

Quality in Finnish Comprehensive Schools

Entrenching Enhancement-Led Evaluation

Jaakko Kauko, Hannele Pitkänen and Janne Varjo

Going against the Grain

The evaluative turn in the public sector that gained momentum from the 1990s has increased the density and scope of the examination of student performance in education policy (Kauko and Varjo 2008; Vedung 1997). The globally common solution has become to devise a national large-scale assessment for measuring student achievement. These assessments are not only for understanding learning performance, but are increasingly used for monitoring and steering purposes. Verger et al. (2018) note that the results of the tests often either directly influence resource allocation, or their open test data are used for rankings which facilitate school competition and markets. Given its popularity, it is perhaps surprising that test-based accountability does not actually result in better schooling. Verger and Parcerisa's (2017: 245) meta review notes that the positive results of test-based accountability in academic performance are usually 'rather weak or mixed'. They also find that the lowest-performing schools are usually unlikely to benefit from test-based accountability (Verger and Parcerisa 2017). Many studies from various countries have concluded that strong testing regimes do not benefit educational aims (e.g. Gurova et al. 2018; Hardy 2021; Koyama 2011; see also Looney 2011: 5, 10–11, 18 and Ravitch 2010). In this regard, the evaluative turn in education cannot be fully justified by the improving academic performance of students. The reasons for having such evaluations must be sought outside education.

Indeed, arguments that back the strong global push for more evaluation do not see evaluation as a means to achieve educational success but rather to support other policy sectors. Persuasive and widespread (Kauko et al. 2018a) arguments suggest that school performance is directly linked with economic prosperity (Hanushek and Woessman 2007). These are supported by the World Bank (Takala et al. 2018) and the OECD (Ydesen 2020), for instance, but they have been repeatedly debunked (Komatsu and Rappleye 2017). Many countries use results from education evaluation to achieve aims in other policy sectors, whether it be a stronger geopolitical position or dealing with governance decentralization (Kauko et al. 2018b), or furthering 'education export' (Schatz 2016) for economic gains.

In other words, the global mainstream in education evaluation is not aimed at developing learning but is a proxy for other aims. Against this tendency to load education policy with—for lack of a better word—‘external’ aims, one could argue that education policy should further learning and do this in its sense of *Bildung*, the wider meaning of education cultivation and culture (see Saari et al. 2017). Taking that latter perspective, we will demonstrate that Finland has been able both to develop and sustain an exceptional system of evaluation in education that has its starting point in educational aims and needs rather than secondary effects on other sectors.

The main feature of Finland’s success is that it has curbed the most adverse effects of test-based accountability by simply not implementing them. Finland supports sample-based testing instead of high-stake standardized testing for entire age cohorts. This going against the grain also entails the publication of only the main results of these evaluations, which makes it impossible to form ranking lists and compare schools based on their performance. Furthermore, Finland has no national schools inspectorate. Instead, its decentralized system relies strongly on local self-evaluations conducted by education providers, mainly the 309 municipalities, and more than 2,000 individual schools. Central government’s role in evaluation is limited to fostering and promoting the local evaluation culture. It has no direct means to control the quality of education at school level (Eurydice 2015). Therefore, the use of these evaluations is geared to the development of education rather than to monitor, control and steer towards secondary purposes. We will argue further on in this section that the Finnish evaluation system, planned to support local development rather than top-down government, can be assessed as a partial, resilient policy success, as described in the PPPE model (see Chapter 1, ‘Introduction: Studying Successful Public Policy in the Nordic Countries’ in this volume) and by McConnell (2010).

Various researchers have highlighted the success of the Finnish evaluation model with different degrees of vigour. Most argue for the *programmatic benefits* of a light-touch evaluation system. The earliest and most direct claims concerning this were made by Sahlberg (2007, 2011), who argues that the Finnish model is more economically efficient and societally fair than the global mainstream. This view receives modest support from international research on the mainstream model’s weaknesses in improving learning outcomes, as already discussed (Verger and Parcerisa 2017).

In relation to the *process success*, the steps in the pathway and the principles of the evaluation scheme have not been publicly discussed. The design process was technocratic: the Finnish model was devised entirely within the education administration (Simola et al. 2009, 2013; Syrjänen 2013). The Nordic tradition of public transparency and collaborative governance does not fit well with this (Wallenius 2020). Hence, the score for process success is mixed: there was thorough analysis and deliberation, but in a closed arena.

Political success is the most apparent, where the quality policy, as uncentralized or even uncoordinated as it is, remains economical and teacher-friendly, bringing political kudos to those attached to it (Kalalahti and Varjo 2020). The main actor, the Finnish Education Evaluation Centre (FINEEC), receives virtually no public criticism—an indication of its good reputation—and signals a strong position as a coordinating and supporting actor in the spirit of what it calls ‘enhancement-led’ evaluation (FINEEC 2020).

The perspective of *endurance*—maintenance of success over time—is highly relevant to assessing and understanding the evaluation of education in Finland. The current system, developed in a specific historical context, has been subject to a considerable path dependency even when addressing the trends of international accountability-based testing policies (Kauko et al. 2020). In this first systematic scrutiny with the help of the PPPE framework, nuances and tensions emerge in addition to these claims of success.

On balance, we assess the system of the evaluation of Finnish comprehensive schools (a nine-year school of primary and lower secondary education, usually ages 7–16) as a partial policy success. The case shares some aspects with other Nordic countries but differs in some (Wallenius 2020). The similarity can be seen in the types of reform adopted following international trends, such as deregulation, decentralization, and the need to monitor and enhance quality. Where Finland differs from its Nordic peers is in the governance tradition and socio-historical peculiarities that are to some extent linked to its proximity to Russia and the lateness of its compulsory education, urbanization, industrialization and internationalization (see Dovemark et al. 2018; Simola et al. 2017).

We describe how the historical programmatic formation took place, and, once subjected to increasing global connections, how it found itself in crisis during the last decades of the twentieth century. The result was that relations between policy actors were recalibrated and then locked in as part of the decentralization of the 1990s. We then focus on how the formation of the quality evaluation policy took place; and how it was coordinated. The fairly recent institutionalization of the dispersed evaluation actors under one roof has consolidated continuities, but the rise of this unified player, FINEEC, and mounting external pressures have the potential to question the long-term design. We will analyse this as we discuss the system’s programmatic and political success and try to understand the robustness of support for the model.

Enduring Historical Threads

The programmatic formation of the evaluation of Finnish schools must be reflected in its long historical context. Our argument in relation to the programmatic success is that the need to formulate a full-scale programme took place only

in the late twentieth century. Before its articulation its formulation had happened in the long historical build-up of governance. These roots of governance came from the Nordic traditions to which Finland was organically attached when part of the Kingdom of Sweden and from the Russian governance regime when Finland was an autonomous Russian Grand Duchy (1809–1917). These eras also mark the first shift in education governance from church to state. Education governance was harnessed during the early days of independence (since 1917) to build a national narrative, and then during the post-war period to help expand education for the baby boomer generations (Lampinen 1998). The main institutional changes were the shift from folk and grammar schools, institutionalized in the 1860s, to a unified comprehensive school in the 1970s. Throughout these governance and institutional changes, according to Kauko et al. (2020; see also Jakku-Sihvonen 2001; Varjo et al. 2016), the programmatic base of the current Finnish evaluation system has sustained the essence of three historical practices: the *inspection of folk education*; the *local supervision (and surveillance) of folk schools by school boards*; and later the *participation in comparative learning studies*. All were readjusted during the economic crisis of the 1990s, which was when the concept of quality started to take root in Finland, and the consequent radical deregulation and decentralization of education governance. However, although inspection was abolished and the local governments' role changed, the essence of these practices was preserved in the new evaluation system. The emphasis on the local level and eased surveillance resulted in a considerable variation in how education providers evaluated the quality of education.

