the third is judgement of this account by the relevant parties. In more traditional vertical accountability systems these phases are more often than not formal steps, with a one-way flow of information and assessment and are nearly always high-stakes. O'Neill critisises these types of accountability systems, calling them *unintelligen*t since they remove the need for independent and competent judgement (2004). However, in horizontal forms of accountability (Bovens 2010), the phases could be informal and practiced continuously, with the intention to perform internal regulation of the profession. These type of accountability systems can open up for a two-way flow of information, reflexivity and constitute a more dynamic, continually self-improving system, i.e. a more intelligent form of accountability, as described by O'Neill (2002, 2013a, b).

However, there is no unified understanding of what intelligent accountability should actually look like when implemented in the education system. In England, the "new relationship with schools" initiative in 2004, as presented by the then Minister of State for School Standards, David Miliband, emphasised *intelligent accountability* as in self-evaluations, as one of the policy's main selling points (Miliband 2004). Ozga (2009, p. 153), is critical of this policy and the problematic combination of school self-evaluations, development planning and performance management together with "short and sharp" inspections. Referring to the same policy Stobart (2008) points to the *potential* of this new inspection system where "they make provisions for schools to set out their own account as the basis for judgement" and claims this is "a move towards an alternative view of change based on building knowledge in order to improve…" (Stobart 2008, p. 136). Sahlberg (2010), who writes about intelligent forms of accountability, points to the experiences from the Nordic countries where horizontal forms of accountability are emphasised, which suggest that trust and school autonomy, as well as good educational leadership are necessary to strengthen individual and collective responsibilities (Sahlberg 2010).

This paper will add to the discussion with a Nordic example giving insight into informal, horizontal, micro level processes and how school leaders take responsibility and hold teachers to account to secure professional development. The concept of trust is used as an analytic tool. The argument is that an intelligent accountability system (O'Neill 2002) implies that schools must critically investigate their own practices. To do this systematically, we will further argue that the school leader must take the lead in a continuous dialogue with teachers. It is a social and complex process that contains judgements of what good practice is and how it can be developed. In terms of accountability, this dialogic and reflective practice is an internal mechanism that over time should contribute to the building trustworthiness so that stakeholders can trust schools.

Research use different but similar concepts to show evidence on what constitutes trustworthiness in relations between people, such as competence, honesty and reliability (O'Neill 2013b) or ability, benevolence and integrity (Mayer et al. 1995). Building trustworthiness in organisations is based on the same principles but requires systematic work. Grimen (2009) argues that an organisation's internal control mechanisms, and making these mechanisms visible for the public, i.e. transparent, contribute to building external trustworthiness. The point is to allow stakeholders to judge if, when, and why they can trust an organisation, so they can take the necessary kind of precautions. Taken together, these perspectives indicate that trust relies on continuous judgements based on the flow of relevant and understandable information.

The very simple definition of trust used in this paper is: Trust is when a person acts without precaution. Taking precautions is a sign of distrust (Grimen 2009). In this perspective, distrust is not negative. It is necessary. The higher the risk at stake is, the more precautions tend to be taken. For example, when implementing new practices affecting the everyday learning for children in school, the quality of the implementation process is pertinent (Fixsen et al. 2005) and taking some precautions is necessary. O'Neill (2002, 2013a, b) stresses that trust is and should be related to specific situations and relations. Therefore, the question is when and under which conditions we can or should trust other persons and institutions. The overarching aim is to place trust *intelligently* in the trustworthy (O'Neill 2013b).

Leaders must balance trust and control and the need for control appears where there is distrust. Leaders must, for example, know who they can trust and when to be alert. Leaders have the responsibility to oversee that the work is being done in accordance with expectations (Cerna 2014). When a teacher's work is made visible, it can be documented, questioned, assessed and improved. This kind of internal, institutionalised active inquiry serves as internal quality assurance and renders the school trustworthy by assuring for example, that parents can through sound arguments trust the school in question. Institutions that have developed standards for their work and routinely assess and control that those standards are followed can be trusted by the public (Grimen 2012). These organisations do not discourage questions and internal disagreements but perceive them as vital for quality improvement.

The paper draws on empirical data from a case study of a large-scale implementation of Assessment for Learning (AfL) in the Norwegian school system (Hopfenbeck et al. 2013), where 101 stakeholders from all levels of the education system were interviewed. We have re-analysed the qualitative interviews with school leaders that addressed questions with relevance for this paper. School leaders relate to both their administrators employed in the municipality and teachers, and therefore they sometimes balancing different or conflicting interests. The aim has been to gain a more profound understanding of how school leaders perceive their responsibility and act in the implementation process to promote professional development. The research question is: *How do school leaders express trust and distrust to municipality leaders and teachers in an implementation process, and how do they act on these perceptions?*

Assessment for learning

The assessment reform group (ARG) defines Assessment for learning as: "The process of seeking and interpreting evidence for use by learners and their teachers, to identify where the learners are in their learning, where they need to go and how best to get there" (2002, pp. 2–3). According to Stobart (2008), AfL is an "attempt to make assessment a productive part of the learning process" and is seen as an essential part of teaching and learning (p. 144). Thus, AfL should be considered a method for quality teaching, not simply an assessment tool (Stobart 2008; Gardner 2011). The publication of Black and William (1998) Assessment and classroom learning, and a series of publications from the ARG, initiated an international spreading of assessment reforms with the goal to change educational practice (Baird et al. 2014). Therefore, implementation of AfL is an interesting case of professional development as it influences the way we perceive teaching in a fundamental way.

Norway embraced the approach, by adding it to the Education Act[1] in 2009 through four principles, which apply for primary and secondary education:

1. Students should understand what they are supposed to learn and what is expected of them.
2. Students should get feedback that informs them about the quality of their work and their level of achievement.
3. Students should be advised on how to improve their learning outcomes.
4. Students should be engaged in their own learning by assessing their own work and their own learning progress (Hopfenbeck et al. 2013, p. 28).

In 2010, a voluntary national programme was initiated by The Directorate for Education and Training (DET) with the intention to align teachers' assessment practices to the above principles and train key stakeholders (local administrators, school leaders and teachers). Participating municipalities worked actively with schools and assessed their work on goal-setting and the development of implementations plans. Some municipalities hired consultants to drive the process while others chose not to engage in the program and, instead, designed their own AfL implementation strategy.

School culture and the educational context where AfL is introduced impact the take up of the program (Florez 2015; Hopfenbeck and Stobart 2015). This means that both the local and the national context where AfL is implemented does affect how the programme is put into practice and what kind of obstacles and challenges different stakeholders face in the process. Based on the research on AfL, and from implementation of AfL in Norway, it can be seen that tensions between top–down and bottom–up-approaches or accountability versus trust affect stakeholders (Black 2015; Hopfenbeck et al. 2015). It is for example, challenging if top-level bureaucrats, based on a shallow understanding of AfL, define characteristics of good

---

[1] The education act: Forskrift til Opplæringslova: https://lovdata.no/forskrift/2006-06-23-724.

practice and subsequently expect teachers to technically implement this despite protests from both teachers and researchers (Hopfenbeck et al. 2015).

