

# School evaluation in Sweden in a local perspective: A synthesis

Anders Hanberger<sup>†\*</sup>, Sara Carlbaum\*, Agneta Hult\*\*,  
Lena Lindgren\*\*\* & Ulf Lundström\*

## Abstract

This article synthesises the role of evaluation at the municipal, school, classroom and parental levels of governance, and discusses the results of the articles appearing in this special issue. The discussion concerns the role of evaluation in school governance, the value of evaluation for local school development, the constitutive effects of evaluation, what explains the present results, how knowledge produced by evaluation can be used, and methodological issues. The results indicate that evaluation systems legitimise and support governance by objectives and results, parental school choice, and accountability for fairness and performance. Evaluation systems emphasise measurable aspects of curricula and foster a performance-oriented school culture. The most important evaluations for improving teaching and schools are teachers' own evaluations. The article suggests two explanations for the actual roles of evaluation in local school governance. First, both the governance structure and applied governance model delimit and partly shape the role of evaluation at local governance levels. Second, how local school actors use their discretion and interpret their role in the education system, including how they respond to accountability pressure, explains how their roles are realised and the fact that actors at the same level of governance can develop partly different roles.

*Keywords:* evaluation systems, local school governance, role of evaluation, consequences of evaluations

## Introduction

This special issue has explored how local school actors in Swedish compulsory education (for students aged 13–15 years) have responded to prevailing evaluation systems and the growing accountability pressure emerging from the recentralisation, marketisation and globalisation of education governance. Strengthened accountability is assumed to enhance education quality and promote school development (OECD 2011, 2015; SOU 2015:22), and a combination of control- and improvement-oriented evaluation systems has been institutionalised at various levels of the school system (Lindgren, Hanberger and Lundström 2016) to promote school development and enhance education quality. However, this development is contested by research claiming that the consequences of the growing accountability pressure are

---

<sup>†</sup>Correspondence to: Anders Hanberger, Department of Applied Educational Science, Umeå University, Sweden. Email: anders.hanberger@umu.se

\*Department of Applied Educational Science, Umeå University, Sweden.

\*\*Department of Education, Umeå University, Sweden.

\*\*\*School of Public Administration, University of Gothenburg, Gothenburg, Sweden.

©Authors. ISSN 2000-4508, pp. 349–371

problematic for school practice (Hoyle and Wallace 2009; Mausethagen 2013; Ravitch 2010). Empirical research into the role and consequences of evaluation systems at the local governance level, and into how local school actors respond to these systems, is very limited. Our research project<sup>1</sup> and this special issue aim to improve our understanding of the role and consequences of evaluation in local school governance. What local decision makers, school providers, principals and teachers consider relevant, useful and actionable knowledge (Stehr and Grundmann 2012) is crucial in understanding the role of evaluation in local school governance.

This article synthesises the articles presented in this special issue and advances discussion of their findings regarding the role and consequences of evaluation at the municipal, school, classroom and parent/citizen levels. Special attention is paid to the value and consequences of various evaluations for local school development.

The next section briefly describes the analytical framework, methods and material of this study. The governance structure in which evaluation is embedded in this context is then described. Our results concerning the roles and consequences of evaluation at four levels of school governance are then synthesised. An extended discussion treats the role of evaluation in school governance, its value for local school development, the constitutive effects of evaluation, what can explain the results, and how this knowledge can be used. The methodological challenges and limitations of the study are also considered. The article ends with conclusions about the role and consequences of evaluation in local school governance.

### ***Analytical framework and material***

The analytical framework that is briefly described here consists of key concepts, three models of decentralised governance and four types of evaluations. The framework is further developed in a separate article (Hanberger 2016).

### ***Key concepts***

*Evaluation* is used generically to refer to, for example, evaluation, inspection, quality assurance, ranking and both stand-alone evaluations and evaluation systems.

*Evaluation system* refers to “the procedural, institutional and policy arrangements shaping the evaluation function and its relationship to its internal and external environment” (Liverani and Lundgren 2007, 241). Evaluation system also refers to routines established for dealing with stand-alone evaluations and to a system producing streams of evaluation information.

*Governance* refers to the new, emerging institutions for collective action that have evolved from traditional forms of government, including negotiated interaction between a range of actors and institutions. Traditional forms of governing are included in the concept.

*Local school governance* refers to governance that occurs in a municipality and in a quasi-market where local school actors govern and influence schooling and education. It includes the efforts of actors and institutions to govern and influence matters such as school policy, education, school climate and school safety.

*Accountability* is both vertical (e.g. the state holding local governments and public agencies to account for national policies and programmes) and horizontal (e.g. parents holding principals and teachers to account for school and student performance), the focus being on accountability in the interests of fairness and performance (Behn 2001).

*School development* refers to changes in schools, teaching, and learning environments conceived as, and claimed to be, for the better. It does not refer to a predefined state or to specific characteristics or conditions in a school or education system.

*Local school development* refers to improvement in school culture, teaching, equity, and school/student performance.

Two concepts are used to identify the role of evaluation in local school governance. *Evaluation use* refers to actors' use of evaluation. *Functions of evaluation* can be indicated by actors' use but go beyond actual use to include e.g. how evaluations facilitate public accountability or legitimise the current governance model. The concept of *consequences* is used in a broad sense and includes the effects and influence of evaluation on governance, policy, institutions and actors.

*Constitutive effects* refers to tacit or indirect effects, for example, how evaluation (systems) can shape discourses, defining what is important in education and school systems.

### ***Brief description of three models of decentralised governance***

In the *state model*, the state governs and delegates responsibilities to lower governance levels. Policy is developed by the state and implemented from the top down. This model requires monitoring and evaluation that focus on achieving national objectives and compliance with the Education Act. Citizens exert democratic control when they vote in national elections, and are assumed to delegate power to the government between elections.

In the *local government model*, local governments govern education and schools. It is based on the discretion delegated to local governments and education committees by the state. Local governments implement national education policy, and develop and implement local school policy. To function, this governance model requires knowledge of the achievement of local objectives and of how schools perform. Democratic accountability is realised when citizens vote in local elections.