The *inspection of folk education* is the oldest thread in the current fabric of Finnish education evaluation. Although the inspectorate was abolished in the early 1990s, its basic function—the external evaluation of education providers—has remained part of the current set-up. Confirmation classes and parish catechetical meetings were the first institutionalized forms of folk education, spreading reading skills and Christian doctrine among the peasantry (Kähkönen 1982). The 1571 Church Order explicitly obliged priests to inspect the folk schools within their jurisdiction (Hanska and Lahtinen 2010). This ecclesiastical practice from Swedish rule was interrupted by the Russian state in 1841, when the tsar decreed that cathedral chapters must prepare a report on the condition of the folk schools in their jurisdiction for the Finnish Senate. Vuorela (1980) maintains that this was the first time such reports equipped the central administration continually to evaluate and control folk schools. When religious and secular authority were separated in the 1860s, the evaluation of folk education became the duty of the latter. This also meant the founding of the National Board of General Education in 1869. School inspections were set as one of its main tasks from the beginning.

This reassignment of inspections took place without any major public debate in the estate-based society of the time. Implementation was the responsibility of the Reverend Uno Cygnaeus, who was the first 'chief inspector of all folk

schools' (Harju 1988). The chief inspector was assisted by a body of inspectors who were charged to 'look after teaching, order, and discipline' (Law 1866) at the folk schools within their districts and report to the National Board of General Education. A century later, the political push from the parties on the left and centre disbanded the folk (elementary) and grammar schools (see Kauko 2019). The resulting implementation of the comprehensive school reform (1972–1977) steered the administrative position of the inspectorate in a formative direction, but regional inspectorate intervention still happened 'if required' (Decree 725/1984 in Varjo et al. 2016). In general, school inspections suited the bureaucratic, norm-driven, and top-down governance that conformed with the planning optimism of the 1960s and 1970s, when it was generally accepted that the comprehensive system should be implemented through strong top-down government (Kauko et al. 2020; Simola et al. 2017; Varjo 2007).

Yet the comprehensive school brought with it newly progressive pedagogical thinking, and the duties of inspectors were redefined as 'guidance,' 'supervision' and 'data collection' at the beginning of the 1980s. These were the first steps towards decentralized governance and the school-based development of education (Pitkänen 2019). The final push came from outside the education system. At the beginning of the 1990s the policies of decentralization and deregulation found an ideological sounding board in the centre-right governments of the time, and the broad governance reforms resulted in the abolition of the school inspection system (Law 182/1991; Rinne et al. 2011). Instead, education providers and their schools were now obliged to participate in external evaluations by order (Law 628/1998; Opetushallitus 1994).

Another thread of *local evaluation* was developed simultaneously with folk education as part of changes in the municipal administration. The Folk School Act came into effect in 1866 (see Law 1866) after the judicial separation of municipalities from the parishes in 1865. Typically, towns had a municipal board for all folk schools, and sparsely populated rural municipalities had a single board for each school. The school board's main task was to supervise the provision of folk education at close hand, 'by means of inspection, regularly visiting schools, attending examinations and comprehensively monitoring school keeping' (Law 1866).

The school board's main task—to supervise the provision of folk education at close hand—is an original element of *endogenous* control in the current fabric of Finnish quality evaluation. A century later the comprehensive school reform largely sustained the position of school boards. The Comprehensive School Act (see Law 476/1983) created a board for each comprehensive school and assigned numerous duties to it based on guidance, supervision, and development. In addition to the traditional tasks of school boards, the supervision of student welfare (e.g. nursing, school meals and transport) was now included in school boards' duties (Kauko et al. 2020). The most radical change took place in the decentralization reforms of the 1990s, when school boards were decreed to be non-obligatory

representative bodies (Law 707/1992; Law 365/1995). School boards are currently required of only private and state education providers, which in their pupil volume constitute a small minority alongside the main bulk of municipal education providers. The change again emphasizes the policies of decentralization and deregulation of the 1990s, which were supported by the centre-right PMs Harri Holkeri (1987–1991) and Esko Aho (1991–1995). In practice, the evaluative function of school boards was replaced by the local self-evaluation of education providers (Law 628/1998). Local evaluation thus wove a thread into the fabric of Finnish evaluation programmes by emphasizing the responsibility of the local education provider to keep track of its education quality.

The third thread started its weaving after the 1950s, when Finland's continuous *participation in comparative learning studies* created the national capacities for evaluation. They laid the foundation for a national evaluation system of learning outcomes (Kangasniemi 2004) and participation in the international development of education indicators (Leimu 2004). After World War II policy-relevant studies in comparative education started to evolve towards a more systematic and rigorous methodological approach: data were collected domestically from each nation, but within a common international research design with the same methods and measurements. According to Alasuutari (1996: 108) between the Second World War and Finland's accession to the European Union in 1995 it was a planning economy, convinced that 'social problems could be solved best by "scientific" planning and organizing. Solutions to problems concerning the national economy or the functionality of state administration were found solely through better planning'. Participation in comparative learning studies since the 1950s was also tightly connected with the period's idea of the planning economy. The International Association for the Evaluation of Educational Achievement (IEA) was built around problematizations associated with the evaluation of school effectiveness and student learning. According to Husén (1967; see also Kauko et al. 2020), the founders of the IEA viewed the world as a natural education laboratory in which different school systems experimented in various ways to obtain optimal results in educating their young people. They assumed that if comparative education research could obtain evidence from a wide range of systems, the variability would be sufficient to reveal important results. Laukkanen (1998) demonstrates that participation in international cooperation enhanced national expertise and capacity-building in the evaluation of education.

The OECD launched its indicators of the 'Performance of Educational Systems Project' in 1970 (see Kauko et al. 2020). The framework comprised 46 indicators, constructed by combining the statistics of student ratios and flows, education resources, and costs with background variables (OECD 1973), for example. In response, the Finnish Ministry of Education funded a 'Project on education indicators' at the University of Jyväskylä to study the opportunities to construct a system of education indicators to assist quantitative education planning, which ran

from 1972–1974 (Olkinuora and Perkki 1974). By the mid-1970s however, indicators had been omitted from the national education policy agenda, because there was no technical capacity, and political interest in the OECD approach had faded. Capacity-building with these national evaluation trials meant that once political interest re-emerged, and the societal and technological conditions allowed it, a new era of indicator-based evaluation could start from the 1990s (Kauko and Varjo 2008).

Tying these historical threads together helps us see how the programme of Finnish education evaluation builds on the division of governance between the central and local authorities. Green et al. (1999) show that Nordic nation states during the nineteenth century commonly delegated control to their subordinate bodies, the municipalities. The municipal provision of folk education and local supervision of schools by school boards are manifestations of contemporary ‘central–local relations of governance’ (Ozga et al. 2011). In turn, the comprehensive school reform of the 1970s, propelled by coalition governments of the Social Democrats and the Agrarian League, marked the peak of centralized governmental education steering. The old control mechanisms—both the inspection of folk education and the local supervision of folk schools by school boards—were effortlessly transposed to a different societal situation than that in which they had originally emerged.

Green et al. (1999) also note that the tradition of ‘Nordic localism’ experienced a revival during the 1980s. The heavily centralized planning and steering system in education, which had been under construction for decades and had reached its peak during the comprehensive school reform (1972–1977) was dismantled in 1988 by a right-left Holkeri government as part of a sweeping ambition to reform the entire state services sector. The new administrative ethos, implemented by a new generation of civil-service leaders (including the formidable director-general of the National Board of Education, Vilho Hirvi) emphasized deregulation and management by results (Simola et al. 2017). The intention, articulated by the Ministry of Education, was now to increase the quality of education by ‘increasing flexibility and choice’ and introducing new evaluation mechanisms (Opetusministeriö 1991). While it had previously been believed that educational goals could be achieved via strict norm steering and careful implementation, national policymakers now moved towards setting national core targets, allowing actors in the decentralized system to pursue these goals as they saw fit, and evaluating the end results *ex post*. This shift entailed also a second reorganization of central–local governance relations (Ozga et al. 2011). The subsequent re-emergence of Nordic localism and decentralization created a situation in which central governance lost its grip on education providers (Eurydice 2004). In this situation it became necessary to begin to both conceptualize and reconceptualize what a suitable evaluation programme for Finnish comprehensive schools could look like.

