

Quality and Evaluation in Finnish Schools

Jaakko Kauko, Janne Varjo, and Hannele Pitkänen

Summary

The quality of education has been a central matter of debate globally in the new millennium. The global trend supports test-based accountability models and increasing national data collection as techniques for supporting and increasing quality in education. In contrast, a central feature of the Finnish education system runs counter to the global trend: it does not have strong top-down quality control mechanisms. The evolution of the Finnish 'model' dates back centuries. The foundations of the Finnish quality system can be traced to participation in international comparative learning studies developing national capacity, the inspection of folk education supporting the tradition of nationally coordinated external evaluation, and the local supervision of folk schools through school boards emphasizing local provision and the quality control of evaluation. These developments culminated during the 1990s with the radical deregulation and decentralization of education governance. The current model is partly unarticulated. However, it is clearly distinguishable: In comprehensive schools (primary and lower secondary) quality is entrusted to education providers and schools. They are expected to conduct self-evaluation regularly. There are no national standardized tests and sample-based testing for development purposes form the core of evaluation data. Only the main evaluation results are published making school rankings impossible. Yet there is a large variation in how the quality of education is approached and evaluated in Finland's more than 300 municipalities. Significantly, the central government has no direct means to control the quality of local education. Its impact is indirect through aims to foster and promote the quality evaluation culture in schools and municipalities. Furthermore, international cooperation and participation in international large-scale assessments have been unable to politicize the national education development discourse. This somewhat uncoordinated yet economical and teacher-friendly quality system raises interesting questions for further research: is this only a Finnish

peculiarity developed in a specific historical context, or does it make possible critical theoretical and societal conclusions that question the dominant global test-based quality trends? The buffering of international accountability-based testing and swimming against the global quality evaluation flow is built on 1) the compartmentalization of international tests, 2) the fact that national coordination began to see a deregulated system as a necessity and virtue, and was long fragmented in different evaluation functions, and 3) the important role the local level has played historically in upholding and evaluating the quality of education.

Keywords

Quality in education, Finland, schools, PISA, evaluation

Quality in education: a global trend

Quality has for decades been of growing interest to international organizations and national policies (Leeuw 2002; Power 1994; Smith 1990), and is now firmly on the global education agenda.

‘Quality education’ [<https://www.un.org/sustainabledevelopment/education/>] is one of the United Nations’ Sustainable Development Goals for 2030. Quality is often thought of as walking hand in hand with *evaluation*, which is another current global trend considered a must for any public organization (Dahler-Larsen 2011). The European Commission (2015) sees the constant monitoring and improvement of the quality of education through effective quality assurance systems as crucial at all levels of education, and necessary for Europe’s economic success and growth. The emphasis on evaluating quality is so strong that recent studies in education policy argue that in the relationship between quality and evaluation the latter has started to define the former (Kauko et al. 2018a). The means to secure quality have thus overshadowed the important and difficult value discussions of what quality in education actually constitutes.

The evaluation of quality constitutes not only a technical measure but a political means. There is a lack of balance in political debate, because it is easier to expend energy in focusing on the concrete conduct of evaluation than on quality, which is an abstract and value-laden concept. Defining quality requires a laborious political debate about values. There is also a strong discourse on how data can be used in any context (Barber 2014), denoting that data is pure of values, and its use needs no theoretical understanding. A good example of the opaque nature of values in evaluating education is the OECD’s PISA programme, where data production is highly complex with different types of decision making and technical decisions, but the process remains concealed from public and political debate and is thus unpoliticized (Carvalho 2013). While quality is a highly normative concept, it is often perceived as a neutral governing tool, resulting in a self-enforcing dynamic: measured quality problems are seen as best resolved with greater quality control (Kauko et al. 2018a; 2018c).