The *multi-actor model* reflects various forms of network governance in which public, private and civil actors govern education and schools through working groups, networks and partnerships. The shaping and implementation of education

policy are not conceived as separate stages of a policy process; rather, policy is conceived as being shaped while being implemented. Network governance requires evaluations that focus on key actors' knowledge needs and on the achievement of network objectives. Democratic control mainly manifests itself through network members' control of network governance and through national and local government control of networks (the models are further explained in Hanberger 2016 and Hanberger 2009).

### ***Four types of evaluation***

*Indicator-based monitoring and evaluation systems*, the first type, produce recurrent data on the inputs, processes and outputs/outcomes of policies and programmes and on school or student performance. *Stand-alone evaluation studies*, the second type, generate quantitative and/or qualitative knowledge and include in-depth evaluations, such as the National Board of Education's thematic evaluations and education committees' in-depth evaluations of schools. *Synthesis studies*, the third type, are evaluation studies that compile knowledge from various type-one and -two evaluations as well as other relevant studies; this type includes synthesis reports produced by, for example, the Swedish Schools Inspectorate (SSI). *Informal, occasional or everyday evaluations*, the fourth type, are conducted by local school actors to meet their personal knowledge needs and are not made official or published; examples of such evaluations are teachers' evaluations of their own or colleagues' teaching.

### ***Empirical material***

Four medium-sized municipalities with populations of 75,000–100,000 were selected strategically to reflect differing local conditions and contextual factors that may affect education and the role of evaluation in local school governance. The municipalities differ in political majority, school performance and share of independent schools (Hanberger *et al.* 2016), and eight schools were selected from them for in-depth study via interviews with personnel. The municipalities are anonymised, being called “North”, “West”, “East” and “South”.

The article is based on the analysis of documents, reports and studies treating global and national evaluation systems, national and municipal policy documents treating school governance and evaluation, minutes from municipal education committee meetings (2011–2013), municipal websites, and 76 interviews. Four politicians from majority parties and three from opposition parties, 10 administrators (i.e. Head of the Education Department, senior administrators, and evaluation experts), five politically elected local auditors, three representatives of independent schools, eight school principals, and 43 teachers were interviewed in person or, in a few cases, by phone.

The interviews lasted from 0.5 to 1.5 hours. In addition, an online questionnaire sent to teachers was used to complement their interviews so as to obtain an overview of teachers' experiences of evaluation in the studied municipalities.

### **The Swedish governance structure**

The discretion for local school governance and the role of evaluation in local school governance are not the same in all countries. Generally, local discretion is greater and the role of evaluation more prominent in decentralised education systems. In nation states, the government has the power to govern lower levels by means of various policy instruments, including evaluations. Local self-government is strongly protected in Sweden (Montin 2007), giving municipalities substantial discretion to adapt education and schooling to local conditions. National decision makers wield the power and mandate given in the country's political system and are confined by current statutes, which they can change if deemed necessary. Recent decades of school reforms in Sweden, characterised by decentralisation (e.g. more power to school principals), recentralisation (e.g. strengthened school inspection) and marketisation (e.g. allowing non-public school providers) were developed as solutions to perceived school problems and to improve education and schooling. It is recognised that these reforms have made the governing of schools complex and the division of responsibility unclear (Holmgren et al. 2012; Montin 2007; Skolverket 2011; SOU 2015:22).

An OECD report reviewing the Swedish education system, in view of the drastic decline in Sweden's PISA results over the last decade, criticises the lack of a clear division of roles and responsibilities between levels of Sweden's education system:

There is a lack of capacity and clarity in roles and responsibilities at various levels of the education administration, and local autonomy is not matched with adequate public accountability. These are key challenges for improving student performance (OECD 2015, 8).

Local decision makers' main roles in the Swedish education system are to implement national education policy, achieve national objectives and comply with the Education Act. The political system also gives discretion to local governments (via education committees) to govern education according to the principle of local self-government, in line with the local government model of governance. This discretion has diminished along with the recentralisation and privatisation reforms. Municipalities can use their discretion to govern schools more or less in line with national objectives and to develop and implement their own municipal school policies. Eagerness to develop and implement school evaluations in line with national objectives and to support parents' school choice can differ between right- and left-wing local governments.

Both levels of government can delegate responsibility to networks and partnerships. In addition, school actors in the public, private and civil spheres can create forums and develop their own school policies to influence national and municipal school policies and education, reflecting the multi-actor model of governance (Hanberger 2016).

Multilevel school governance rests on a division of power granting school actors and institutions different and overlapping responsibilities. It is recognised that actual governance may not correspond to the formal division of power and responsibilities.

## **Evaluations in local school governance**

### ***Municipal level***

Overall, decision makers have used school evaluations at the municipal level (Hanberger *et al.* 2016) as administrative tools to support local school governance in various ways. For those in power and the administration, the local evaluation system served mainly to sustain and legitimise the applied governance model, support the monitoring and governance of schools, and hold schools accountable for goal achievement and systematic quality work. Evaluations gave them an overview of how schools performed and helped them identify problems related to poor performance. Evaluations were used to foster school development, conceived as improved school and student performance, and greater equity, conceived as helping students finish grade nine with passing grades in all subjects.

Education committees used all four types of evaluations, but data from indicator-based monitoring and evaluation systems (the first type) appeared to be the most essential at this level. Stand-alone evaluations, synthesis studies and informal evaluations were also used as complements. Similarly, independent school providers also relied on recurrent performance data and set up their own evaluation systems by integrating data from national evaluations with their own evaluations. Large school companies developed company-wide evaluation systems, monitoring schools and holding them accountable for performance. According to our interviews, if school and student performance declined or if other problems arose independent school providers could take action more easily than could public schools.

Decision makers responded differently to national evaluations depending on how they perceived the validity and usability of these evaluations. The response to school inspections, for example, varied from unquestioned to questioned compliance. One strong school community questioned the validity of the school inspection, including how the local schooling system was portrayed, and considered the criticism from the SSI as shallow and unfair. A few national key performance measures were taken at face value and integrated into the four municipalities' evaluation systems. The SALSA<sup>2</sup> measure, adjusted for structural conditions increasing or decreasing

student performance, was often used when municipalities and schools benchmarked themselves. SALSA was conceived as a valid measure of school performance.