Overcoming Political Fault Lines

As noted, in the early 1990s the evaluation programme was reformulated, but with familiar features from the past. This work laid down the fundamentals of Finnish education evaluation, which [Simola et al. \(2009, 2013\)](#) have recognized as a strategy of public ‘non-articulation’ of policy principles, according to which there should be no public data, no rankings and only sample-based tests providing data for development purposes. Interviews of the sample-based evaluation model’s main architects indicate that economic and pragmatic considerations also played into it: sample-based evaluation was cost-effective and provided enough information for national policymakers’ purposes. They also shared a belief that high-quality teacher training would drive continuous improvement within schools and superimposing annual rankings exercises would only create and exacerbate inequalities ([Syrjänen 2013](#): 86–90). In sum, many parts of the programme drew on a combination of the historical pathways described in the previous section and the need to make a virtue of the necessity to face the new situation of decentralized and deregulated governance (see [Kauko et al. 2020](#)).

In the deregulated and decentralized environment of the 1990s many actors undertook their own, only partially overlapping initiatives in the emerging field of quality evaluation. This directly affected political relations. Decentralization reinforced the gap between national and local governments’ preferences, while deregulation ensured that the central government had fewer tools to hand for influencing the municipalities. The National Board of General Education and the National Board of Vocational Education were replaced by the National Board of Education (NBE) in 1991, and this new agency was set up more to advise and foster policy-relevant knowledge than as a top-down regulator ([Government Bill 225/1990](#)). At the same time the various municipal collaboration organizations were organizing, forming the Association of Finnish Local and Regional Authorities (AFLRA) in 1993 ([Government Bill 347/1992](#)).

The new organizational set-up represents only one current in the larger stream of change but it certainly contributed to the formation of what could be called a current evaluation programme. Preparations for it had already started in the early 1980s, when the Ministry of Education tasked the National Board of General Education with sketching a model to evaluate the effectiveness and outcomes of comprehensive and upper secondary education. This was motivated by the legislative change in education and related development plans ([Kouluhallitus 1982](#); [Opetusministeriö 1983](#)), while more responsibility for school development was directed to the local level. The working group set up for this task drafted its suggestion for a national system to evaluate the outcomes and performance of education in the mid-1980s ([Kouluhallitus 1986a, 1986b](#); [Lyytinen et al. 1989](#); see also [Lahelma 1984](#)). In addition to national data gathering (e.g. from pre-existing registers, databases and studies) for the purpose of external evaluation,

the working group introduced the idea of local and school self-evaluation. In the spirit of emerging decentralization and regulation policies, the group suggested that in addition to ongoing research on and evaluation of educational outcomes, the performance of education institutions should increasingly be evaluated at local level, and especially by the schools themselves (Kouluhallitus 1986b; Lyytinen et al. 1989). This was where education took place. Although this model (Kouluhallitus 1986a, 1986b) was never adopted, it was the first to raise the issue of education performance to the level of a policy programme.

The process of developing the national evaluation model was continued by a Ministry of Education working group (Opetusministeriö 1990), but with little effect (Laukkanen 1998). The task was then reassigned to the newly established National Board of Education. The model's development became a timely issue, as the Ministry of Education also highlighted the need for quality evaluation in its newly established strategic steering document for education (Opetusministeriö 1991). As a result, under the leadership of long-term Finnish school developer Ritva Jakku-Sihvonen, *The Framework for Evaluating Educational Outcomes* (Opetushallitus 1995, 1998) emerged in the deliberations of a working group consisting of officials of the National Board of Education and university experts. Its development was helped by several educational institutions experimenting with practical use of the framework in municipalities and schools.

Corresponding to the New Public Management doctrine of the era, the framework held that economy, efficiency and effectiveness ought to be the main evaluation criteria not only at national level, but also in regional, municipal and school-level evaluations. The model also introduced the structure of the evaluation system, consisting of complementary and reciprocal elements of international, national external and local internal evaluations, which closely followed the historical threads described in the previous section. This division of stages, complementary elements and national responsibilities has continued until today (Karvi 2020). Although the model was given central and programmatic status, it was officially only a recommendation (Simola et al. 2009). However, its programmatic centrality remains visible in that the model's basic ideas, for example the idea of the three Es, is still present in current Finnish evaluation policy documents and programmes (e.g. Opetus- ja kulttuuriministeriö 2012).

In another process in the early 1990s the new National Board of Education (Opetushallitus 1994) coordinated and drafted the new curriculum, which called for school self-evaluation, introducing it as an integral part of continuous school and curriculum development (Pitkänen 2019). The curriculum claimed that school self-evaluation should be comprehensive and systematic. Self-evaluation should focus piece by piece on the school's diverse areas, aspects and operations so that it would finally encompass all areas of schooling and the curriculum. The task of the municipality, the education provider, was to support and frame the school's self-evaluation. (Opetushallitus 1994.)

The increased importance of the local level brought with it the risk of great disparities of approaches, and the vulnerabilities of small jurisdictions lacking the capacity to deliver created a need for the central government to build evaluation skills. Non-binding national recommendations or guidelines for fostering local evaluation were published by the different sides: the national education authorities ([Opetushallitus 1995, 1998](#), [Opetusministeriö 2010](#); [Opetus- ja kulttuuriministeriö 2012](#)) and the Association of Finnish Local and Regional Authorities, representing the Finnish municipalities ([Suomen Kuntaliitto 2003a, 2003b, 2006a, 2006b](#)). As decentralization had obliterated most legal requirements for evaluation, there was a strong push for a quality evaluation culture in schools and municipalities promoted by projects, courses and the vast evaluation literature ([Pitkänen 2019](#)). It was ultimately up to the municipalities to decide what to choose and what to leave out from the broad array of suggestions.

This extensive local freedom was partly curbed in the Basic Education Act (see [Law 628/1998](#)), in which the principle of local evaluation received legal status rather than the status of a recommendation. The [Government Bill \(86/1997\)](#) leading to this act was issued by the PM Social Democrat Paavo Lipponen's first left-right government led by the (1995–1999). The quality policy section of this law was politically uncontroversial ([Kalalahti and Varjo 2020](#)). It drew on four years of preparation at the Ministry of Education that had explicated the need to clarify and conceptually define the terms, (national) external and (local) self-evaluation. According to the law each education provider should self-evaluate the education they organized and participate in external evaluations ([Law 628/1998](#)). The schools should also conduct self-evaluations as part of the local evaluation of education ([Opetus- ja kulttuuriministeriö 2012](#): 36).

More recently, to support education providers and schools in their obligation of self-evaluation in line with Matti Vanhanen's second centre-right government (2007–2010) programme, the Ministry of Education drafted a new recommendation, Quality Criterion for Basic Education ([Opetusministeriö 2009, 2010](#); [Opetus- ja kulttuuriministeriö 2012](#)). This was done in cooperation with other experts, authorities and stakeholders, for example the NBE, Finnish Education Evaluation Council, the Association of Finnish Local and Regional Authorities, Finnish universities, regional and municipal authorities, and school personnel. Nevertheless, the implementation of the Quality Criterion recommendation in schools and municipalities has been voluntary. The main principle has been that each education provider has strong autonomy in determining the objects, targets and methods of their self-evaluation ([Kauko et al. 2020](#)). In 2016, 41 per cent of education providers used the Quality Criterion as their principal tool of self-evaluation ([Harjunen et al. 2017](#)).

Binding these developments, the self-evaluation conducted at the local level since the 1990s, both in municipalities and schools, has been argued to constitute the foundation of the Finnish quality evaluation and evaluation system of

education, which the other levels of evaluation system support (Pitkänen 2019: 121). The complementary and supportive role of the national and local evaluation, along with the idea of enhancement-led evaluation, has become deeply embedded in the Finnish quality evaluation discourse. It can be traced in subsequent legislation (e.g. [Government Decree 1317/2013](#) on the Finnish Education Evaluation Centre), as well as in the quality criterion for basic education ([Opetus- ja kulttuuriministeriö 2012](#)) and the National Core Curriculum ([Opetushallitus 2014](#)). The latter states:

The purpose of both the internal and external evaluation is to support educational development and to improve conditions for learning. The monitoring, regular evaluation and development of the local curriculum and annual plans are part of this duty... . In their self-evaluations, the education provider and schools can draw on the results of national evaluations and development projects and the national quality criteria for basic education. Cooperation with pupils, guardians and other actors promotes transparent and constructive self-evaluation.