The political nature of quality evaluation in education can arguably be simplified into two main streams of thinking. The first, global mainstream relies on the idea that tools such as national large-scale assessments of education achievement are stringent in providing test-based accountability through which education policy goals are attained. This approach usually entails elements of accountability, standards, and decentralization. It is globally mainstream and has reached countries that until recently did not subscribe to it (Verger et al. 2018). Sahlberg (2016) describes this as a ‘Global Educational Reform Movement’, which supports ‘competition, choice, prescribed curricula, standardized testing, and privatization’ and sees the effect of such an approach on quality as negative. The second way to understand quality in education is to see it as swimming upstream against the global flow. It relies on the professional capacity of teachers and principals, favours non-sanctioning development-led use of data, and is wary of ranking lists and national large-scale assessments (Simola et al. 2013). It also relies on public rather than private education, and seeks to support collaborative practices and trust (Sahlberg 2016). The global mainstream, which relies on numbers and data, is used in various countries, such as the Anglo-American countries (most of the UK, Australia, and the USA), large and heterogeneous countries (Brazil, China, and Russia), and some smaller countries (Singapore). The global upstream approach is used in small countries or states with a high level of education (Finland, Norway, and Scotland) and some federal states (e.g. Alberta in Canada), for example (Sahlberg 2006; Green 1997; 1999).

Describing quality of education in the general terms of these two streams does not enhance the understanding of why and how the politics of quality in education are constructed so differently, and why the global mainstream and upstream are constructed. The debate concerning whether global influences or local histories and circumstances, or perhaps a combination, are more dominant in formulating policy is a key question in comparative research (e.g. Marginson and Rhoades 2002; Anderson-Levitt 2012). In institutional research many decades of studies have shown that historical path dependencies (Pierson 2000) and norms that define institutions (March and Olsen 1989) create

a situation where understanding the local context becomes essential to understanding change. The construction of this Finnish version, swimming upstream against the global model of quality and evaluation in education, can be understood by examining the historical paths and lines of quality evaluation at the international, national, and local levels, because they have all been effective in constructing the present rationale and model of quality evaluation in Finnish comprehensive education.

The Finnish comprehensive school is a nine-year compulsory school for the whole age cohort, and it consists of six lower classes for ages 7–13 and three upper classes for ages 13–16. The schools are free, including all related costs (meals, travels and study materials, for example).

Comprehensive schools are typically run by the over 300 Finnish municipalities. There are also some independent education providers, which are non-profit, free, and follow the same legislation and requirements as all other schools. The historical roots of what is now understood as the ‘Finnish’ education system were formed under the Swedish Crown (until 1809) and as an autonomous Russian archduchy (until 1917, when Finland declared independence). Religious and estate-based education became institutionalized and secularized during Russian rule (Joutsivuo 2010), and education provision was eventually transferred from the church to the state and municipalities in the wake of industrialization (Leino-Kaukiainen and Heikkinen 2011). Between 1860 and 1870 the school system was reorganized into elementary schools (*kansakoulu*, literally ‘folk schools’) and grammar schools (*oppikoulu*, literally ‘learning schools’). However, it was only in the 1940s that elementary schools began to reach the country’s remoter areas (Simola 2002). In the 1960s the increasing change in economic structure from an agrarian to a more industrial and service-based economy demanded a more educated workforce (Kettunen et al. 2012), which was explicitly stated in the government bill (*Hallituksen esitys 44/1967*) to establish comprehensive schooling (see Kauko 2019). The focus of this chapter is on the quality evaluation of the comprehensive school.

Most of the roots of the current quality evaluation system in comprehensive education can be traced to the emergence of the formal comprehensive education system, but some are found even in medieval times. According to Varjo, Simola, and Rinne (2016; see also Jakku-Sihvonen 2001) the foundations of the Finnish quality system can be traced to *participation in comparative learning studies, the inspection of folk education, and the local supervision of folk schools by school boards*. It is important to note that these three early forms of quality assurance and evaluation have remained, despite the fact that the 1960s and 1970s brought remarkable changes to the architecture, functions, and procedures of the central Finnish administration. For example, the state's sphere of operations expanded to the fields of education, healthcare, and social insurance. The rapid growth in the number of civil servants and administrative bodies was an indication of a completely new phase of national development (Varjo 2007; Simola, Kauko, Varjo, Kalalahti, and Sahlström 2017). The quality of comprehensive schools is currently evaluated in three categories: international testing; national external evaluation; and local self-evaluation. These three historical paths affect the current system and the construction of this Finnish version of swimming upstream against the global flow in the quality and evaluation of education.