Politicians from the opposition parties in two municipalities claimed that the political majority used school evaluations for promotional purposes. One interviewee asked why the local new public management (NPM)-oriented governance model, which he thought was problematic, had not been evaluated, indicating that evaluation has been used to support single-loop learning (Argyris and Schön 1996) and to maintain this governance model.

National evaluations induced local decision makers to comply with national statutes and policy objectives, reinforcing their responsibility for student and school performance. Contextual factors such as ‘confidence in the governance model and evaluation system’, ‘school performance’, and ‘self-perception of the municipality as a school community’ affected how accountability functioned in the state model of governance (Hanberger et al. 2016). The consequences of this accountability pressure were dampened by developing local evaluation systems and using national evaluations selectively.

### ***School level***

Evaluations appeared to be an important but time-consuming task for school principals who conducted evaluation work with varying degrees of enthusiasm and effort (Hult, Lundström and Edström 2016). The most important benefits of evaluation from the principals’ perspective concerned developing teacher work in the classroom and gaining an overview of schools and of how the students and teachers were performing. Evaluations served other purposes as well, but were intertwined with, or overlapped, the general goal of improving education. Evaluations were used to gain an overview, expose problems and improve teaching and student achievement. Some principals strongly advocated the use of hands-on evaluations. The use of informal, occasional or everyday evaluations was therefore emphasised.

National statistics (indicator-based evaluations) were used for benchmarking to understand school performance and ranking, but these data were not used in everyday school activities. National test results were also used to compare schools’ results with national standards and to detect deficiencies in education. Schools’ results were used for comparisons in terms of both goal achievement and to improve competitiveness in the school market. Stand-alone evaluation studies and synthesis studies from the SSI or the National Agency for Education were also used to a minor extent by principals. However, the SALSA measure, compensating for socio-economic factors affecting student and school performance, was used by some principals. In particular, principals in low-performing schools used SALSA in communicating with teachers.

While the principals expressed varying opinions about the value and quality of school inspections, the SSI has apparently strengthened its position and principals must increasingly respond to its requirements. Principals are also subjected to governance and accountability requirements from local governments; this has made principals critical of municipal evaluation demands, which they largely perceived as burdens that increase the workload rather than being useful for school improvement.

The main function of evaluations evident at the school level was school development (as manifested in school and student performance), but evaluations also had various accountability functions at the school level, reflecting the principals' central and 'in-between' position in the chain of command. They are responsible for realising state and school provider policies and demands, and are held to account for how they succeed in the eyes of the state and school providers. At the same time, they are responsible for holding teachers and other school staff accountable for school results. The principals have been pressured by demands to improve performance and quality and, simultaneously, to meet internal requirements from staff and students – a situation that created a need for risk reduction and for evaluation data collection and reporting routines. One way to manage this accountability pressure was to ensure that all reports, documents and figures were in place, helping to create the image of an effective school. This strategy and symbolic function of evaluation was spurred by the SSI's demands.

### ***The classroom level***

The overall picture of the role of evaluation at the classroom level is that evaluations of everyday teaching and schoolwork were used to learn continuously about what works for different groups and individuals and for school development (Hult and Edström 2016). Teachers also had to participate in many external evaluations that increased their workload without offering any perceived value for them.

Evaluations that mattered and were used were those performed close to teachers' practice and intended to develop their teaching for the good of the students. These evaluations were performed daily by teachers themselves and in ongoing dialogue with their students and often with colleagues. To improve teaching, teachers need evaluations that inform them about, for example: what works; student results; students' views of lessons, interests and preferences; student well-being; and student awareness of results and goals. Together with colleagues, teachers reflected on common problems and issues related to their teaching, subjects and students. Mutual observations of lessons followed by feedback with colleagues in their school or municipality were other informal evaluations that teachers appreciated. They expressed a great need for continuous evaluations of their own teaching as a way of enacting their professionalism and responsibility for students and schools.



The interviewed teachers expressed a desire to continuously improve education and increase learning opportunities for the students.

Although critical of most external evaluations, teachers generally complied with and participated in them. Consequences of this were less time available for creativity and lesson planning as well as a feeling of distrust. Most external evaluations were perceived as indications that teachers were not trusted by parents, principals, school providers or the state. Some teachers refused to complete some external evaluations as a way to manage the accountability pressure and defend their professionalism and responsibility for students.

In contrast to this negative picture of external evaluations, there are a few examples of external evaluations perceived as valuable by teachers. Some principals have developed and used evaluation systems close to teachers' practice that teachers considered relevant and useful for improving their teaching and daily work. Although most teachers were very critical and without positive experiences of external evaluations, some still hoped that a good evaluation model could be developed to support the ongoing development of teaching and schools.

The main functions of classroom-level evaluations, besides school development, are vertical and horizontal accountability. Teachers were held accountable for student performance by national and local governments, school providers and parents. Many teachers reported increased use of tests and assignments in response to horizontal accountability. They felt a need to establish legitimate measures of student accomplishments when questioned by parents about how they graded their students, for example.

### ***Parental level***

While we had insufficient empirical material to explore whether and, if so, how parents and citizens actually used evaluations, our analysis of their role in local school governance was confined to how evaluations were presented on websites and intended to be used by citizens, primarily parents.

The overall picture of the role of evaluation at the parental level of governance, explored in "Customers, partners and rights-holders: School evaluations of websites" (Carlbaum 2016), is that evaluations addressing parents are intended to facilitate informed school choice and support public (or customer) accountability. The schools' and education committees' websites provide indicators of student and school performance and quality such as grades, test results and SSI inspection reports. This information can be used by parents to hold sub-standard schools responsible for poor performance and can encourage parents whose children are students in low-scoring schools to move their children to other schools. Taken together, these evaluations are intended to promote customer-driven school development – that is,

schools must improve and adapt to customer wants and needs in order to survive in the school market.