Research states the importance of productive collaboration between students, teachers, and leaders when implementing AfL (Ofsted 2008; Department for education and skills 2007; Condie et al. 2005; Hayward and Spencer 2010; Hopfenbeck et al. 2015; Kirton et al. 2007; Webb and Jones 2009; Kellard et al. 2008). It is broadly assumed that leadership and knowledge sharing within and between schools is crucial for successful implementation (Department for education and skills 2007; Kellard et al. 2008; Ofsted 2008). Few studies have, however, empirically investigated school leaders' engagement in the implementation of AfL (Prøitz et al. 2017).

In a systematic review of prerequisites for successful implementation of AfL, Heitik et al. (2016) found collaboration, teacher autonomy, and less pressure from the accountability system to be the most important factors. Based on this finding, we may assume that conditions for implementing AfL are good in Norway, where the accountability system is labelled as 'soft' (Hall 2017) and with a high level of trust (OECD 2017). Hudson (2011), however, warns that assumedly soft systems may be as powerful as more direct control methods. Paulsen and Høyer (2016) find that a system can be soft in terms of control by testing but still exercise substantial control of school leaders' and teachers' behaviour through standardised training programs.

Culture of trust in Norwegian education system

It is generally assumed that trust in teachers' professional judgment is integral to the education system (Barber 2002), albeit to a varying degree in different countries (Czerniawski 2011). There is a vast amount of studies on different aspects and levels of trust in education (for example, Cerna 2014; Gray et al. 2016; Handford and Leithwood 2013; Harris and Jones 2018; Louis 2007; Tschannen-Moran and Gareis 2015). We will in this section limit ourselves to focus solely on how a culture of trust is evident at the system level of Norwegian education.

One trait of trusting cultures is, according to O'Neill (2014), that they only selectively and sparingly use sanctions, as is the case in Norway. Educational provision in Norway is largely based on trust to teachers and local initiatives, and these values are embedded in the education system. However, soft accountability systems also have inbuilt control mechanisms and tensions can emerge.

Norway has national tests, but no national test-based accountability system, and little high-stakes testing. The evaluation system relies on trust to teachers and is characterised as "soft-touch" (Hall 2017). Students in primary education (year 1–7) receive qualitative feedback from the teacher but no grades. Students' first externally rated exam is in grade 10 (age 15), when they have one written and one oral exam in a randomly selected subject. Children aged 6–14 are evaluated and graded by their teachers. There is no national system for teacher assessment but initiatives have been made to establish agreement on certain principles for teacher evaluation (Lillejord et al. 2014, 2018).

Norwegian education policy is developed nationally and 422 municipalities (primary education) and 18 counties (upper secondary) are responsible for local

adaptation, reform implementation, and quality assurance in schools. In the Norwegian quality assessment system (NQAS) for education, the County Governor (CG) represents the state in two respects. Firstly, the CG regularly inspects selected schools on specific requirements in the Education Act and secondly acts as appellate authority in cases of complaints. Schools are accountable to the school owners (municipalities or county councils), school owners are accountable to the CG, and the CG to the Ministry of Local Government and Administration.

The national level also conducts normative steering through different kinds of programs implemented by the Directorate of Education and Training (DET) (Paulsen and Høyer 2016). Previous research on the implementation of AfL in Norway (Hopfenbeck et al. 2015) noted that the DET was actively engaged in the implementation process while simultaneously responsible for the implementation of the programme. DET was criticised for having a surface interpretation of AfL and for having developed a standard 'correct' AfL practice. Schools, therefore, risk being corrected in inspections for not practicing AfL 'correctly'. At the time of the interviews, this controversy about what was the 'correct' AfL practice challenged school leaders and teachers, who felt they had to balance contradictory messages.

Research on implementation of AfL and school leadership

Complex and critical approaches to policy enactment have highlighted the way in which teachers tend to resist, translate and interpret external initiatives (Houghton et al. 2015; Maguire et al. 2018). They do not immediately trust new ideas from the administration and do not always see how new working procedures align with their beliefs (Rubie-Davies et al. 2015) or improve their pedagogical practice. When tested against their experience, teachers may find new requirements counterproductive.

Research on implementation of AfL argues that resistance normally occurs and must be handled professionally to be constructive for the implementation process (Hermansen 2014; MacPhail and Halbert 2010; Department for education and skills 2007; Kirton et al. 2007). Researchers highlight the crucial role of school leaders. Black et al. (2004) argue that teachers who want to change should not invest effort under unfavourable conditions. Policy reports about AfL emphasise school leadership and the role of senior staff as a key factor in implementation as well as stating the importance of distributed leadership as ideal for implementation of AfL (Department for education and skills 2007). Motivation, commitment and enthusiasm of senior staff taking the programme forward is seen as critical (Kellard et al. 2008). Whole school visions of teaching, learning, and assessment, driven by senior staff, preferably headteachers, were the most important factor for effective AfL practice (Ofsted 2008). Trust is implicitly assumed in phrases like high expectations from the leaders, support for knowledge and practice development and knowledge sharing within and between schools (Kellard et al. 2008; Department for education and skills 2007; Ofsted 2008). Leaders should stir emotions through "winning the hearts, minds and belief of all staff" (Department for education and skills 2007, p. 43).

Apart from these policy reports, few studies have empirically investigated AfL implementation and leadership. One study (Davies et al. 2014) draws on experiences from 8 jurisdictions from different parts of the world, when studying how Assessment for Learning can be used both as a strategy for change and the object of change. The researchers have interviewed one leader from each jurisdiction, responsible for many students (ranging from 350 to 185,000). All leaders were expected to use AfL as a leadership tool and have strived to implement AfL for years. For over a decade, they also attended the same conference on AfL. Davies et al. (2014) find that "these positional leaders described actions that go beyond 'servant leadership' and 'distributed leadership' and what is typically thought of as 'instructional leadership'…" (p. 586) and conclude that AfL is a powerful tool for student-, teacher-, school-, and system learning.

In Norway, team leadership is frequently found (Abrahamsen 2018; Eide 2015), and described as a fine balance where the school leaders are placed in a continuous tension of building trusting relations with and between the employees (Helstad and Møller 2013). Expectations and demands from stakeholders, new knowledge from research, and contextual changes force the school leaders to establish new kinds of relations with the teachers and make every one of them responsible not only for their own work with the students but also for the collective work. To manage this balance, school leaders must have deep knowledge about schools as organisations and systems, communication skills, and the courage and capacity to lead in situations of resistance (Andreassen et al. 2010).