([Opetushallitus 2014](#))

However, there is another side to this story. Despite the emphasis on local evaluations, since the implementation of the Basic Education Act ([Law, 628/1998](#)) there has been continuing concern that education providers and schools have not met the expectations regarding self-evaluation (see e.g. [Löfström et al. 2005](#); [Sivistysvaliokunnan mietintö 11/2002, 9/2009](#)). According to an evaluation report only around 25 per cent of education providers had used the National Board of Education model in their self-evaluations in 1998. The Balanced Scorecard was second with 20 per cent, while EFQM (European Foundation for Quality Management) and other quality models were much less popular. However, the vast majority, around 70 per cent, reported that they used a combination or their 'own' model, which leaves open the question of how systematic and robust those local evaluations are ([Rajanen 2000](#): 18).

Furthermore, the State Auditor's Office noted that the local level was largely happy with the 'information steering' at the national level, yet it recommended that such steering should be used only as a complement to other forms (laws, decrees and resources), and that securing rights such as basic education should rely on stronger steering forms ([Valtiontalouden tarkastusvirasto 2009](#)). According to a recent evaluation report on the state of local evaluation ([Harjunen et al. 2017](#)), the evaluation culture had not been bedded down sufficiently more than 15 years after the Basic Education Act came into force in 1999. In many cases there appears to be a lack of a functioning self-evaluation system or a systematic assessment culture. Examples of the opposite are few. According to these evaluations of evaluation, no education provider has achieved a level of excellence in their self-evaluation

system. Most providers (60%) report being at the initial level of implementation (Harjunen et al. 2017: 42).

Alternative views advocating a more stringent and top-down evaluation system have been scarce because of the political commitment to the existing programme. Test-based accountability has not been seriously debated in Finland since the 1990s (Simola et al. 2009), but the lack of standardized tests in the Finnish model has recently raised some scholarly eyebrows. Unevenness in grading due to the model's non-standardized nature and teacher-based grading—instead of mandatory national testing—has led some evaluation researchers to argue that the resulting variety in grades at the completion of comprehensive school threatens equal access to further education, especially in the capital region (Hotulainen et al. 2016: 50–51). This may be used to support moves towards standardized tests. In general, the concerns have not gained political leverage, possibly because of the very strict tradition of non-politicization concerning this matter (Wallenius 2020). However, in response to this challenge, the National Agency for Education (formerly the National Board of Education) recently provided clearer guidelines for evaluating students at this stage, starting from 1 August 2021 (Opetushallitus 2021).

In sum, through the institutional changes and decentralization policies since the early 1990s the concepts of quality and quality evaluation have entered and been consolidated in the Finnish evaluation discourse and practices of education governance in the field of comprehensive education. At the same time, it has become obvious that the quality of education should not only be evaluated at the national level but where the provision of education takes place—in municipalities and schools. The data acquired at both levels should interact in producing an overall picture of the quality and performance of comprehensive education (Pitkänen 2019). In this set-up, the National Board of Education has aimed to develop education through national external evaluations, and the municipalities' association has represented the education providers responsible for the local development and evaluations. Whereas the former has focused on evaluation especially from the perspective of education and learning, the latter has focused on education as a municipal basic service as part of other municipal services and local policies. This has led to some conflict between these coalitions concerning whether education constitutes a distinct sphere separate from societal services, or whether it is also evaluable according to the same principles and methods as other public services.

This dispute has become less visible in recent years, possibly due to the foundation of FINEEC and the official formation of the national evaluation programme, and the legal recognition of education providers as autonomous local evaluation actors (see Pitkänen 2019; Simola et al. 2009). This understanding of the production of complementary evaluation data at local and national levels was accompanied by what was framed and later named as *a programme of enhancement-led evaluation*. This principle fitted the decentralized landscape well,

because it suggested evaluation data should be used for development rather than control and accountability (Kauko et al. 2020; Pitkänen 2019; Simola et al. 2009). However, a built-in feature after the 1990s was an uncoordinated division of labour between the national and local, which leads us to the conclusion that success in the programme's coalition forming may be contested, while other aspects of the process could be seen as successful, especially the preservation of original ideas, legitimacy, symbolic value, and eventually a lack of opposition (McConnell 2010: 352). We will continue the analysis of coalitions in the next section, which will focus on the institutionalization of this programme through changes in actor relations.

Institutionalization: Towards a Stronger Quality Policy?

The formulation of the quality programme created a latent conflict between advocates of the national and the local levels, and the resulting evaluation model was a compromise between these two parties. Likewise, there were divisions among the national evaluation actors, who were dispersed and doing partly overlapping work.

The problematic role of the National Board of Education as the evaluator and executive central government body emerged at the debate on the Basic Education Act in 1998 (Temmes et al. 2002; Varjo et al. 2016). The doubts were reinforced when parliament required the Ministry of Education to evaluate the evaluation function after the new law came to force. The Ministry report concluded that an independent evaluation organ should be formed. The Standing Committee for Education supported the formation of a new organization alongside but independent of the National Board of Education (Sivistysvaliokunnan lausunto 5/2001vp). The Finnish Education Evaluation Council (FEEC) was founded in 2003. However, the field remained dispersed, and the idea of a large merger was taken forward by the ministry and then implemented by the left-right government of Jyrki Katainen (2011–2014), as we describe in Kauko et al. (2020: 9):

‘The Finnish Education Evaluation Council (FEEC) started in 2003 in conjunction with the Ministry of Education. This meant there were now three organizations to conduct the external evaluation of comprehensive schools: the FEEC was responsible for institutional and general evaluations, the National Board of Education for sample-based learning outcomes testing, and provincial governments conducted theme-based evaluations or indicator reports for municipal education. In addition, a Finnish Higher Education Evaluation Council (FINHEEC) was formed in 1996. The FEEC was tasked to draft an evaluation plan and provide information for education policy and setting new goals (Opetusministeriö 2004). The role and work of the FEEC as part of the evaluation landscape was supported and problematized in subsequent years. A working group of the Ministry of Education (Opetusministeriö 2007) recommended that the FEEC and FINHEEC

should continue as they were, and a rapporteur suggested they should be combined as a national evaluation agency (Lankinen 2007). Both reports suggested that the evaluation of learning outcomes should be the responsibility of the National Board of Education. The rapporteur also saw the evaluation conducted by the provincial governments as problematic, because the provincial governments were also responsible for legality control (Lankinen 2007). The idea of a single evaluation council hosting all the evaluation functions was adopted in the next government programme (Government 2011), and the Finnish Education Evaluation Centre (FINEEC) was inaugurated in 2014' (Kauko et al. 2020: 9).

In other words, the National Board of Education lost its jurisdiction in evaluating learning outcomes (Jakku-Sihvonen 2014). Now all evaluation activities were concentrated in a single organ. The long-term problem had been how evaluation would be credible if it were conducted by the same organs that governed education (e.g. Opetus- ja kulttuuriministeriö 2012).

FINEEC brought greater rigour and trustworthiness to external evaluation, employing a model that had previously been used in the higher education sector (for FINHEEC). Its founding was the culmination of a development that had started in the 1990s as the post-deregulation reorganization of evaluation functions—while the basic principles for evaluation design remained the same—to assist in the development of local quality processes and provide external evaluation data, not to use data to control the education providers (Kauko et al. 2020).

FINEEC's programme relies on the long-evolved and sustained principles of enhancement-led education: the promotion of 'participation and interaction'; tailoring of evaluation according to the evaluand; and the provision of 'current state analysis, conclusions and recommendations that can be used both by actors at the national, regional and local level and by education and training providers, higher education institutions and stakeholders' (FINEEC 2020: 5). These principles follow the same aims that were drafted in the National Board of Education models during the 1990s. They also imply that the use of evaluation is not to control but to develop, and that the use of data takes place among professionals. The previously implicit goals (Simola et al. 2009, 2013) of not compiling ranking lists and publicizing performance data except for development purposes were given a firm footing (FINEEC 2020: 7): 'By trust we mean that the evaluation participant can talk openly about successes, development needs and shortcomings without having to fear sanctions. We collect and use evaluation data for the benefit of those who are evaluated. FINEEC does not compile any ranking lists'. These goals had endured.