East municipality used customer satisfaction surveys that could help parents and students to choose schools that reported having satisfied parents and students and to avoid or leave schools that did not. As demonstrated in “Navigating the evaluation web” (Hanberger *et al.* 2016), decision makers in East considered it intrinsically good to provide parents with evaluation information, particularly to facilitate school choice, even though it is not widely used. Although this evaluation information is accessible on websites, it was not known whether all parents knew of its existence.

Analysis of evaluations provided on school providers’ and schools’ websites indicates that parents have access to varying amounts of information that could be used to support informed school choice and monitor how their children’s and other schools are developing and performing. However, whether parents use the existing evaluations to choose, change or influence schools is unknown.

Evaluations can also have legitimating or symbolic functions. Publishing evaluations such as SSI reports, Open Comparisons and systematic quality reports could be a way for schools to build legitimacy. Parents will feel safer enrolling their children in schools that follow rules and regulations, work extensively on school development, and realise improvements through applying monitoring and evaluation results, for example. The use of customer satisfaction surveys can also function to legitimise current school governance, with parents’ opinions being collected to amass evidence of well-functioning schools or, symbolically, to demonstrate that evaluations are being conducted, though not being used in any specific way.

## **Discussion**

The results of the study partly confirm the OECD’s assessment of unclear responsibilities and accountability problems in the Swedish education system (OECD 2015), but they also indicate that local school actors are trying to overcome the problems engendered by the multiple overlapping evaluation systems. More importantly, though, the consequences of national evaluations are extensive and in some sense problematic for local actors. Evaluations intended to support national governance, accountability and school development appear mainly as a burden at local levels. Local school actors spend considerable time responding to and providing information for national evaluations, and developing response strategies to manage the accountability pressure. A few school providers and principals, however, have succeeded in developing evaluations that support teachers’ ongoing work and help them improve their teaching. This illustrates how the focus and design of evaluations and pedagogical leadership are crucial if evaluations are to support collegial learning and improve teaching and schools.

Given the problems of unclear responsibilities and accountability in the Swedish education system, some national policy makers are strong proponents of the need to strengthen the 'chain of command' and use evaluation to improve compliance with national objectives (OECD 2015; SOU 2015:22, 17), reflecting a top-down perspective on the education system and policy implementation. However, policy analysts and bottom-up researchers recognise that the implementation of national policies and evaluations is not a linear, instrumental process (Apple 2014; Colebatch 2009; Fazekas and Burns 2012; Hanberger 2011), as the present study also demonstrates. Instead, policy implementation is a complex process of mediation, interpretation and translation at various levels (Ball, Maguire and Braun 2012). Similarly, before evaluation conclusions and recommendations are endorsed and implemented by an organisation, it will first respond to and, possibly, challenge them (Hanberger 2011). This study demonstrates that if school actors have confidence and believe in what they are doing, they can respond analytically and critically to external evaluations. In light of our results, simply increasing the number of national evaluations cannot be expected to improve school development in terms of improving teaching and school practice.

### ***The role of evaluation in school governance***

The analytical framework supports two interrelated explanations of the actual roles of evaluation in local school governance. First, the governance structure and applied governance model delimit and partly shape the role of evaluation at local governance levels. Second, how local school actors use their discretion and interpret their role in the education system, including how they respond to accountability pressure, explains how their role is realised and how actors at the same level of governance can develop partly different roles.

Similarly, the consequences of evaluation for school practice can be interpreted as effects of the applied governance model, of how the state, local governments and school providers govern principals and teachers, and of whether and how they hold them accountable for school and student performance. Consequences also emerge as a result of how school administrations present and give parents access to evaluations.

The use of evaluation differed somewhat between politicians in power and opposition politicians (Hanberger et al. 2016), between school principals who act as managers and those who act as pedagogical leaders (Hult, Lundström and Edström 2016), and between teachers who use their own and school providers' and principals' evaluations to develop their teaching and those who rely on their own informal evaluations (Hult and Edström 2016). Therefore, how these actors interpret their discretion, empower themselves and perform their roles affect the consequences of evaluation for school practice. Some school providers and principals tried to manage the negative effects of evaluation and the growing accountability pressure

by adapting evaluations to the needs of teachers' daily practice and by protecting teachers from too many external evaluations, which reduced the negative consequences of accountability.

How can the three models of governance, explored in a separate article (Hanberger 2016), help interpret and explain these results? The *state model* is in operation when national evaluations are used to increase compliance with national regulations and achievement of objectives.<sup>3</sup> In the ideal model of national school governance, local governments and school providers implement national education policy, but policy enactment is not, as demonstrated, simply a linear implementation process from the centre to the classroom level. Local decision makers and principals spend considerable time responding to national evaluations and managing the growing accountability pressure, strongly influenced by school inspections and a few key performance measures. These performance measures or numbers are used by decision makers, school providers and some principals to increase performance demands on teachers, illustrating how the state implements governance using a few key numbers (Lingard 2011).

Yet not all national evaluations are implemented fully or implemented with resistance, as local actors do not always conceive the state's portrayal and assessment (e.g. via school inspection) of a municipality's education system and schooling as valid. Local decision-maker responses to national evaluations were related to whether they perceived the state's assessment of the local school system as valid and whether they were confident as a school community (Hanberger *et al.* 2016). In a decentralised education system, local actors are not only implementers of national policy but have further responsibilities and are accountable within a broad accountability environment (Behn 2001; Hanberger 2009). This makes them more inclined to pursue local objectives, addressing national objectives only if they are aligned with local endeavours or when forced to. Taken together, this partly explains why local governments, school providers and schools are not easily governed by national evaluations, and why the present articles report variations in compliance with national objectives.