This discussion around AfL implementation and leadership will be addressed in connection to the interpretation of empirical results.

Method

Sample and data

Data for this paper is drawn from a case study of the implementation of Assessment for Learning in Norway (Hopfenbeck et al. 2015). During the implementation process, several actors were interviewed, including OECD employees, and representatives from the Norwegian government, schools and municipalities. Interviewees were chosen using a combination of strategic sampling of informants, such as Ministers of Education, bureaucrats from the Directorate for Education and Training (DET), researchers, and convenience sampling of six municipalities and nine schools representing urban and rural schools, as well as small and large municipalities (ranging from 8000 to 275,000 inhabitants). The first and fourth author conducted 56 interviews with a total of 101 participants. For the purpose of this study, seven school leader interviews were re-analysed (See Table 1) and context information was retrieved from previous interviews with teachers and municipality leaders and in some cases used for confirmation and comparison of perspectives.

The interviews were semi-structured and different themes with relevance for implementation of AfL were elaborated in each interview. The paper reports from 7 of the 9 schools we visited because these are the interviews where relevant

Table 1 Participating schools and interviewed school leaders

	Interviewed
School A	School leader
School B	School leader
School C	School leader and department leader
School D	School leader
School E	School leader
School F	School leader
School G	School leader and leader group

information for this study occurred. For simplicity, the schools were named A–G. (The logic of the order is shown in Table 3). All interviews were recorded and transcribed ad verbum to facilitate data analysis.

Data analysis

In the main study, we chose content analysis to reduce the amount of data into organised segments (Silverman 2006). One of nineteen coding categories in NVivo was labelled: "Trust among the participants, respect, dialogue, consensus, shared understanding". However, there was a requirement to go through the full transcripts of interviews from the school leaders again, since the theoretical framework of this particular study might open up for more segments to be analysed. The creation of condensed text based on segments from the interviews is theory informed, but also inductive (Silverman 2006). Therefore, we worked through several inductive and theoretical phases as in hermeneutical analyses.

The inductive phase: We sorted the text from the interviews in categories. The two main categories are school leader's descriptions of relations to teachers and relations to municipality leaders. We then identified reported actions, how the leaders interpreted the outcome of the actions they undertook, and their reflections about the process. We looked for segments in the interviews where school leaders' relations to teachers and to municipality leaders were mentioned and described e.g. through narratives of how they organised work and stories about incidents.

The theoretical phase: A theory of action perspective on trust (Grimen 2009; O'Neill 2002, 2013a, b) suggests that we should look for trust in relations. To grasp this perspective, we focused on what a school leader reported doing to implement AfL, and how actions of trust and dis-trust to different stakeholders play out. In what situations do, for example, the school leaders report that he or she acts with precaution? We also looked for instances of, for example, routines that were set up to secure quality and progression of the work and how for example teachers perceived these routines.

The first author read through the interviews, marked all text that described relations (inductive) and could be related to the issue of trust (theory informed), and then wrote a condensed version of these segments. The fourth author read the

condensed text and secured the quality of the interpretations according to the interviews. The presentation of results is based on a systematisation of the condensed text.

Presentation of results

In the results section, the school leaders' relations with the municipality administrators, are first explored. Then, certain aspects of trust in relations between school leaders and teachers are investigated from the leaders' perspective.

In Norway, 422 municipalities are responsible for primary education (grade 1–10) and 18 county councils are responsible for upper secondary (grade 11–13). The size of the municipalities varies from 201 to 666 759 inhabitants. In this study, we have data from small (8000 inhabitants) to mid-sized municipalities, and also one of the largest (275,000).

Municipality size influences the relationship between schools and municipality administration. Small, rural municipalities have few school administrators. Sometimes one person is responsible for all the schools in a small municipality. By contrast, the largest urban municipalities have 100–200 employees committed to administrative and professional support. Not all school leaders can expect support from the municipality in a specific area (like, for example, the implementation of AfL). They can, however, call for assistance in the implementation process from other municipalities, the DET or professional consultants.

Table 2, below, gives an overview over the various municipal contexts and the interviewed school leaders perceptions of demands, support, control and trust in their relation to the municipality level.

According to O'Neill (2013b), people trust those who they perceive as competent, reliable and honest. The excerpts that follow from the interviews show how school leaders—in their different contexts—trust or distrust administrators at the level above them. A central theme crossing through the interviews is how school leaders judge the competence amongst municipality administrators and what level of demands and support they can expect. This is relevant for intelligent accountability because when taking responsibility for a school's development, the school leaders need to know what kind of demands they have to meet, what capacity they can draw on, and what kind of obstacles there are in the way to reach the goal. In this respect, they need to know who are the trustworthy allies in the municipality administration and how and with what they can help.

The two school leaders in school C and D, both located in small municipalities, said they had experienced demands and expectations in the Education Act (on AfL) in 2009 as challenging. Early in the project they had little or no support from the municipality administration but in 2011 external consultants were hired to support schools. The leader in School D was very pleased with the involvement in the program and had high expectations for what could be achieved. She appreciated the possibility to join the AfL-program, where meetings were organised for school leaders, and did not expect any more support from the municipality. Also the leader in school C was positive but expected little support from

Table 2 How school leaders express trust relations to the municipality level

	Interviewed	Description of trust relation
School A Medium sized municipality	School leader	Specific demands and control. Reciprocal distrust between school leader and municipality leaders
School B Medium sized municipality	School leader	Relation to municipality not discussed
School C Small sized municipality	School leader and department leader	Generally, few demands and little support, but some concerning AfL because of attendance in the national program
School D Small sized municipality.	School leader	Generally, few demands and low support, but some concerning AfL because of attendance in the national program
School E Large sized municipality	School leader	High competence at the municipality level and generally mutual trust. However—in the case of AfL the leader expressed reservations on the role of the municipality
School F Medium sized municipality	School leader	Relation not discussed
School G Large sized municipality	School leader and leader group	High competence and general, mutual trust

the municipality after the DET program had finished. "*There is generally, no follow up,*" he says. The interview with the municipality leader confirmed this. The leader in school H, from a mid-sized municipality, had similar experiences, and says: "*This is the problem with state-initiated projects. Some municipalities have capacity to handle it and support schools professionally, while others don't*". They all mention informal networks between school leaders as vital professional support.