Learning evaluation methods from international comparisons

The 1940s saw the increasing importance of educational assessment and the developing of measurements for educational achievement within the fields of psychological and educational research (Konttinen 1995). Since the 1950s Finland's continuous participation in comparative learning studies has laid the foundation for a national evaluation system of learning outcomes (Kangasniemi 2004) and participation in the international development of education indicators (Leimu 2004). As more reliable measurements have developed, the understanding of educational achievements' multidimensionality and their objective measuring has gradually grown. Advances in methodology have paved the way for international research projects. After the Second World War studies in comparative education started to evolve methodologically more systematically and

rigorously: data was collected domestically from each nation but within a common international research design with the same methods and measurements (Leimu 2004). The flows of data and ideas between universities and nations therefore became more constant (Lawn and Grek 2012).

The most prominent of the early comparative projects started in 1958 under the auspices of UNESCO. *The International Association for the Evaluation of Educational Achievement* (IEA) was built around problematizations associated with the evaluation of school effectiveness and student learning. The core idea was that effective evaluation required the examination of both inputs to education and its outcomes, such as knowledge, attitudes, and participation. Educational achievements, measured by a common research design, were articulated as the organization's main objective (Leimu 2004). According to Husén (1967) 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 across a wide range of systems, the variability would be sufficient to reveal important relationships that would otherwise fail to have been revealed within a single education system. The first IEA study, conducted in 1960 in twelve countries, assessed thirteen-year-old students' attainment in mathematics, reading comprehension, geography, science, and non-verbal ability. The aim of this research, known as the Pilot Twelve-Country Study, was to investigate the feasibility of more extensive assessments of educational achievement. The study demonstrated the overall feasibility of conducting large-scale crossnational surveys – and, simultaneously, paved the way for numerous comparative studies within the OECD and the EU, for example.

Finland was one of the twelve founding members of the IEA. According to Laukkanen (1998), participation in international cooperation was a necessity to develop national expertise in the evaluation of education. For instance, until the 1970s Professor Matti Takala, who contributed

substantially to the establishment of the organization in the 1950s, was also a central figure in the developments in comparative education and psychometric measurements in Finland.

The OECD launched its Indicators of the 'Performance of Educational Systems project' in 1970. The first set of comparative indicators – A Framework for Educational Indicators to Guide Government Decisions (OECD 1973) – was published in 1973. The framework suggested forty-six indicators which were constructed by combining, for example, the statistics of student ratios and flows, education resources, and costs with background variables. The Finnish Ministry of Education subsequently funded a 'project on education indicators' from 1972 to 1974 at the University of Jyväskylä to 'study the possibilities of constructing a system of education indicators to help quantitative education planning' (Olkinuora and Perkki 1974, XI). Nevertheless, indicators were omitted from the national education policy agenda by the mid-1970s. None of the forty-six OECD indicators was implemented as such in Finland; they were more of an underlying factor behind the national enterprise of indicator elaboration, which was soon to be faded out when administrative interest dwindled (Kauko and Varjo 2008).

Education indicators and quantitative comparative studies again became current in the OECD's sphere of interest. Issues of accountability in education had also gained importance (Kauko and Varjo 2008). The 1990s saw an active quest for international cooperation in the education sector. A Ministry of Education (1990, 26) working group memo highlights the OECD indicator development as an interesting way of collaborating internationally. However, it identifies theoretical limitations in it that have become evident in previous decades. Finnish education's comparability and its international presentation as an interesting environment are emphasized in the 1993 revision of the 1991 education development plan, which was written during an economic crisis exacerbated by the fall of the Soviet Union (Opetusministeriö 1991; 1993). The same thoughts are replicated in the 1997 education evaluation development strategy (Opetusministeriö 1997). The ministry was also active in publishing analysis and translations of OECD-related topics (Laukkanen 1994; 1997) and

seeing international evaluations as an essential part of quality in education (Hämäläinen and Jakku-Sihvonen 2000). This view is linked to a strengthening agenda which sees education as an important part of staying competitive internationally (Opetusministeriö 2004, 12).