The *local government model* of governance is in operation when local decision makers (i.e. Education Committees) interpret their discretion and selectively apply national evaluations in the local evaluation system. The main function of local evaluation systems was to support and legitimise local governance by objectives and results, through monitoring and evaluating schools and student performance and using performance results to take action to improve student performance and equity (e.g. to increase the number of students with passing grades in all subjects). Interviewed politicians from the majority and opposition parties and minutes from education committees indicated that local evaluation systems have been set up and used mainly as administrative tools (cf. Dahler-Larsen 2000), for single-loop learning (Argyris and Schön 1996; Nilsson 2005) and to support the current governance model.

The local governance model had not been evaluated in any of the studied municipalities, indicating that evaluations serve to maintain the current governance model. Local auditors in Sweden are commissioned to evaluate whether local governments control the local chain of governance and whether their school governance is effective according to municipal objectives. However, this task was rarely performed in the studied municipalities.

The *multi-actor model*, characterised by the delegation of responsibilities by the state or local government, or by actors empowering themselves in emerging institutions, was not found to prevail extensively. School administrators and principals used yearly quality reports (summarising systematic quality work by schools and municipalities) to deliberate on school status and on how to resolve problems and improve performance. However, this is a fairly closed form of network governance. Teachers and principals also used their own, colleagues' and students' evaluations to govern and improve schools in networks and in intra-school working groups. Parents' boards and other forums for network governance exist, but were not explored here.

The three models of governance exist in various forms and combinations in the Swedish school system, and in other education systems embedded in multilevel governance. Overlapping governance models and responsibilities as well as actors' various objectives, endeavours and knowledge needs can explain how evaluation is used in different ways to support governance, accountability and school development, and how evaluations create multiple governance and accountability problems.

Our study of the interplay between evaluation and governance illustrates how multilevel governance and governance by objectives and results shape the roles of evaluation. We did not identify the open and critical role of evaluation in democratic governance (Hanberger 2013), for example, to provide knowledge of the consequences of NPM-oriented governance and evaluations that can be used to change local school governance and school policy.

### ***Implications of evaluation for school development***

Evaluations and evaluation systems have contributed to various aspects of school development, serving mainly to maintain a performance-oriented *school culture* reflecting an ideal of education in the interest of economic growth, promoting performativity, standardisation and decontextualisation (cf. Carvalho 2012; Keddie 2013; Lawn 2011; Mangez and Hilgers 2012). Although local school cultures differ in many respects, all have become more performance oriented in recent years. Recurrent evaluation has become the norm and school and student performance the most valid criteria defining and measuring school development. In contrast, evaluations promoting personal development and the creation of reflective and culturally aware citizens, reflecting the enlightenment ideal, are not found.

Reflecting developments in school effectiveness research, recurrent evaluations are used to take action to enhance *student and school performance*, which also fosters a performance-oriented school culture. Although our study indicates some resistance (among opposition politicians, some principals and some teachers), all school actors must consider and respond to the SSI's inspections and be responsible for key performance data regarding student and school performance.

A main result of our study is that school development conceived as developing *teaching* is based mainly on teachers' own informal evaluations of their own or colleagues' teaching regarding what works for various groups and students. Except for a few school providers' and principals' evaluations, external evaluations were not used by teachers in developing teaching because their results cannot help teachers improve their school practice.

Evaluations are also used for school development conceived as increasing *equity*, as is understood in school effectiveness research (Reynolds *et al.* 2014). Education committees, school providers, principals and teachers highly prioritise supporting students who do not have sufficiently high grades to pass grade nine. In addition, a few school providers have evaluated the causes of student failure and studied effective measures to reduce failure, finding that key performance measures alone cannot be used to promote school development conceived in this way. The main equity aspect monitored by the evaluation systems is student performance, which decision makers use to take action to close the gap between high- and low-performing students. Apparently, evaluation systems do not measure other aspects of equity such as gender, economic, social and cultural diversity, and fairness (Englund *et al.* 2008).

Our definition of local school development is synthesised from school effectiveness and school development research, but this does not mean that we give equal weight to the different aspects of the concept. School development conceived as improved teaching and school culture appears to us to be a more valid representation of school development than student and school performance captured by a few key performance measures.

### ***Constitutive effects of evaluation***

The articles have identified several constitutive effects. Evaluation systems measuring grades and results in national tests have helped shape the notion of what is worth knowing about school performance and thus what constitutes quality in education. This has affected all school actors' conceptions of education quality. The average final grade in compulsory school "is the most important output to make visible for the school provider, the environment and ourselves", according to one interviewed decision maker (Hanberger *et al.* 2016). This key performance measure

also contributes to reductionism (Dahler-Larsen 2012a, b; 2013), that is, evaluation of the curriculum is narrowed to measurable goals, while broad qualitative goals disappear. Some principals warned that indicator-based evaluation systems could be reductive, promoting strategic behaviour such as teaching to the test. This is noted by researchers such as Lingard (2011, 357), who says that “policy as numbers has become the reductive norm for contemporary education at all levels”, and by Lundahl (2009, 208), who claims that the language with highest legitimacy in the public debate wins, that is, “numbers have a tendency to win versus words”, which he says constitutes an obstacle to pedagogical development. The monitoring of key performance measures and using them for benchmarking constitute the norm that benchmarking is something one ought to do to promote education quality.

Local evaluation systems have reinforced the conception of parents and students as customers and of schooling as constituting a market, creating the expectation that customers should make repeated school choices to ensure the best possible education. How evaluations are presented to parents reinforces and legitimises marketisation in education, in which schools compete for students and parents. This supports a notion of education as a private rather than a public good in which the parental right to educational authority remains unquestioned. Further, this notion risks reducing parental interest in collective action and co-management in favour of individual action and problem-solving strategies, such as school choice, school exit and complaints. This notion also risks changing the professional role of teachers in that they need to adapt to the attitude that ‘the customer is always right’, putting pressure on teachers to change grades and assessments, questioning their authority and resulting in an overload of individual demands.

The explored evaluation systems mainly support roles of evaluation and key actors in school governance promoted by NPM. For example, principals’ use of evaluations to hold teachers to account for student performance promotes the managerial role of principals, which can be conceived as a constitutive effect. In contrast, principals’ use of evaluations to support teachers, promote collegial learning and protect core pedagogical values fosters a pedagogical leadership role. When the SSI inspects the performance of pedagogical leadership, the principals’ roles as managers and as pedagogical leaders are not easily separated, with the possible constitutive effect that the two roles may be conceived as integrated in the profession.