The school leader in school A had a different perspective. He had been school leader in a small municipality, where he appreciated the possibility to develop AfL practices with teachers and local administrative staff members. When he moved to a larger municipality with a bigger administration and support from consultants, his motivation disappeared; "*I felt I became more passive*", he explained. The municipality administration focused on national test results and did not initiate dialogues with the schools. While he tried to be involved in the development of AfL practice, neither his initiatives were appreciated or his competence acknowledged. Instead, he felt he was supervised from above. This was difficult as he did not trust their competence. On the other hand, the municipality administration did not want the implementation process to be "privatised" in the sense that school leaders should work alone, they said. Administrators, therefore, observed teachers, supervised teachers, developed "to-do-tasks" for the teachers, and mentored school resource groups. The school leader felt that the administration's interpretation of AfL was superficial, and wrong and he experienced the municipality level follow-up strategy as micro-management, control, and distrust. The school leader also felt that they overestimated their own competence.

In schools E and G, both located in a large municipality, the school leaders trusted the municipality administration's competence, routines and supportive follow up. When talking about problems in the process, competence was not an issue. In the interviews with municipality leaders, however, it was revealed that they did not trust the DET program. They felt they were abreast on AfL, knew how to do it and who to contact for assistance if needed. They were self-confident about their implementation skills and felt that they also could assist smaller municipalities in the region if they were asked. While this municipality also had challenges, which will be explored in the next section, the general impression was that the school leaders trusted the municipality administrators and the actors held each other intelligently to account.

School leaders experience municipal support and demands differently and the challenges that leaders experience are context dependent (Moos et al. 2008). These data show how some school leaders rely on the competence and capacity at the administrative level; others don't. In one case (School A) we found that the school leader felt limited by municipal leaders who interpreted AfL differently and had other priorities than he had, while school leaders in other municipalities appreciated support from the administration. School leaders who had no support from the municipality sought assistance from other school leaders. In the next section, these various experiences schools have with the municipality level is used as background information when analysing leader–teacher relations characterised by trust, distrust and, the balance between trust and distrust.

Trust and blind trust in school A, B and C

In three of the interviews with school leaders (schools A, B, and C), there was no mention of precautions taken towards teachers and how they worked when implementing AfL. Apparently, implementation went smoothly. The leader in school A talked about how they worked as a team and collectively found practical solutions in the implementation of AfL. In school B, the leader had delegated the responsibility for AfL implementation to the school's team leaders. He said he instructed the team leaders, told them what he expected them to do, and emphasised how important their work was for the implementation. He provided no specific information about his own follow up. Teacher interviews, however, revealed that leaders "… *check if we have done what we are supposed to do, because everything is on the computer*…" and if something was not done properly, the leaders asked why. This was described as a useful approach to all development projects, including AfL. In school C, interviews with the leader indicated a high level of trust in teachers. He referred to teachers' work, mentioned materials they had developed, and how some teachers used the DET webpage to learn about AfL. He apparently was in control, and knew what teachers were doing since teachers' work was transparent. Teacher interviews from this school confirmed that they worked autonomously with AfL and the leader did not tell them what to do or control what they did.

Interviews with school leaders in schools A, B and C, provide three examples of how leaders trusted, worked with, instructed, or just followed the teachers' implementation of AfL. According to the teachers in school B, their middle leaders had some critical comments, but apart from this, distrust was not expressed by any of these leaders and our data gives no indication if this is a sign of good implementation and trustworthy teachers, or if the school leader placed trust in the teachers without actively judging if they were trustworthy. At least from what they say, the leaders in these three schools did not interfere directly and did not actively control.

How relations of distrust complicate implementation in school D and E

In two other interviews, leaders from schools D and E, reported far more precautions in leader–teacher relations. These leaders did, in certain respects, trust their teachers, but frequently met teacher resistance and lack of motivation and therefore, struggled to organise the collective work and keep the implementation process on track.

The school leader in school D, who had high expectations of the programme, said that the school had been working on AfL since it was introduced in the Education Act in 2009. However, with no demands or support from the municipality they gained little success and the school leader felt left alone. At the time of the interview, they had recently joined the DET program. According to the school leader, some teachers resisted the work while others could not wait to get started. Some teachers prepared lessons with the students while others took the opportunity to "*concentrate on their planning, which they preferred to do in their own way*", as she phrased it.

According to the school leader, teachers who resisted did not trust the program or the leaders who wanted them to change their practices. When asked to explain why teachers took these precautions, she said that teachers might fear being exploited by being expected to put in more work hours. Early in the implementation phase, an assessment routine was established in School D. When a researcher who visited the school told the teachers that such a routine was no prerequisite to succeed with AfL, the process came to a halt. The school leader said she then realised that the teachers distrusted their own leaders, ignored their contribution, and instead ascribed authority to the researcher and the DET program consultants. In one sense, this school leader was in control but the schools' implementation mechanisms seemed weak and her leader strategy, she said in the interview, was to seek support from an external authority that teachers might trust or—at least—listen to.

The school leader in school E had tried to understand the theoretical underpinnings of AfL. Her ambition was that the teachers should get a profound understanding of AfL but the school never got to this point. The teachers had read and discussed a book about AfL and the school leader gave them additional time to work with students. In the interview, she problematised the demographics of the staff. Many teachers had been there for decades, and the few recently appointed teachers were rapidly absorbed by the established school culture. The school leader referred to the "union thing" in a not very favourable way. The "union thing" occurred when teachers resisted leader initiatives and did not want to engage in collective development efforts unless it had the union's explicit support.

She referred to an episode when the municipality arranged a competence development project on AfL. The school leader sent two teachers, who after the training felt no commitment to share the knowledge they had gained with their colleagues. They argued that municipality representatives had told them they had no responsibility for this kind of sharing. This became, according to the school leader, a "union thing", and she was left alone with the responsibility for AfL, quite contrary to what she expected when she sent teachers to the training. She was very disappointed with the teachers and the municipality.

The school leader interpreted the resistance as integral to how teachers perceive their work. *"When a child doesn't succeed, the teachers claim it is something wrong with the child, not the teaching"*, she explained. Among her AfL-initiatives, one initiative was to give each teacher one additional hour each week to be used for learning dialogues with the students. She suspected, however, that this hour was not used as intended and added: "*I must follow that up. I have tried to be clear about my expectations, but when you think you are clear, you realise you are not*". In hindsight, she realised that additional time vanishes in a busy day: "*The principal or the department head cannot 'give' teachers time and tell them how to use it*". Student interviews confirmed that they sometimes talked with their teacher alone, but they were quite unclear about the content of these dialogues. The teachers, on the other hand, complained about heavy workloads. This example of distrust might be related to the school culture or the school leader's approach. She had a clear vision for what she wanted to achieve but struggled in the implementation. At the time of the interview, she had, in a way, become resigned and was waiting for the municipality to help her solve the problem with teacher resistance. This strategy resembles the one

in school D. Both school leaders sought support from an external authority that they were hoping teachers might trust or—at least—listen to.

Leadership: Balancing trust and challenge in school F and G

Two other schools (F and G) report a quite different approach to implementation challenges.