The local–national cleavage still exists, but the allure of the national camp has been strengthened by the establishment of the Finnish Education Evaluation Centre. During 2003 and 2013 national steering became more uniform and more influential. Indeed, evaluation during the FINEEC era has become more

prominent now that all evaluative functions are under one roof. FINEEC drafts an evaluation plan every five years that combines all levels of education. It has also suggested increasing the role of evaluation with the aid of a quality assessment strategy (Väättäinen 2019: 26).

Recently, the centre-left PM Sanna Marin government (2020–) has issued an education policy report that aims to set the direction for education for the next two decades. In the report before parliament the direction appears to differ from the long-term evaluation track. The reasons for this change seem to be related to an understanding that demographic challenges to equal education provision can be solved with the aid of technology and data, a solution that could tip the governance scale in favour of the national level (Valtioneuvosto 2021). The report suggests that ‘clear and binding quality goals and indicators to follow their implementation’ should be established for the organization of early childhood education, pre-primary education and the comprehensive school (Valtioneuvosto 2021: 22). If this were implemented nationally, it would mean a substantial turnaround in the Finnish quality evaluation policy, which has previously aimed to develop rather than control the nature of evaluation, and has operated on the basis of trust in municipalities and school professionals.

Conclusion: From Tacit Knowledge to Institutionalized Assessment

Historical path dependencies formed some of the basic conditions under which the evaluation of quality in Finnish schools was established. In this chapter, we have sought to analyse the more or less articulated goals and the purposeful and coincidental actions that shaped the pursuit of these goals. Parts of the policy process could not be reconstructed because much of the work occurred behind closed doors. Paradoxically, though undermining the ‘democratic’ component of process quality, the disciplined attention given to the challenge of building a more meaningful evaluation system befitting Finnish conditions was also a key part of both the programmatic and the political success of the evaluation model that was created.

The designers of the National Board of Education model have reported how the design process was carefully fitted to balance the values of equality and cost-effectiveness. The internal debate on the matter became enervated through the misalignment of local and national roles and interests, as represented by the National Board of Education and the Association for Regional and Municipal Authorities. And yet, because both parties shared the same values and vision, they could eventually come to an agreement.

Thirty years on, the delivery of the quality model remains unfinished business. Many municipalities still effectively lack a proper quality system in education. Even with this feature, the education system has remained steadfast in its

commitment to not competing for students with the aid of indicators, and to the principle that the same school services should be provided to every pupil as much as possible.

Finland has been able to develop and sustain a quality evaluation model for schools that is distinctive from most other education systems in the world. This ability to buffer an equality-based system against more competitive models from international examples has previously been recognized (Simola et al. 2017). The idea of enhancement-led evaluation is the cornerstone of FINEEC (2020). While the National Board of Education evaluation model failed to flourish in most municipalities, its main aims are institutionalized in the basic philosophy of the national evaluator, FINEEC. The original aims of equality are supported by the features of what was long an unarticulated though now stated programme of no rankings or standardized tests for the whole age cohort.

The paradox thus emerges: it was precisely the combination of the political silence (Wallenius 2020) and professional monopolization of evaluation that has given birth to the formation of a distinctive model. The leeway it provided to experts to do their work without a political steer has certainly contributed to the model's coherence. In other words, as more politically sensitised policymakers tend to be enthused by being seen to be embracing international trends, in this instance the lack of appetite for policy importation due to the preponderance of local professional experts in the policy process may have opened a space for the formulation of a distinctive, 'Finnish' model.

The Finnish evaluation model has resulted in a political success scenario with many winners and strong reputational benefits. At the national level it is noticeable how little public discussion about this approach to evaluation there has been. The strong continuity between governments of practically all the main coalition combinations since the 1990s is no less noteworthy. At the local level Finland lacks the usual complaints of teachers overwhelmed by the testing and inspection regime (see, for example, Gurova et al. 2018). Nor have the municipalities attempted to politicize the issue.

Only the latest developments, with FINEEC's growing importance and the Marin government's new report opening possible avenues for more stringent evaluation, may indicate a future ripple. There are also worrying trends in relation to school segregation through parental choice (Kosunen 2016; Seppänen 2006; Seppänen et al. 2015) and increasing social and spatial inequalities (Bernelius and Huilla 2021; Bernelius et al. 2021; Huilla 2022), which usually do not benefit from more evaluation. Until now the base level of the Finnish system in international comparison has provided equal results, and the lack of ranking has restrained segregation.

The programme of enhancement-led evaluation has remained the same for three decades. The basic idea of local and national evaluation has remained the same for more than a century, but was always reshaped according to the reform

needs of different times. In sum, we argue that the programme and political elements of the Finnish comprehensive school's quality evaluation have been a success that has also shown temporal durability. However, the success is only partial, given the processual shortcomings generally linked to the lack of transparency and discussion. These questions have been better articulated with the centralization of evaluation functions in FINEEC, which now wields a lot of power in evaluation. A more unified form of evaluation may lead to more centralized and sudden changes, possibly risking the temporal endurance. FINEEC has a professional monopoly position from which to argue for potential changes in evaluation, which makes change easier. In addition, as the latest government report shows (Valtioneuvosto 2021), global trends of data-based government are gaining ground in Finnish education. What we know from previous studies of evaluation is that it creates more evaluation (Kauko et al. 2018b). It remains to be seen if FINEEC and future governments will maintain the long-standing tradition or go with the international flow. Their hands are relatively free, for the general public has not been accustomed to the discussion of evaluation.

Questions for discussion

1. What changed in the 1990s in Finnish education evaluation?
2. How intentional were the effects of deregulation in Finland?
3. Which were the main organizational players in sustaining continuity?
4. What similarities and differences are there between Finnish and global mainstream education evaluation?

Links to online resources

SAWA Atlas, 'information about the international adoption of large scale assessments, the main features and uses of these policy instruments, and the accountability measures attached to them':

<http://reformedproject.eu/sawa-atlas/>.

The Finnish Education Evaluation Centre:

<https://karvi.fi/en/>.

References

- Alasuutari, P. 1996. *Toinen tasavalta: Suomi 1946–1994* [Second Republic: Finland 1946–94]. Tampere: Vastapaino.
- Bernelius, V. and H. Huilla. 2021. 'Koulutuksellinen tasa-arvo, alueellinen ja sosiaalinen eriytyminen ja myönteisen erityiskohtelun mahdollisuudet'. Valtioneuvosto. <https://doi.org/http://urn.fi/URN:ISBN:978-952-383-761-4/>.