Another indirect or constitutive effect of monitoring performance and evaluating school failures is that evaluation systems used for performance accountability purposes yield negativism (Schillemans and Bovens 2011). This fixation on poor results and problems in the school environment has created frustration among teachers in problem schools, for example. The strong audit culture cultivated by the applied governance model and evaluation systems has also fostered a culture of silence, particularly in independent schools.

Teachers' participation in the great number of compulsory evaluations has led to fatigue and stress, reduced the time available for reflection, and eroded creativity and work satisfaction. In addition, feelings of mistrust are an indirect effect of all evaluations. Teachers have experienced external evaluations and required documentation as impoverishing the teaching profession. A constitutive effect of evaluation, however, is the reprofessionalisation of teachers as accountability holders producing quality in education and improved grades.

### ***Methodological challenges and limitations of this study***

The advantages and limitations of using an open approach to capture the role of evaluation systems in local school governance and the consequences of evaluation for school practice are discussed in a separate article (Hanberger 2016). A comparative case study of four municipalities in one country does not allow the drawing of far-reaching conclusions regarding the effects and consequences of evaluations in local school governance, and that was not the purpose of this project. Still, the findings and patterns described in this issue are likely to be recognised in other local contexts and cases (Larsson 2009) as the evaluation trend is widespread in decentralised school systems.

In this issue, 'constitutive effects' refers to how evaluation systems constitute the conditions and modes of school governance and actors' role in governance, and to other tacit effects of evaluation in the education context. One can interpret constitutive effects as resulting from previous political decisions (e.g. school reforms) and from routinisations that continuously change the conditions for school governance. Some of these effects reflect political endeavours. Advocates of NPM, for example, promote the role of parents as customers and argue that principals should act as managers. Different constitutive effects emerge from how actors conceive and communicate their role in governance and from how they act in their daily practice, and these effects can go beyond these actors' formal mandates. Some constitutive effects are not intended or foreseen (e.g. negative stress and loss of creativity). Our approach is eclectic and borrows elements from different research traditions. It is based on a constructivist epistemology in which quantitative and qualitative data are interpreted in complementary ways. Whether the identified effects we refer to as constitutive are best represented by this concept is open to debate. Peter Dahler-Larsen (2012b; 2013), who introduced the concept of constitutive effects, suggests that it can be used as an alternative to the commonly used notion of the intended and unintended consequences of indicators and evaluation systems, which he conceives as problematic. Indicators help constitute our notion of a reality and according to Dahler-Larsen (2013, 15) they also



define a strategic landscape in which practitioners must navigate. This landscape is political, as it consists of policy-related categories by means of which society seeks to manage itself and thereby represents itself and its values.

As demonstrated in the articles, the state and local governments have institutionalised evaluation systems, but none of them steps back after creating the evaluation arena or the 'strategic landscape'. They continue to govern society and local actors by means of evaluation in various ways. In this issue, we have used the concept of constitutive effects, but not as an alternative to intended and unintended consequences because our research concerns the obvious and tacit effects and consequences of evaluations. As Dahler-Larsen (2013) emphasises, what is intended and unintended does not remain stable and intentions and evaluation systems change, yet does this mean that constitutive effects cannot be intended, that active governance disappears once the rules of the game (i.e. the strategic landscape) are created and that society governs itself? Some indicator-based evaluation systems seem to have that intended function, namely, to provide benchmarking data so that local school actors can govern themselves (e.g. Open Comparison). However, some evaluation systems are used to actively govern principals and teachers not only in terms of steering them to govern themselves by numbers, and steering how they should conceive of reality, education quality, accountability and good governance, but also to reinforce their role as accountability holders responsible for taking improvement action before and after inspection in line with national standards. 'Hard' governance through strengthened accountability is manifested when the SSI inspects and holds school providers accountable for meeting national standards and achieving national education objectives, threatening or imposing sanctions to ensure compliance. The latter has a constitutive effect in terms of how school providers, principals and teachers conceive themselves as accountability holders, that is, accountable to the state for school and student performance.

When democratic control and what politicians and parties can be held to account for are discussed, the concept of constitutive effects without reference to intention is of little use. It is an ambiguous concept meaning different things to different researchers, and is also difficult to use in communication with practitioners. Yet the concept's ambiguity is also a strength in that it opens one's eyes to all kinds of important but elusive and indirect effects of evaluations.

The three governance models provide three maps for analysing the role of evaluation in multilevel governance, but there are limitations to their empirical application. The data collection and articles were not structured according to these maps; they were mainly used during interpretation to enhance our understanding of why evaluations are developed and used in various ways in local school governance. Here, the local government model is applied to the Swedish education system,

although it is also applicable to systems where regional governments have a role in school governance (Hanberger 2009).

### ***What is the value of this knowledge?***

This knowledge can be used in various ways. It can help local school actors better understand the governance conditions in which they operate, and encourage them to use their freedom of choice when demanding and using evaluations to promote school development. The results regarding the consequences of evaluation can be used when deliberating about clarifying and improving the role of evaluation in national and local school governance, to recognise the importance of teachers' daily, occasional and informal evaluations for improving teaching (and in supporting and providing more time for these evaluations), and in managing the negative consequences of evaluations. Decision makers and evaluation system architects can use these results when deliberating about which evaluation systems are needed and which systems can be terminated or merged. Knowledge of how evaluations work in practice can also be used in more philosophical and political discussions of the role of evaluation in democratic school governance. The three governance models, reflecting different democratic notions, can inform this discussion. Which evaluations are needed to support different notions of democratic school governance? What should the role of evaluation be in democratic school governance, and how can existing evaluation systems and evaluations be managed and changed by democratic institutions?