In school F, development projects were team-organised. Teams met once a week and plans were made for the whole semester. Once a month, teachers had learning dialogues with students; a process that was well integrated in the school culture. The school leader emphasised the importance of keeping activities focused as the school is in constant change; "*Because teachers cannot handle too much change*", he said. In the process, he asked teachers to keep it simple because "*motivation is also about not having too much to do at the same time and to leave some periods calm*". This leader did not problematise teachers' resistance or refer to concrete episodes. When outlining how the school developed AfL he presented a structured, almost technical, implementation approach. From how he described the work, it can therefore be assumed that he thought a rigid structure would facilitate the implementation of AfL.

In school G, the entire leader–team was interviewed. Their implementation strategy resembled the one chosen by school F but here, the leader explicitly addressed resistance and focused on action and reflection. The process had lasted for 6–7 years, required substantial efforts and pressure from the leader and leader–team at all levels, and had had several setbacks. As a team, they had tried to be very clear and persevering. The school leader presented a rigid system of breaking goals into sub-goals and developing assessment criteria to support new assessment approaches. The leader asked teachers to follow schemes and fixed routines as he did not trust them to change practice based on what they knew about AfL. The Education Act requires documentation and the rigid system was, as he said, his way of "feeding the documentation ghost". Documentation was a problem for the teachers but collectively and over time they had moved away from the rigid practice of "breaking down goals" and found an approach they experienced as more meaningful. At the time of the interview, AfL met less resistance but the teachers still found the work difficult and felt tired of it. The leader–team in school G recognised how time consuming the work was for the teachers and understood their resistance. They tried to focus on very practical AfL solutions in the classroom because this was what the teachers preferred.

The leader–group in school G motivated teachers to experiment with assessment practices, expected substantial teacher effort in the implementation, and opened a room for 'trial and error'. Instead of focusing on what teachers did not achieve, the leaders followed up and supported them. There appeared to be a high level of trust between the leader–team and the teachers. This was confirmed in interviews with teachers and students. Frustrated leaders from other schools came to school G to ask for advice, they said.

When school leaders in schools F and G met teacher resistance, they acknowledged the challenges and listened to the critique and adjusted plans. However, they still kept focus on the systematic work without delaying the implementation process. The two school leaders did not expect teachers to enthusiastically change their practice. They argued that teachers expect practical solutions to complex problems and need someone to help them sort out what is important and what can wait. In the interviews, teachers said they perceived these approaches as reasonable. The question is, however, if the practice of providing teachers with externally developed practical solutions to complex problems, might weaken their professional discretion and undermine their autonomy.

Summing up leader: Teacher relations

Data from the main study (Hopfenbeck et al. 2015) shows that in schools F and G, where the school leaders chose to actively trust and challenge teachers, students reported that their teachers worked in accordance with principles from AfL. Therefore, we concluded that the implementation in these schools was successful. In the two other approaches: (1) schools A, B and C, where leaders trusted teachers blindly and (2) schools D and E, where leaders distrusted teachers, implementation was less successful. In the discussion, this is highlighted as important background information.

In the implementation of AfL in schools, the interviewed school leaders saw it as their main responsibility to work with the teachers, in accordance with the Education Act. Based on available data about leader–teacher interactions, three approaches to taking responsibility for the implementation of AfL were identified. Leaders in schools A, B, and C were inclined to trust the teachers' knowledge, competence, and work processes. They took no precautions in the process of implementing AfL and appeared to blindly trust the teachers. In the second approach, leaders from schools D and E referred to incidents of distrust and suspicion and explained that their solution was to ask an external authority to tell teachers how they should implement AfL. In schools D and E, a conflict ridden and frustrating implementation process stopped when the leaders abdicated because they did not feel competent to argue convincingly for their decisions. In the third approach, (schools F and G), school leaders had a knowledge-based, inquiring approach to the implementation process. They worked collaboratively with teachers, inquired into and assessed practices, and motivated teachers to argue for their judgments and choices of method. This is visualised in Table 3.

Discussion

This case-study was conducted to answer the question *How do school leaders express trust and distrust to municipality leaders and teachers in an implementation process, and how do they act on these perceptions?* The data analysis has shown that the municipal context the leaders work within is very different and the level

Table 3 Trust, knowledge and leaders acts in the implementation process

Trust, knowledge and action	Schools A, B, C	Schools D and E	Schools F and G
leader–teacher relations	Trust teachers blindly	Distrust teachers	Balance trust and challenge
Knowledge	Exists internally	Exists externally	Existing knowledge must be challenged
How leaders act in the implementation	Passive, leave the work to teachers	Seek external support	Collaborate with teachers

of support and demands from municipality administration varies. Also school cultures vary. The seven school leaders, dependent on how they perceived knowledge and how they positioned themselves in relation to the teachers, chose three different approaches to the implementation process. One approach was to blindly trust teachers' knowledge and implementation skills. A second approach was to distrust teachers and instead seek external administrators' knowledge and competence. The third approach was to both trust and challenge teachers' knowledge with the intention to strengthen the teachers' professional discretion.

Educational reforms are typically initiated at the national level and local administrators are expected to secure effective implementation. According to Ball et al. (2012), policy implementation should be perceived as ideas based on certain principles that are reinterpreted at different levels of the system to create new forms of practice. For this to succeed, school leaders must understand their role not merely as facilitators but also as *contributors* to the work by being involved in the process and challenging teachers when needed. Interestingly, the three approaches to implementation identified in this study indicate how the school leaders position themselves in relation to the teaching profession—as outside the profession (schools A, B, C, D and E) or as responsible participators in the profession's knowledge-building process (schools F and G).

While the first approach, where school leaders trust the teachers blindly, may be perceived as somewhat withdrawn and naïve, it is in accordance with the theory of educational organisations as loosely coupled (Weick 1976). One characteristic of loosely coupled organisations is division of labour (Ingersoll 1991) and school leaders and teachers have a long tradition of working relatively independently from each other. In this tradition, school leaders can leave the responsibility for the professional work to the teachers where it rightfully belongs.

Researchers have found that school leaders historically have distanced themselves from the teachers by developing a role as professional bureaucrat or administrator (Scheerens 2012; Mehta 2013). If school leaders identify with the administration, administrative knowledge and ethos prevail in schools (Mehta 2013) where teachers' professional knowledge base is weak (Hermansen 2017). A possible countermove on the accountability reform movement is, therefore, to intelligently strengthen the teaching profession through closer connecting teachers, school leaders, and researchers through dialogues about what is important in schools and how this can be practiced and assessed.

This case from the Norwegian trust-based system reveals that even educational systems with little external control have inbuilt quality control mechanisms (Bovens 2010). Interviews reveal that school leaders, teachers and municipality administrators 'softly' hold each other to account. The study shows, however, a variance in how systematically school leaders work with implementation. Based on how they judge the situation and perceive their responsibility for the result, they make personal choices. The professional knowledge base for these judgements, is, however, not explicit.