Finland has performed well in international comparisons in the new millennium. In the OECD's PISA Finland ranked first (2001) and second (2009) in reading (2001), first (2003) and twelfth (2012) in mathematics, and first (2006) and fifth (2015) in science (OECD 2002; 2004; 2006; 2010; 2014; 2016). In the IEA Trends in International Mathematics and Science Study (TIMSS) Finnish eighth-graders ranked tenth in science (1999) and fourteenth (1999) and eighth (2011) in mathematics. Fourth-graders ranked eighth (2011) and seventeenth (2015) in mathematics and seventh (2015) in science (Mullis et al. 2000; 2012; 2016; Martin et al. 2000; 2016). Fourth-graders in Finland ranked third (2011) and fifth (2016) in the Progress in International Reading Literacy Study (PIRLS) (Mullis et al. 2012; 2017). These high results were used in domestic policies (Rautalin 2013) and in exporting Finnish education (Schatz et al. 2015). Nationally, the effect of the success in the PISA tests is disputed. Sahlberg (2011) argues that PISA success caused a lack of innovation in primary and secondary education, while Seppänen et al. (2019) suggest that comprehensive education policy during the new millennium has been noticeably active.

In the evaluation plan for 2012–2015 international comparisons are treated separately, and they are used as sources for comparison with other countries rather than of evidence for national development, or as supplementary information (Opetus- ja kulttuuriministeriö 2012, 5–8). This follows the general logic with which international comparisons have tended to be treated. A report focusing especially on a dip in results does not follow suit in this sense: in the foreword the Minister of Education states that in drawing on international and national observations the government has initiated many actions like the Future Comprehensive School project (as part of curriculum reform) and digitalization (Väljärvi and Kupari 2015, 7). However, following the

argument developed by Seppänen et al. (2019), these policy goals are more in line with the government programme predating the PISA dip (Government 2011) than anything that might be deduced from the results of the international large-scale assessment.

Finland has been interested in international education comparisons for as long as they have existed. Although they may be assumed to have had an effect on developing evaluation capacity, they have had no effect on the Finnish education system's basic philosophy. The reasons for this lie in the longer path dependencies in which the national and local levels built on the previous work.

The development of nationally coordinated external evaluation

Confirmation classes and parish catechetical meetings were the first institutionalized forms of folk education. Their function was to evaluate and control reading skills and command of the basics of Christian doctrine among the peasantry (Kähkönen 1982). Since medieval times bishops have regularly inspected their dioceses' parishes. As parish priests started to play a more active role in folk education, the evaluation of reading and the basics of Christian doctrine were included in dioceses' inspection schedules during the sixteenth century. The 1571 Church Order explicitly obliged priests to inspect the folk schools within their jurisdiction (Hanska and Lahtinen 2010).

In 1841 the Tsar decreed that cathedral chapters must prepare a report for the Finnish Senate on the condition of the folk schools in their jurisdiction. Vuorela (1980) maintains that this was the first time that such reports equipped the central administration to evaluate and control folk schools continually. When religious and secular authority were separated in the 1860s, the evaluation of folk education became the duty of the latter. The National Board of General Education was founded in 1869. School inspections were set as one of its main tasks from the beginning, and the Reverend Uno Cygnaeus was nominated as the 'chief inspector of all folk schools' (Harju 1988). The chief inspector was assisted by a body of inspectors which were charged to 'look after teaching, order, and discipline' (Act 12/1866) at the folk schools within their districts and report to the National

Board of General Education. As the number of folk schools increased, new inspection districts were established. In 1921 there were already twenty-one districts in Finland (Nikander 2011).