### **Conclusions**

This study has demonstrated that multiple accountability problems emerge as a result of overlapping evaluation systems in a crowded policy space (Lindgren *et al.* 2016) and that local decision makers set up their own evaluation systems to adapt the evaluation web to meet the needs of municipal school governance (Hanberger *et al.* 2016). Principals and teachers are subjected to many external evaluations that usually have little value for them; instead, they mainly use teachers' own informal evaluations to improve teaching and school practice (Hult, Lundström and Edström 2016; Hult and Edström 2016). School providers and schools present their evaluations to parents on websites, shaping the parents' role as customers; this evaluation information helps parents choose schools and hold schools and teachers accountable for student performance (Carlbaum 2016).

Most of the more than 30 evaluation systems identified in Swedish compulsory education (for students aged 13–15 years) produce quantitative data capturing measurable aspects of education, whereas data capturing other parts of the curriculum, more difficult or impossible to measure (e.g. how schools have succeeded in achieving

democracy, sustainability and solidarity objectives), are lacking. A few key performance measures are used in several systems (Lindgren et al. 2016).

The ultimate intended function of evaluation is to support school development, yet in practice the identified evaluation systems have many and partly conflicting functions at the municipal, school, classroom and parent/citizen levels, creating multiple governance and accountability problems.

National and local evaluation systems legitimise and support governance by objectives and results, parental school choice, and accountability in the interests of fairness and performance. They also support school development conceived as alleviating problems affecting student performance and fostering a performance-oriented school culture. At the municipal and school levels, evaluations promote single-loop learning and confirm rather than question current policies and routines.

The identified evaluation systems induce local school actors and institutions to think and act according to the principles of NPM; these are aligned with most decision makers' and managerial-oriented principals' endeavours but not with those of all local school actors. This indicates that evaluations in local school governance are politically loaded and serve to support and legitimise the applied governance model and current education policy. Stakeholder evaluations that can provide a more multifaceted understanding, including critical accounts that school actors can use for informed deliberation about the status of schools, consequences of current school policy, and where to go in the future, are not found in our case communities.

The workload and accountability pressure have increased for both principals and teachers. The effects have been the most negative for teachers, however, as external evaluations have questioned their professional competence and authority, unintentionally damaging teacher motivation. The external evaluation systems had little or no value in terms of helping teachers improve their teaching practice. Instead, teachers used their own evaluations regarding what works for various groups and students to continuously improve teaching and schools. A few school providers and principals succeeded in developing evaluations addressing the needs of teachers that were used in developing teaching and daily practices.

There is a need to commission the evaluation of and further research into the roles and consequences of evaluation at different levels of school governance, particularly at local levels. How do national and local decision makers govern and manage the existing evaluation systems? Could the negative consequences of most external evaluations for school practice encourage decision makers or other actors (e.g. global organisations and teachers' unions) to initiate open, reflective, probing and sceptical evaluations of today's evaluations and evaluation systems based on assumptions and analytical frameworks other than those of the OECD's assessments? Or will this continue to be a task for researchers alone?

It is hoped that this issue can encourage further research (e.g. case, survey and comparative studies) into the shaping and implementation of local school evaluation policies at different local levels, and into the interplay between evaluation/evaluation systems and governance. Research into parents' and citizens' roles in school governance is also needed, for example, exploring how parents use existing evaluation systems and what their own informal evaluations entail, how parents use evaluations for horizontal accountability, and how individual and collective action affects schools. A closer look at open and probing evaluations that emerge outside evaluation systems is warranted, as is examination of how evaluation systems can support teachers and collegial learning work. Although such research is currently rare, it can provide essential knowledge and inspire local school actors to empower themselves, helping them find ways to manage external evaluation systems.

---

*Anders Hanberger* is Professor in Evaluation and director of research at the Umeå Centre for Evaluation Research, Umeå University. His research areas cover policy analysis and evaluation research and focus on public policy, governance, evaluation, evaluation systems, and methodology. Special interests include the role of evaluation in democratic governance. Email: anders.hanberger@umu.se

*Sara Carlbaum* is a Doctor in Political Science and a researcher at Umeå University, Sweden. Her research centres on education policy and governance, with a focus on evaluation, school inspection, and marketisation. Particular interests are citizenship constructions, social justice, gender, rights and equity. Email: sara.carlbaum@umu.se

*Agneta Hult* is Associate Professor at the Department of Education, Umeå University. Her research interests mainly focus on the implications and consequences of evaluations and assessment in the field of education. Email: agneta.hult@umu.se

*Lena Lindgren* is Associate Professor at the School of Public Administration, Gothenburg University. She works with teaching and research concentrating on evaluation and policy analysis, and has also conducted several evaluations on commission for government agencies, particularly in the education field. Email: lena.lindgren@spa.gu.se

*Ulf Lundström* is Associate Professor at the Department of Applied Educational Science, Umeå University. His main research interests lie within the areas of education policy (e.g. focusing on marketisation and inclusion), evaluation and the teaching profession. Email: ulf.p.lundstrom@umu.se

## **Notes**

- 1 This special issue emerged from the research project “Consequences of evaluation for school practice: governance, accountability, and school development” (2012–2015), financed by the Swedish Research Council.
- 2 SALSAs stands for Skolverkets Arbetsverktyg för Lokala Sambands Analyser (NAE’s tool for local correlation analysis).
- 3 Results of national evaluations and international assessments of the Swedish education system (PISA in particular) are also used by political parties when developing education policy (interviews with education spokespeople from five political parties).