School leaders who blindly trust the teachers, rely on teachers' initiatives and teachers' knowledge. In doing so, however, they abdicate as leaders (Hult et al. 2016). It is difficult for teachers, who have busy days teaching students, to develop

professionally without leader support and clear priorities and expectations (Honingh and Hooge 2014). Professional work develops through collective work and a joint knowledge base that also leaders adhere to.

Leaders who continuously distrust teachers undermine teachers' professional learning by assuming that teachers lack knowledge and by rejecting the value of teachers' experience-based knowledge. While the blind trust approach creates a room for professional development, an abdicated leader cannot successfully support the teaching professions' collective collaboration. In the first two cases, implementation efforts are dependent on personal knowledge and random and informal teacher initiatives. The third, balanced approach is, however, worth exploring in terms of intelligent accountability. Here, leaders actively engage in teachers' professional learning, assess and judge when teachers are on the right track and abreast of the situation, and when precautions must be taken. The success of this approach requires that certain school leaders work closely with teachers. First, professional knowledge draws on various knowledge sources—a blend of personal knowledge, experiences shared within the professions and amongst colleagues, summative data and research. Second, professions typically collaborate, judge their own and each other's work informally and related to standards, to secure the quality of the services they offer to the public (Lillejord and Børte 2019). To become a continuous part of the profession's work and a prerequisite for their public trustworthiness, such collective work processes need leadership. Third, professional teachers refer to their research- and experience-based knowledge when they publicly argue for what they do, how and why they do it. This is how professions use their discretion to show trustworthiness (Grimen 2008).

Conclusion

The study has analysed how school leaders relate to accountability expectations and act in an implementation process. The study drew on a case study of implementation of Assessment for Learning and contributes to research about how school leaders perceive knowledge and position themselves in relation to teachers. Interviews with seven school leaders were re-analysed and context information retrieved from previous interviews with teachers and municipality leaders. The interview data revealed that relations of *trust* contribute to explaining why school leaders, who are expected to mediate reform initiatives, choose to act as they do in the implementation process.

Norway has no test-based accountability system and the education system is based on a high level of trust. This does *not* mean that all actors trust each other no matter what, but that trust in the local schools is perceived as a prerequisite for a successful system. The analysis of the interview data has nuanced this general notion of trust in the Norwegian educational system. While some school leaders trust their teachers blindly, other school leaders instead trust national or district level authorities to have the necessary knowledge and authority to assist them when they meet teacher resistance. A third approach identified is when school leaders trust, support and encourage the teachers to use their professional discretion, while they simultaneously challenge the teachers' professional competence.

Intelligent accountability requires the possibility for stakeholders to place trust intelligently in the trustworthy. Professions typically build societal trust by having systems in place for quality control and continuously assess and improve their work. School leaders who blindly trust their teachers are no hindrance to teachers' professional learning but rely on teachers to take the lead in their own professionalisation efforts. School leaders who mainly distrust teachers, identify with and prefer the authority of the administration will not contribute to increased professionalism in schools. On the contrary, they confirm the distrust inherent in the accountability reform movement. The most promising approach is the third, chosen by leaders in schools F and G where leaders and leader–teams worked closely and collaboratively with teachers and placed trust intelligently. The leaders challenged the teachers' arguments, inquired into their practices and motivated them to argue for their choices of method. Teachers were motivated to experiment and innovate. This is an example of how leaders and teachers shared the responsibility of transforming practice as in horizontal forms of accountability (Bovens 2010) but the school leader took the lead and main responsibility when it was deemed necessary. In order to develop a more intelligent accountability system, this kind of leadership could offer a meaningful way forward.

References

Abrahamsen, H. (2018). Redesigning the role of deputy heads in Norwegian schools–tensions between control and autonomy? *International Journal of Leadership in Education, 21*(3), 327–343.

Abrahamsen, H., Aas, M., & Hellekjær, G. O. (2015). How do principals make sense of school leadership in Norwegian reorganised leadership teams? *School Leadership & Management, 35*(1), 62–78.

Andreasen, K. E., Rasmussen, A., & Ydesen, C. (2013). Standardized testing. In Ainsworth, J. (Ed.) *Sociology of education* (pp. 737–740). Thousand Oaks: Sage Publications, Incorporated.

Andreassen, R. A., Irgens, E. J., & Skaalvik, E. M. (2010). *Competent school leadership*. Bergen: Fagbokforlaget.

Assessment Reform Group. (2002). *Assessment for learning. 10 principles*. Cambridge: University of Cambridge.

Baird, J., Hopfenbeck, T. N., Newton, P., Stobart, G., & Steen-Utheim, A. T. (2014). *State of the field review: Assessment and learning*. Oslo: Report for the Norwegian Knowledge Centre for Education.

Baldwin, A. (2004). Enhanced accountability for electronic processes. In *Trust management* (pp. 319–332).

Ball, S., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools* (1st ed.). New York: Routledge.

Barber, M. (2002). A new context for accountability and inspection. In A. Pollard (Ed.), *Readings for reflective teaching*. London: Continuum.

Black, P. (2015). Formative assessment—An optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice, 22*(1), 161–177.

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Black, P., et al. (2004). *The nature and value of formative assessment for learning (draft)*. London: King's College.

Bovens, M. (2010). Two concepts of accountability: Accountability as a virtue and as a mechanism. *West European Politics, 33*(5), 946–967.

Brown, C., Schildkamp, K., & Hubers, M. D. (2017). Combining the best of two worlds: A conceptual proposal for evidence-informed school improvement. *Educational Research, 59*(2), 154–172.

Cerna, L. (2014). *Trust: What it is and why it matters for governance and education*. OECD Education Working Papers No. 108. OECD Publishing. http://dx.doi.org/10.1787/5jxswcg0t6wl-en. Retrieved 10 October 2015.

Cochran-Smith, M., Baker, M., Burton, S., Chang, W. C., Cummings Carney, M., Fernández, M. B., et al. (2017). The accountability era in US teacher education: Looking back, looking forward. *European Journal of Teacher Education, 40*(5), 572–588.

Condie, R., Livingston, K., & Seagraves, L. (2005). *Evaluation of the assessment for learning programme: Final report*. Glasgow: University of Strathclyde.

Cowie, M., & Croxford, L. (2007). *Intelligent accountability: Sound-bite or sea-change?*. Edinburgh: University of Edinburgh.

Czerniawski, G. (2011). Emerging teachers–emerging identities: Trust and accountability in the construction of newly qualified teachers in Norway, Germany, and England. *European Journal of Teacher Education, 34*(4), 431–447.

Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York: Teachers College Press.

Darling-Hammond, L., Wilhoit, G., & Pittenger, L. (2014). Accountability for college and career readiness: Developing a new paradigm. *Education Policy Analysis Archives, 22*(86), 1–34.