- Bernelius, V., H. Huilla and I. Ramos Lobato. 2021. “Notorious Schools” in “Notorious Places”? Exploring the Connectedness of Urban and Educational Segregation. *Social Inclusion*, 9(2): pp. 154–165.
- Decree (725/1984). Asetus lääninhallituksen kouluosastosta annetun asetuksen 2 ja 4 §: n muuttamisesta.
- de la Porte, C., G.B. Eydal, J. Kauko, D. Nohrstedt, P. 't Hart and B.S. Tranøy. 2022. ‘Introduction: Studying successful public policy in the Nordic countries’. In *Successful Public Policy in the Nordic Countries*, Chapter 1. Oxford: Oxford University Press.
- Dovemark, M., S. Kosunen, J. Kauko, B. Magnúsdóttir, P. Hansen and P. Rasmussen. 2018. ‘Deregulation, privatisation and marketisation of Nordic comprehensive education: Social changes reflected in schooling’. *Education Inquiry*, 9(1): pp. 122–141.
- Eurydice. 2004. *Evaluation of schools providing compulsory education in Europe*. Brussels: Eurydice, European Unit.
- Eurydice. 2015. Assuring quality in education: Policies and approaches to school evaluation in Europe. Eurydice Report. Luxembourg: Publications Office of the European Union.
- FINEEC. [Finnish Education Evaluation Centre] 2020. ‘Enhancement-led evaluation at the Finnish Education Evaluation Centre. Summaries 11:2020’. Authors Sirpa Moitus and Päivi Kamppi. Helsinki: FINEEC. https://karvi.fi/app/uploads/2020/04/KARVI_T1120.pdf. Accessed 10 June 2021.
- Government. 2011. ‘Programme of Prime Minister Jyrki Katainen’s government’. June 22. <https://vm.fi/en/publication?pubid=3606/>. Accessed 14 December 2021.
- Government Bill (225/1990). ‘Hallituksen esitys eduskunnalle laiksi opetushallituksesta’. <https://www.eduskunta.fi/valtiopaivaasiakirjat/HE+225/1990/>. Accessed 6 December 2021.
- Government Bill (347/1992). ‘Hallituksen esitys Eduskunnalle laiksi Suomen Kuntaliiton perustamiseen liittyvistä järjestelyistä’. <https://www.edilex.fi/he/fi19920347.pdf>. Accessed 6 December 2021.
- Government Bill (86/1997). ‘Hallituksen esitys Eduskunnalle koulutusta koskevaksi lainsäädännöksi’. <https://www.eduskunta.fi/valtiopaivaasiakirjat/HE+86/1997/>. Accessed 6 December 2021.
- Government Decree (1317/2013) on the Finnish Education Evaluation Centre. ‘Valtioneuvoston asetus Kansallisesta koulutuksen arviointikeskuksesta’.
- Green A, A. Wolf and T. Leney. 1999. *Convergence and divergence in European education and training systems*. London: Institute of Education, University of London.
- Gurova, G., Candido, H. and Zhou, X. (2018). Effects of quality assurance and evaluation on schools’ room of action. In Kauko J., Rinne R. & Takala T. (eds.) *Politics of quality in education: A comparative study of Brazil, China, and Russia*, pp.137–160. London: Routledge.
- Hanska, J. and A. Lahtinen. 2010. ‘Keskiajalta 1500-luvun lopulle’ [From the Middle Ages to the end of the sixteenth century]. In *Huoneentaulun maailma. Kasvatus ja koulutus Suomessa keskiajalta 1860-luvulle*, edited by J. Hanska and K. Vainio-Korhonen, pp. 17–111. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Hanushek, E.A. and L. Woessmann. 2007. *Education quality and economic growth*. Washington, DC: The World Bank.
- Hardy, I. 2021. ‘The quandary of quantification: Data, numbers and teachers’ learning’. *Journal of Education Policy*, 36(1): pp. 44–63.

- Harju, A. 1988. *Maalaiskansakoulujen tarkastustoimi vuosina 1861–1921* [The inspection of rural schools in 1861–1921]. *Studia historica Jyväskylälensia* 38. Jyväskylä: Jyväskylän yliopisto.
- Harjunen, E., R. Hietala, L. Lepola, A. Räisänen and A. Korpi. 2017. 'Arvioinnilla luottamusta. Perusopetuksen ja lukiokoulutuksen järjestäjien laadunhallinta- ja itsearviointikäytänteet'. Helsinki: Kansallinen koulutuksen arviointikeskus. 4:2017 https://karvi.fi/app/uploads/2017/01/KARVI_0417.pdf. Accessed 6 December 2021.
- Hotulainen, R., A. Rimpelä, J. Hautamäki, S. Karvonen, J.M. Kinnunen, S. Kupiainen, ... and T. Wallenius. 2016. 'Osaaminen ja hyvinvointi yläkoulusta toiselle asteelle: Tutkimus metropolialueen nuorista.' [Learning and well-being: Transition from lower to upper secondary education in the Helsinki Metropolitan Area.] Helsinki: University of Helsinki, Faculty of Behavioural Sciences. <https://trepo.tuni.fi/handle/10024/100405/>. Accessed 6 December 2021.
- Huilla, H. (2022). *Kaupunkikoulut ja huono-osaisuus*. [Disadvantaged Urban Schools.] Helsinki: Helsingin yliopisto. Accessed 4 May 2022: <https://helda.helsinki.fi/handle/10138/341436>
- Husén, T. 1967. *International study of achievements in mathematics: A comparison of twelve countries*. Volume II. Uppsala: Almqvist and Wiksell.
- Jakku-Sihvonen, R. 2001. 'Arvioinnin kehittyminen osaksi julkishallintoa ja koulutuksen ohjausjärjestelmää' [The development of evaluation into a public administration and steering system of education]. In *Johdatus koulutuksen uudistuvaan arviointikulttuuriin*, edited by R. Jakku-Sihvonen and S. Heinonen, pp. 9–21. Helsinki: Opetushallitus.
- Jakku-Sihvonen, R. 2014. *Arviointia akanvirrassa: Opetushallituksen arviointitoiminnan historiikki*. Helsinki: Opetushallitus.
- Kähkönen, E. 1982. 'Kansanopetus Suomen Lapissa ennen kansakoulua' [Folk education in Lapland before folk schools]. *Lapin korkeakoulun kasvatustieteiden osaston julkaisuja*. Sarja A:1. Tutkimuksia. Rovaniemi: Lapin korkeakoulu.
- Kalalahti, M. and J. Varjo. 2020. 'Revisiting universalism in the Finnish education system.' *Research on Finnish Society*, 12(1): pp. 1–16.
- Kangasniemi, E. 2004. 'Matematiikalla se alkoi' [It all started with mathematics]. In *Kansainväliset IEA-tutkimukset Suomi-kuvaa luomassa*, edited by K. Leimu, pp. 99–121. Jyväskylä: Koulutuksen tutkimuslaitos.
- Karvi. 2020. *Koulutuksen arviointisuunnitelma 2020–2023*. Helsinki: Kansallinen koulutuksen arviointikeskus.
- Kauko, J. 2019. 'The Finnish comprehensive school: Conflicts, compromises, and institutional robustness'. In *Great policy successes*, edited by P. 't Hart and M. E. Compton, pp. 122–142. Oxford, UK: Oxford University Press.
- Kauko, J., T. Takala and R. Rinne. 2018a. 'Comparing politics of quality in education'. In *Politics of quality in education: A comparative study of Brazil, China, and Russia*, edited by J. Kauko, R. Rinne and T. Takala, pp. 1–17. London: Routledge.
- Kauko, J., R. Rinne and T. Takala. 2018b. Conclusion. In *Politics of quality in education: A comparative study of Brazil, China, and Russia*, edited by J. Kauko, R. Rinne and T. Takala, pp. 180–190. London: Routledge.
- Kauko, J. and J. Varjo. 2008. 'Age of indicators: Changes in the Finnish education policy agenda.' *European Educational Research Journal*, 7(2): pp.219–231.
- Kauko, J., J. Varjo and H. Pitkänen. 2020. 'Quality and evaluation of Finnish schools.' *Oxford research encyclopedia of education*.