## References

- Argyris, C. and Schön, D. (1996). *Organizational learning II: theory, method and practice*. Reading, MA: Addison-Wesley.
- Ball, S.P., Maguire, M. and Braun, A. (2012). *How schools do policy. Policy enactments in secondary schools*. London: Routledge.
- Behn, R. (2001). *Rethinking democratic accountability*. Washington, DC: Brookings Institution Press.
- Carlbaum, S. (2016). Customers, partners and rights-holders: school evaluations of websites. *Education Inquiry*, 7(3), 29971, doi: <http://dx.doi.org/10.3402/edui.v7.29971>
- Carvalho, L. (2012). The fabrications and travels of a knowledge-policy instrument. *European Educational Research Journal*, 11(2), 172–188.
- Colebatch, H.K. (2009). Governance as a conceptual development in the analysis of policy. *Critical Policy Studies*, 3(1), 58–67.
- Dahler-Larsen, P. (2000). Surviving the routinization of evaluation: the administrative use of evaluations in Danish municipalities. *Administration & Society*, 32(1), 70–92.
- Dahler-Larsen, P. (2012a). *The evaluation society*. Stanford, CA: Stanford Business Books.
- Dahler-Larsen, P. (2012b). Evaluation as a situational or universal good? Why evaluability assessment for evaluation systems is a good idea, what it might look in practice, and why it is fashionable. *Scandinavian Journal of Public Administration*, 16(3), 29–46.
- Dahler-Larsen, P. (2013). Constitutive effects of performance indicators: getting beyond unintended consequences. *Public Management Review*, 16(7), 969–986.
- Englund, T. et al. (2008). *Vadå likvärdighet?: studier i utbildningspolitisk språkanvändning*. Göteborg: Daidalos.
- Fazekas, M. and Burns, T. (2012). Exploring the complex interaction between governance and knowledge in education. *OECD Education Working Papers*, No. 67, OECD Publishing.
- Holmgren, M., Johansson, O., Nihlfors, E. and Skott, P. (2012). Local school governance in Sweden: boards, parents and democracy. *Journal of School Public Relations*, 33(1), 8–28.
- Hoyle, E. and Wallace, M. (2009). Leadership for professional practice. In *Changing Teacher Professionalism. International Trends, Challenges and Ways Forward*. S. Gewirtz, P. Mahony and A. Cribb (eds.), pp. 67–80. London; New York: Routledge.
- Hanberger, A. (2009). Democratic accountability in decentralized governance. *Scandinavian Political Studies*, 32(1), 1–22.
- Hanberger, A. (2011). The real functions of evaluation and response systems. *Evaluation*, 17(4), 327–349.
- Hanberger, A. (2013). Framework for exploring the interplay of governance and evaluation. *Scandinavian Journal of Public Administration*, 16(3), 9–28.
- Hanberger, A. (2016). Evaluation in local school governance: a framework of analysis. *Education Inquiry*, 7(3), 29914, doi: <http://dx.doi.org/10.3402/edui.v7.29914>
- Hanberger, A., Lindgren, L. and Lundström, U. (2016). Navigating the evaluation web: evaluation in Swedish local school governance. *Education Inquiry*, 7(3), 29913, doi: <http://dx.doi.org/10.3402/edui.v7.29913>
- Hult, A., Lundström, U. and Edström, C. (2016). Balancing managerial and professional demands: school principals as evaluation brokers. *Education Inquiry*, 7(3), 29960, doi: <http://dx.doi.org/10.3402/edui.v7.29960>
- Hult, A. and Edström, C. (2016). Teacher ambivalence towards school evaluation: promoting and ruining teacher professionalism. *Education Inquiry*, 7(3), 30200, doi: <http://dx.doi.org/10.3402/edui.v7.30200>
- Keddie, A. (2013). Thriving amid the performative demands of the contemporary audit culture: a matter of school context. *Journal of Education Policy*, 28(6), 750–766.

- Larsson, S. (2009). A pluralist view of generalization in qualitative research. *International Journal of Research & Method in Education*, 32(1), 25–38.
- Lingard, B. (2011). Policy as numbers: accounting for educational research. *The Australian Educational Researcher*, 38(4), 355–382.
- Lawn, M. (2011). Standardizing the European education policy space. *European Educational Research Journal*, 10(2), 259–272.
- Lindgren, L., Hanberger, A. and Lundström, U. (2016). Evaluation systems in a crowded policy space: implications for local school governance. *Education Inquiry*, 7(3), 30202, doi: <http://dx.doi.org/10.3402/edui.v7.30202>
- Liverani, A. and Lundgren, H. (2007). Evaluation systems in development aid agencies: an analysis of DAC peer reviews, 1996–2004. *Evaluation*, 13(2), 241–256.
- Lundahl, Christian (2009). *Varför nationella prov? Framväxt, dilemman, möjligheter*. Lund: Studentlitteratur.
- Mangez, E. and Hilgers, M. (2012). The field of knowledge and the policy field in education: PISA and the production of knowledge for policy. *European Educational Research Journal*, 11(2), 189–205.
- Mausethagen, S. (2013). Accountable for what and to whom? Changing representations and new legitimation discourse among teachers under increased external control. *Journal of Educational Change*, 14(4), 423–444.
- Montin, S. (2007). *Moderna kommuner* [Modern municipalities] (3rd ed.). Malmö, Sweden: Liber.
- Nilsson, M. (2005). The role of assessments and institutions for policy learning: a study on Swedish climate and nuclear policy formation. *Policy Sciences*, 38(4), 225–249.
- OECD (2011). *OECD reviews of evaluation and assessment in education, SWEDEN*. Paris: OECD Publishing.
- OECD (2015). *Improving schools in Sweden: an OECD perspective*. Paris: OECD Publishing.
- Ravitch, D. (2010). *The death and life of the great American school system: how testing and choice are undermining education*. New York: Basic Books.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C. & Stringfield, S. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement*, 25(2), 197–230.
- Schillemans, T. and Bovens, M. (2011). The challenge of multiple accountability: does redundancy lead to overload? In *Accountable Governance. Problems and Promises*, M. J. Dubnick, and H. G. Frederickson (eds.), pp. 3–21. Armonk: ME Sharpe.
- Skolverket. (2011). *Kommunalt huvudmannaskap i praktiken: en kvalitativ studie* [Municipal responsibility in practice: a qualitative study]. Stockholm: Skolverket.
- SOU 2015:22. *Rektorn och styrkedjan. Betänkande av Utredningen om rektorernas arbetssituation inom skolväsendet* [The principal and the chain of command. Investigation of principals' work situation in compulsory education]. Stockholm: Fritzes.
- Stehr, N. and Grundmann, R. (2012). How does knowledge relate to political action? *Innovation: The European Journal of Social Science Research*, 25(1), 29–44.