Datnow, A., & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change, 17*(1), 7–28.

Davies, A., Busick, K., Herbst, S., & Sherman, A. (2014). System leaders using assessment for learning as both the change and the change process: Developing theory from practice. *The Curriculum Journal, 25*(14), 567–592.

Department for education and skills (2007). *Secondary national strategy for school improvement. assessment for learning & schools project report*. http://dera.ioe.ac.uk/7600/1/1f1ab286369a7ee24df53c863a72da97-1.pdf Retrieved 02 November 2018.

Ehren, M. C., Altrichter, H., McNamara, G., & O'Hara, J. (2013). Impact of school inspections on improvement of schools—Describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability, 25*(1), 3–43.

Ehren, M., & Perryman, J. (2018). Accountability of school networks: Who is accountable to whom and for what? *Educational Management Administration & Leadership*, *46*(6), 942–959.

Eide, H. M. K. (2015). Narrating the relationship between leadership and learning outcomes. *University of Bergen*. https://doi.org/10.1080/01442872.2014.971729.

Ellison, S. (2012). Re-thinking the concept of "accountability" in the popular discourse of education policy. *Journal of Thought*, *47*(2), 19–41.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M. & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa: University of South Florida, The National Implementation Research Network (FMHI Publication #231).

Flores, M. A. (2012). The implementation of a new policy on teacher appraisal in Portugal: How do teachers experience it at school? *Educational Assessment, Evaluation and Accountability, 24*(4), 351–368.

Flórez Petour, M. T. (2015). Systems, ideologies and history: A three-dimensional absence in the study of assessment reform processes. *Assessment in Education: Principles, Policy & Practice, 22*(1), 3–26.

Gardner, J. (2011). *Assessment and learning* (2nd ed.). Thousand Oaks: Sage.

Gray, J., Kruse, S., & Tarter, C. J. (2016). Enabling school structures, collegial trust and academic emphasis: Antecedents of professional learning communities. *Educational Management Administration & Leadership, 44*(6), 875–891.

Grimen, H. (2008). Profesjon og kunnskap. In A. Molander & L. I. Terum (Eds.), *Profesjonsstudier*. Oslo: Universitetsforlaget.

Grimen, H. (2009). *What is trust?*. Oslo: Universitetsforlaget.

Grimen, H. (2012). Gode institusjoners betyding for tillit (How institutions build trust). In: H. Skirbekk, & H. Grimen (Ed.), *Norwegian trust*. Res Publica.

Hall, J. B. (2017). "Governing by templates" through new modes of school inspection in Norway. *Journal of Educational Change, 18*(2), 161–182.

Handford, V., & Leithwood, K. (2013). Why teachers trust school leaders. *Journal of Educational Administration, 51*(2), 194–212.

Harris, A., & Jones, M. (2018). Why context matters: A comparative perspective on education reform and policy implementation. *Educational Research for Policy and Practice, 17*(3), 195–207.

Hayward, L., & Spencer, E. (2010). The complexities of change: Formative assessment in Scotland. *Curriculum Journal, 21,* 161–177.

Heitik, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review., 17*(2016), 50–62.

Helstad, K., & Møller, J. (2013). Leadership as relational work: Risks and opportunities. *International Journal of Leadership in Education, 16*(3), 245–262.

Hermansen, H. (2014). Recontextualizing assessment resources for use in local settings: Opening up the black box of teachers' knowledge work. *The Curriculum Journal, 25*(4), 470–494.

Hermansen, H. (2017). Knowledge relations and epistemic infrastructures as mediators of teachers' collective autonomy. *Teaching and Teacher Education, 65,* 1–9.

Hodgson, G. (2011). *Creating a climate of intelligent accountability in all-through academies*. Research Associate. Full Report. National College for School Leadership.

Honingh, M., & Hooge, E. (2014). The effect of school-leader support and participation in decision making on teacher collaboration in Dutch primary and secondary schools. *Educational Management Administration & Leadership, 42*(1), 75–98.

Hopfenbeck, T., Flórez Petour, M. T., & Tolo, A. (2015). Balancing tensions in educational policy reforms: Large-scale implementation of Assessment for Learning in Norway. *Assessment in Education: Principles, Policy & Practice, 22*(1), 44–60.

Hopfenbeck, T., & Stobart, G. (2015). Large-scale implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice, 22,* 1–2. https://doi.org/10.1080/0969594x.2014.1001566.

Hopfenbeck, T., Tolo, A, Florez, M. T., & El.Masri, Y. (2013). *Balancing trust and accountability? The assessment or learning programme in Norway*. A Governing Complex Education Systems Case Study. OECD working Papers No. 97. Paris: OECD Publishing.

Houghton, L., Ruutz, A., Green, W., & Hibbins, R. (2015). I just do not have time for new ideas: Resistance, resonance and micro-mobilisation in a teaching community of practice. *Higher Education Research & Development, 34*(3), 527–540.

Hubbard, L., Datnow, A., & Pruyn, L. (2014). Multiple initiatives, multiple challenges: The promise and pitfalls of implementing data. *Studies in Educational Evaluation, 42,* 54–62.

Hudson, C. (2011). Evaluation—The (not so) softly–softly approach to governance and its consequences for compulsory education in the Nordic contraies. *Education Inquiry, 2*(4), 671–687.

Hult, A., Lundström, U., & Edström, C. (2016). Balancing managerial and professional demands: School principals as evaluation brokers. *Education Inquiry, 7*(3), 29960.

Ingersoll, R. (1991). Loosely coupled organizations revisited. In *Paper presented at the annual meeting of the American Educational Research Association* (April 1991). https://files.eric.ed.gov/fulltext/ED351774.pdf.

Kellard, K., Costello, M., & Godfrey, D. (2008). *Evaluation of the developing thinking and assessment for learning development programme*. Birmingham: BMG Research.

Kirton, A., Hallam, S., Peffers, J., Robertson, P., & Stobart, G. (2007). Revolution, evolution or a Trojan horse? Piloting assessment for learning in some Scottish primary schools. *British Educational Research Journal, 33,* 605–627.

Kools, M., & Stoll L. (2016). *What makes a school a learning organisation*?. OECD Education Working Papers, No. 137. Paris: OECD Publishing. http://dx.doi.org/10.1787/5jlwm62b3bvh-en.

Lillejord, S., & Børte, K. (2019). Trapped between accountability and professional learning? School leaders and teacher evaluation. *Professional Development in Education*. https://doi.org/10.1080/19415257.2019.1585384.

Lillejord, S., Børte, K., Ruud, E., Hauge, T. E., Hopfenbeck, T. N., Tolo, A., et al. (2014). *Former for lærervurdering som kan ha positive innvirkning på skolens kvalitet. En systematisk kunnskapsoversikt*. Oslo: Kunnskapssenter for utdanning.