- Komatsu, H. and J. Rappleye. 2017 'A new global policy regime founded on invalid statistics? Hanushek, Woessmann, PISA, and economic growth.' *Comparative Education*, 53(2): pp.166–191.
- Kosunen, S. 2016. *Families and the social space of school choice in urban Finland*. Studies in Educational Sciences 267. Helsinki: University of Helsinki, Institute of Behavioural Sciences.
- Kouluhallitus. 1982. *Peruskoulun kehittämissuunnitelma 1980-luvulle*. Kouluhallituksen ehdotus opetusministeriölle 30. 11.1982. Helsinki: Kouluhallitus.
- Kouluhallitus. 1986a. *Peruskoulun ja lukion tuloksellisuus. Ehdotus tuloksellisuuden arviointijärjestelmäksi Kouluhallituksessa sekä peruskoulun tilannekartoitusta vuodelta 1986. Osa 1. Performance of comprehensive and upper secondary schools. Kouluhallituksen alaisen koulutoimen ja vapaan sivistystyön tuloksellisuutta käsittelevä työryhmä*. Helsinki: Kouluhallitus.
- Kouluhallitus. 1986b. *Koulutuksen tuloksellisuuden prosessiarviointi. Osa 2. Kouluhallituksen alaisen koulutoimen ja vapaan sivistystyön tuloksellisuutta käsittelevä työryhmä, Koulun prosessiarviointijaosto*. Muistio Kouluhallitukselle 15.12.1986. Helsinki: Kouluhallitus.
- Koyama, J. P. 2011. 'Generating, comparing, manipulating, categorizing: reporting, and sometimes fabricating data to comply with No Child Left Behind mandates.' *Journal of Education Policy*, 26: 5.
- Lahelma, E. 1984. *Lähtökohtia koulutuksen vaikuttavuuden seuranta järjestelmän luomiselle Kouluhallituksessa*. Helsinki: Kouluhallitus.
- Lampinen, Osmo 1998. *Suomen koulutusjärjestelmän kehitys*. Helsinki: Gaudeamus.
- Lankinen, T. 2007. *Opetushallituksen asema, rooli ja tehtävät sekä koulutustoimialan ohjaus muuttuvassa toimintaympäristössä*. [The position, role and tasks of the Finnish National Board of Education and the steering of the educational branch in a changing operating environment.] Opetusministeriön työryhmämuistioita ja selvityksiä 2007:46. Helsinki: Opetusministeriö.
- Laukkanen, R. 1998. *Opetustoimen keskushallinnon evaluaatioajattelun kehitys Suomessa 1970-luvulta 1990-luvulle*. Jyväskylä: Koulutuksen tutkimuslaitos
- Law 1866. 1866 Keisarillisen Majesteetin Armollinen Asetus kansakoulutoimen järjestämisestä Suomen Suuriruhtinaanmaassa.
- Law 476/1983. Comprehensive School Act. [Peruskoululaki].
- Law 182/1991. Laki opetushallituksesta. Law on the National Board of Education.
- Law 707/1992. Law on the Amendment of the Comprehensive School Act. 1992. Laki peruskoululain muuttamisesta 707.
- Law 365/17.3.1995. Municipal Act [Kuntalaki].
- Law 628/1998. Perusopetuslaki. [Basic Education Act].
- Leimu, K. 2004. 'Esipuhe' [Foreword]. In *Kansainväliset IEA-tutkimukset Suomi-kuvaa luomassa* [International IEA studies creating the image of Finland], edited by K. Leimu, pp. 3–5. Jyväskylä: Koulutuksen tutkimuslaitos.
- Löfström, E., J. Metsämuuronen, J.E.K. Niemi, K. Salmio and K. Stenvall. 2005. *Koulutuksen paikallinen arviointi vuonna 2004* [The local evaluation of education in 2004]. Helsinki: Opetushallitus.
- Looney, J. 2011. 'Integrating formative and summative assessment: Progress toward a seamless system?' OECD Education Working Papers No. 58. OECD.
- Lyytinen, H.K., H. Jokinen and S. Rask. 1989. *Koulun työkäytäntöjen arviointi. Prosessiarvioinnin näkökulma* [Evaluating school work practices: A view of process evaluation]. Helsinki, Finland: Kouluhallitus.

- McConnell, Allan. 2010. 'Policy success, policy failure and grey areas in-between'. *Journal of Public Policy*, 30(3): pp. 345–62.
- OECD, Organization for Economic Co-operation and Development. 1973. *A framework for educational indicators to guide government decisions*. Paris, France: OECD.
- Olkinuora, E. and L. Perkki 1974. 'Koulutusindikaattorijärjestelmän edellytykset, lähtökohdat ja käytännön konstruointimahdollisuudet' [The theoretical foundations of a system of educational indicators and possibilities for its practical construction]. Reports from the Institute for Educational Research 243/1974. Jyväskylä: University of Jyväskylä.
- Opetus- ja kulttuuriministeriö. 2012. 'Perusopetuksen laatukriteerit: Perusopetuksen, perusopetuksen aamu- ja iltapäivätoiminnan sekä koulun kerhotoiminnan laatukriteerit'. <http://urn.fi/URN:ISBN:978-952-263-170-1/>. [Quality criteria for basic education: Quality criteria for basic education, morning and afternoon activities, and school's club activities]. Opetus- ja kulttuuriministeriön julkaisuja 2012:29. Helsinki, Finland: Opetus- ja kulttuuriministeriö.
- Opetushallitus. 1994. *Peruskoulun opetussuunnitelman perusteet* [Core curriculum for the comprehensive school]. Helsinki, Finland: Opetushallitus.
- Opetushallitus. 1995. *Koulutuksen tuloksellisuuden arviointimalli* [The Framework for Evaluating Educational Outcomes]. Helsinki, Finland: Opetushallitus.
- Opetushallitus. 1998. [The Framework for Evaluating Educational Outcomes]. Helsinki, Finland: Opetushallitus.
- Opetushallitus. 2014. 'Perusopetuksen opetussuunnitelman perusteet' [Core curriculum for basic education]. Helsinki, Finland: Opetushallitus. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/perusopetuksen-opetussuunnitelman-perusteet>.
- Opetushallitus. 2021. 'Oppiaineiden päättöarviointi'. [Final grading of subjects]. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/oppiaineiden-paattoarviointi#6a9ce513/>. Accessed 4 February 2021.
- Opetusministeriö. 1983. 'Peruskoulun kehittäminen. Opetusministeriön vastauskirje 24.3.1983, n:o 8271/27/82 Kouluhallituksen kirjeeseen 30.11.1982'. [Developing the Comprehensive School.] In R. Luhtanen (ed.) *Koululait. Lakikokoelma 13* (pp. 208–215). Helsinki: Valtion painatuskeskus.
- Opetusministeriö. 1990. 'Koulutuksen tuloksellisuuden arvioinnista. Arviointimenettelyä selvittäneen työryhmän muistio'. [On school performance evaluation: Memo of the working group analysing evaluation processes]. Opetusministeriön työryhmien muistioita 1990:23. Helsinki, Finland: Opetusministeriö.
- Opetusministeriö. 1991. 'Koulutuksen ja korkeakouluissa harjoitettavan tutkimuksen kehittämissuunnitelma vuosille 1991–1996'. [Development plan 1991–1996 for education and research in higher education institutions]. Helsinki, Finland: Opetusministeriö.
- Opetusministeriö. (2004). *Koulutus ja tutkimus 2003–2008. Kehittämissuunnitelma* [Education and research 2003–2008. Development plan]. Helsinki, Finland: Opetusministeriö.
- Opetusministeriö. (2007). *Koulutuksen arviointijärjestelmän kehittämistyöryhmän muistio* [Memorandum of the working group for developing the evaluation system]. Opetusministeriön työryhmämuistioita ja selvityksiä 2007:27. Helsinki: Opetusministeriö.
- Opetusministeriö 2009. 'Perusopetuksen laatukriteerit'. [Quality criteria for basic education] Opetusministeriön julkaisuja 2009: 12.

- Opetusministeriö 2010. 'Perusopetuksen laatukriteerit'. [Quality criteria for basic education] Opetusministeriön julkaisuja 2010:6.
- Ozga, J., H. Simola, J. Varjo, C. Segerholm and H. Pitkänen. 2011. 'Central-local relations of governance'. In *Fabricating quality in education: Data and governance in Europe*, edited by J. Ozga, P. Dahler-Larsen and H. Simola, pp. 107–126. London: Routledge.
- Pitkänen, H. 2019. 'Arviointi, tieto ja hallinta. Peruskoulun paikallisen arvioinnin genealogia'. [Evaluation, knowledge and power: Genealogy of local quality evaluation in the field of comprehensive education]. Helsinki Studies in Education 50. Helsinki, Finland: University of Helsinki. <http://urn.fi/URN:ISBN:978-951-51-5238-1/>.
- Rajanen, J. 2000. *Selvitys koulutuksen paikallisen tason arvioinnin tilasta* [The state of local level evaluation]. Helsinki: Opetushallitus.
- Ravitch, Diane. 2010. *Death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.
- Rinne, R., H. Simola, M. Mäkinen-Streng, S. Silmäri-Salo and J. Varjo. 2011. *Arvioinnin arvo: Suomalaisen perusopetuksen laadunarviointi rehtoreiden ja opettajien kokemanä*. [The value of the evaluation. QAE in Finnish comprehensive education as experienced by head teachers and teachers.] Turku: Suomen kasvatustieteellinen seura.
- Saari, A., S. Salmela and J. Vilkkilä. 2017. 'Bildung- ja curriculum-perinteet suomalaisessa opetussuunnitelma-ajattelussa. [Bildung and curriculum traditions in Finnish curriculum thinking]'. In *Opetussuunnitelmatutkimus: keskustelunavauksia suomalaiseen kouluun ja opettajankoulutukseen*, edited by T. Autio, L. Hakala and T. Kujala, pp. 61–82. Tampere: Tampere University Press.
- Sahlberg, P. 2007. 'Education policies for raising student learning: The Finnish approach'. *Journal of Education Policy*, 22(2): pp. 147–171, doi: 10.1080/02680930601158919.
- Sahlberg, P. 2011. 'The fourth way of Finland'. *Journal of Educational Change* 12: pp. 173–185. <https://doi.org/10.1007/s10833-011-9157-y/>.
- Schatz, M. 2016. 'Education as Finland's hottest export?: A multi-faceted case study on Finnish national education export policies'. Thesis. Helsinki: Faculty of Behavioural Sciences. <http://urn.fi/URN:ISBN:978-951-51-2092-2/>.
- Seppänen, P. 2006. *Koulunvalintapolitiikka perusopetuksessa: Suomalaiskaupunkien koulumarkkinat kansainvälisessä valossa*. [School-choice policy in comprehensive schooling: School markets of Finnish cities in the international perspective]. Research in Educational Sciences 26. Turku: Finnish Educational Research Association.
- Seppänen, P., M. Kalalahti, R. Rinne and H. Simola 2015. 'Suomalainen lohkoutuva peruskoulu vanhempien näkemänä' [The Finnish segmenting comprehensive school as seen by parents]. In *Lohkoutuva peruskoulu: Perheiden kouluvalinnat, yhteiskuntaluokat ja koulutuspolitiikka* [Segmenting comprehensive school: School choices, social classes and education politics], edited by Piia Seppänen, Mira Kalalahti, Risto Rinne and Hannu Simola, pp. 511–529. Research in Educational Sciences 68. Jyväskylä: Finnish Educational Research Association.
- Simola, H., J. Kauko, J. Varjo, M. Kalalahti and F. Sahlström. 2017. *Dynamics in education politics: Understanding and explaining the Finnish case*. London: Routledge.