Lillejord, S., Elstad, E., & Kavli, H. (2018). Teacher evaluation as a wicked policy problem. *Assessment in Education: Principles, Policy & Practice*. Special issue.

Louis, K. S. (2007). Trust and improvement in schools. *Journal of Educational Change, 8*(1), 1–24.

Lovdata: Forskrift til Opplæringslova (the education act). https://lovdata.no/forskrift/2006-06-23-724. Retrieved 26 October 2018.

MacPhail, A., & Halbert, J. (2010). We had to do intelligent thinking during recent PE: Students' and teachers' experiences of assessment for learning in post-primary physical education. *Assessment in Education Principles Policy and Practice, 17*(1), 23–39.

Maguire, M., Braun, A., & Ball, S. (2018). Discomforts, opposition and resistance in schools: The perspectives of union representatives. *British Journal of Sociology of Education*, *39*(7), 1060–1073.

Mandinach, E. B., & Jimerson, J. B. (2016). Teachers learning how to use data: A synthesis of the issues and what is known. *Teaching and Teacher Education, 60,* 452–457.

Mayer, R. C., Davis, J.-H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy og Management Review, 20,* 709–734.

Mehta, J. (2013). From bureaucracy to profession: Remaking the educational sector for the twenty-first century. *Harvard Educational Review, 83*(3), 463–488.

Miliband, D. (2004). Personalised learning. Building a new relationship with schools. https://pdfs.semanticscholar.org/7df7/ceb2d78a797f3b9f2eeb09d36270083c22ac.pdf. Retrieved 24 May 2019.

Moos, L., Krejsler, J., & Kofod, K. K. (2008). Successful principals: Telling or selling? On the importance of context for school leadership. *International journal of Leadership in Education, 11*(4), 341–352.

Nguyen, T. D., & Hunter, S. (2018). Towards an understanding of dynamics among teachers, teacher leaders, and administrators in a teacher-led school reform. *Journal of Educational Change, 19*(4), 539–565.

OECD. (2017). How's life in Norway. https://www.oecd.org/statistics/Better-Life-Initiative-country-note-Norway.pdf Retrieved 30 September 2018.

Ofsted. (2008). *Assessment for learning: The impact on national strategy support*. Reference no: 070244.

O'Neill, O. (2002). *A question of trust: The BBC reith lectures 2002*. Cambridge: Cambridge University Press.

O'Neill, O. (2013a). Intelligent accountability in education. *Oxford Review of Education, 39,* 1. https://doi.org/10.1080/03054985.2013.764761.

O'Neill, O. (2013b). What we don't understand about trust. TED talk. https://www.ted.com/talks/onora_o_neill_what_we_don_t_understand_about_trust. Retrieved 20 June 2016.

O'Neill, O. (2014). Trust, trustworthiness, and accountability. In Morris, N. & Vines, D. (Eds.) *Capital failure: Rebuilding trust in financial services* (pp. 172–192). Oxford: Oxford University Press.

Ozga, J. (2009). Governing education through data in England: From regulation to self-evaluation. *Journal of Education Policy, 24*(2), 149–162.

Paulsen, J. M., & Høyer, H. C. (2016). External control and professional trust in Norwegian school governing: Synthesis from a Nordic research project. *Nordic Studies in Education, 36*(02), 86–102.

Philpott, C., & Oates, C. (2017). Professional learning communities as drivers of educational change: The case of learning rounds. *Journal of Educational Change, 18*(2), 209–234.

Prøitz, T. S., Mausethagen, S., & Skedsmo, G. (2017). Data use in education: Alluring attributes and productive processes. *Nordic Journal of Studies in Educational Policy, 3*(1), 1–5. https://doi.org/10.1080/20020317.2017.1328873.

Rubie-Davies, C. M., Peterson, E. R., Sibley, C. G., & Rosenthal, R. (2015). A teacher expectation intervention: Modelling the practices of high expectation teachers. *Contemporary Educational Psychology, 40,* 72–85.

Russell, M., Madaus, G., & Higgins, J. (2009). *The paradoxes of high stakes testing: How they affect students, their parents, teachers, principals, schools, and society*. Mumbai: IAP.

Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change, 11*(1), 45–61.

Scheerens, J. (2012). Summary and conclusion: Instructional leadership in schools as loosely coupled organizations. In Scheerens (Ed.) *School leadership effects revisited* (pp. 131–152). Dordrecht: Springer.

Schildkamp, K., Poortman, C., Luyten, H., & Ebbeler, J. (2017). Factors promoting and hindering data-based decision making in schools. *School Effectiveness and School Improvement, 28*(2), 242–258.

Schillemans, T., Van Twist, M., & Vanhommerig, I. (2013). Innovations in accountability: Learning through interactive, dynamic, and citizen-initiated forms of accountability. *Public Performance & Management Review, 36*(3), 407–435.

Silverman, D. (2006). *Interpreting qualitative data: Methods for analyzing talk, text and interaction.* Thousand Oaks: Sage.

Simons, P. R. J., & Ruijters, M. C. (2014). The real professional is a learning professional. In Billett, S., Harteis, C. & Gruber, H. (Eds.) *International handbook of research in professional and practice-based learning* (pp. 955–985). Dordrecht: Springer.

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

Tschannen-Moran, M., & Gareis, C. R. (2015). Faculty trust in the principal: An essential ingredient in high-performing schools. *Journal of Educational Administration, 53*(1), 66–92.

UNESCO. (2017). *Global education monitoring report. Accountability in education: Meeting our commitments.* Paris: UNESCO publishing.

Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice, 16,* 165–184.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *21*(1), 1–19.

Wilson, M. (2017). Making measurement important for education: The crucial role of classroom assessment. *Educational Measurement: Issues and Practice, 37*(1), 5–20.

Zhang, X. F., & Ng, H. M. (2011). A case study of teacher appraisal in Shanghai, China: In relation to teacher professional development. *Asia Pacific Education Review, 12*(4), 569–580.

Affiliations

Astrid Tolo[1] [iD] **· Sølvi Lillejord[2,3] · María Teresa Flórez Petour[4] · Therese N. Hopfenbeck[5]**

Sølvi Lillejord
solvi.lillejord@uib.no

María Teresa Flórez Petour
mtflorez@u.uchile.cl

Therese N. Hopfenbeck
therese.hopfenbeck@education.ox.ac.uk

[1]   Department of Education, University of Bergen, Bergen, Norway

[2]   University of Bergen, Bergen, Norway

[3]   Department of Education, University of Oxford, Oxford, UK

[4]   Pedagogical Studies Department, University of Chile, Santiago, Chile

[5]   Oxford University Centre for Educational Assessment, Department of Education, University of Oxford, Oxford, UK