- Simola, H., R. Rinne, J. Varjo and J. Kauko. 2013. 'The paradox of the education race: How to win the ranking game by sailing to headwind.' *Journal of Education Policy*, 28(5): pp. 612–633. doi: 10.1080/02680939.2012.758832.
- Simola, H., R. Rinne, J. Varjo, H. Pitkänen and J. Kauko. 2009. 'Quality assurance and evaluation (QAE) in Finnish compulsory schooling: A national model or just unintended effects of radical decentralisation?'. *Journal of Education Policy*, 24(2): pp. 163–178.
- Sivistysvaliokunnan lausunto (5/2001vp). 'Sivistysvaliokunnan lausunto. Opetusministeriön selvitys koulutuksen arviointitoiminnasta.' [Statement by the Education and Culture Committee. Ministry of Education study on evaluation in schools]. <https://www.eduskunta.fi/FI/vaski/sivut/trip.aspx?triptype=ValtiopaivaAsiakirjat&docid=sivl+5/2001/>. Accessed 4 February.
- Sivistysvaliokunnan mietintö 11/2002vp/28.11.2002. 'Sivistysvaliokunnan mietintö. Hallituksen esitys laeiksi eräiden koulutusta koskevien lakien sekä Opetushallituksesta annetun lain 4§: nmuuttamisesta.' [Report by the Education and Culture Committee. Government bill on changing education legislation and the amendment of the fourth section on the law on the Finnish Agency for Education].
- Sivistysvaliokunnan mietintö 9/2009vp/4.11.2009. 'Sivistysvaliokunnan mietintö. Hallituksen esitys laeiksi eräiden koulutusta koskevien lakien muuttamisesta.' [Report by the Education and Culture Committee. Government bill on changing education legislation].
- Suomen Kuntaliitto 2003a. 'Yleiskirje 14/ 80/2003. Sivistystoimen arvioinnin indikaattorit.' [Circular 14/ 80/2003. Indicators for evaluating education]. Suomen Kuntaliitto.
- Suomen Kuntaliitto 2003b. 'Sivistystoimen arvioinnin indikaattorit—keskustelualoite.' [Education evaluation indicators – discussion initiative]. Suomen Kuntaliitto.
- Suomen Kuntaliitto 2006a. 'Yleiskirje 17/ 80/2006 Suositus arvioinnista kuntien valtuustoille.' Suomen Kuntaliitto. [Circular 17/ 80/2006. A recommendation on evaluation for municipal councils].
- Suomen Kuntaliitto 2006b. 'Suositus arvioinnista kuntien valtuustoille. Perustelut ja täsmennykset.' [A recommendation on evaluation for municipal councils. Justifications and clarifications]. Suomen Kuntaliitto.
- Syrjänen, P. 2013. 'Uutta perinnettä rakentamassa—Suomen peruskoulujen otospohjaisen arviointimallin synty politiikan dynamiikan näkökulmasta.' [Building a new tradition—Birth of the sample-based evaluation model of Finnish comprehensive schools. Political dynamics perspective]. Master's thesis. Institute of Behavioural Sciences, University of Helsinki.
- Takala, T., J. Kauko, R. Rinne and J. Kallio 2018. 'One size for all? Policy advice of the World Bank and the OECD on quality assurance and evaluation of school education in Russia, Brazil and China.' In *Cross-national comparative, evidence-based educational policymaking and reform*, edited by A.W. Wiseman and P.M. Davidson, pp. 301–319. International Perspectives on Education and Society Series. Bingley: Emerald Publishing.
- Temmes, M., P. Ahonen and P. Ojala. 2002. 'Suomen koulutusjärjestelmän hallinnon arviointi.' [An evaluation of the governance of the Finnish education system]. Helsinki, Finland: Opetusministeriö.
- Väättäin, H. (ed.). 2019. 'Kannatteleeko koulutus? Yhteenveto kansallisen koulutuksen arviointitoiminnan tuloksista.' [Does education carry? Summary of the results

- from national evaluation on education]. Helsinki: Finnish Education Evaluation Centre.
- Valtioneuvosto. 2021. *Valtioneuvoston koulutuspoliittinen selonteko 2021. Valtioneuvoston julkaisuja 2021:24*. Helsinki: Valtioneuvosto.
- Valtionalouden tarkastusvirasto 2009. 'Perusopetuksen ohjaus- ja rahoitusjärjestelmä'. [Steering and funding system of basic education]. Valtionalouden tarkastusviraston tuloksellisuustarkastuskertomukset 192/2009. Helsinki: Edita Prima.
- Varjo, J. 2007. 'Kilpailukykyvaltion koululainsäädännön rakentuminen: Suomen eduskunta ja 1990-luvun koulutuspoliittinen käänne'. [Drafting education legislation for the competitive state: The Parliament of Finland and the 1990s change in education policy]. Kasvatustieteen laitoksen tutkimuksia 209. Helsinki, Finland: Helsingin yliopisto, käyttäytymistieteiden laitos.
- Varjo, J., H. Simola and R. Rinne. 2016. 'Arvioida ja hallita: Perään katsomisesta informaatio-ohjaukseen suomalaisessa koulupolitiikassa'. [To evaluate and govern – From "looking after" to management by data in Finnish education politics]. Kasvatusalan tutkimuksia 70. Jyväskylä, Finland: Suomen kasvatustieteellinen seura.
- Vedung E. 1997. *Public policy and program evaluation*. New Brunswick, NJ: Transaction Publishers.
- Verger, A. and L. Parcerisa 2017. 'A difficult relationship: Accountability policies and teachers—International evidence and key premises for future research'. In *International handbook of teacher quality and policy*, edited by M. Akiba and G. LeTendre, pp. 241–254. New York: Routledge.
- Verger, A., L. Parcerisa and C. Fontdevila 2018. 'The growth and spread of large-scale assessments and test-based accountabilities: A political sociology of global education reforms'. *Educational Review*, 71(1): pp. 5–30.
- Vuorela, T. 1980. 'Opetusministeriön historia II. Taantumuksesta uudistuksiin 1825–1868'. [The history of the Ministry of Education II: From recession to reform 1825–1868]. Helsinki: Opetusministeriö.
- Wallenius, T. 2020. 'Schools, performance and publicity: Contrasting the policy on publicising school performance indicators in Finland with the other Nordic countries'. Thesis. Helsinki Studies in Education 92.
- Ydesen, C. 2020. *The OECD's historical rise in education: The formation of a global governing complex*. Cham: Palgrave Macmillan.