Inspectors' tasks were very broad. They included legality control, economic management, pedagogy, and learning results, governing every aspect of schoolkeeping (Lyytinen and Rautama 2002). The routines of school inspections remained the same until the 1970s; nevertheless, the legislation on which they were based became increasingly detailed as the decades passed. The implementation of the comprehensive school reform (1972–1977) altered the administrative position of the inspectorate only slightly. Inspectors were now posted to the provincial governments. Simultaneously, their duties were redefined as *guidance, supervision, and data collection*.

Finnish basic schooling was subjected to two decades of extremely bureaucratic, norm-driven, and top-down governance. The implementation of comprehensive schooling conformed with the planning optimism of the 1960s and 1970s, when it was generally accepted that the comprehensive system should be implemented by strong top-down government. Normative and detailed legislation was deemed necessary to guarantee equality in different parts of the country. Legislation defined all the education services to which a citizen was entitled, as well as how these should be provided. For example, subjects, curricula, teachers' qualification requirements, school books, the rights and duties of students and staff, and the architecture and procedures of local school administration were all prescribed in laws, decrees, and normative decisions at the central administrative level.

Centralized legislative and statutory steering was finalized with a procedure requiring local education authorities to subordinate in advance their decisions to regional state administrative agencies or the National Board of Education (Varjo, 2007; Simola et al. 2017).

The steering mindset changed radically during the 1980s and 1990s. These changes belonged to a larger international decentralization and deregulation wave influenced by New Public Management ideas. In the Finnish case the economic crisis at the beginning of the 1990s provided another

argument for moving decisions on budget cuts to the local level. Funding and other legislative reforms had the effect of significantly increasing municipalities' autonomy (Law 707/1992; 365/1995). Following an amendment to a decree on provincial governments, the legal basis for school inspections was dismantled (Decree 1010/1990), and they ceased within a few years in every province. This dismantling of central control set in motion a trajectory which culminated only in 2015, when all external evaluation activities were placed under one national organization.

In policy texts the political move towards evaluation took effect throughout the administrative levels. Nationally, the idea of developing a new evaluation culture was visible in various elements. The government started a four-year cycle of five-year education development plans and emphasized the need for continuous evaluation (Valtioneuvosto 1990). The first plan explicitly stated that 'Education steering will shift from norm-based regulation into setting national educational goals and their evaluation' (Opetusministeriö 1990). The Ministry of Education (1990) working group connected evaluation to democracy and self-development, and explicitly argued against using evaluation primarily as a means of control. The National Board of Education's role was also redefined, becoming more development- and evaluation-oriented, and it was merged with its vocational education counterpart. During the 1990s the agency published performance evaluation models (Opetushallitus 1995; 1998), which were intended to support the education providers in developing an evaluation and quality culture.

The national model was and still is based on a set of principles, which various documents and interviews with officials demonstrate. For example, although the industrial lobby (Teollisuuden koulutusvaliokunta 1991) called for national standardized achievement tests for whole age cohorts, tests are sample-based, and there are no school ranking lists. Many official documents declared that the use of evaluation data was above all for officials and school development, not for control (Opetusministeriö 1997, 3). This approach is aligned with inspection becoming obsolete (Simola et al. 2009).

Although initial attempts were to develop a quality and evaluation culture, evaluation became less rooted the more closely the perspective was locally focused. A ministry-commissioned national evaluation of the steering system by three academics expressed suspicions ‘that our country would have succeeded in shifting from a centralised education policy into an “evaluation-centred education policy”’, because budgetary and institutional structures were increasingly defining features (Temmes et al. 2002, 91). They also drew attention to the fact that national-level evaluation was dispersed to different units, and were critical of the National Board of Education’s evaluation role. The evaluation report suggested that the Standing Committee for Education should heed a previous rapporteur and establish a separate council for this purpose (Temmes et al. 2002, 94, 132). These suggestions became reality in 2003.

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)* for higher education was formed in 1996. The FEEC was tasked to draft an evaluation plan and provide information for education policy and setting new goals (Ministry of Education 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 (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 they were also responsible for legality control (Lankinen 2007, 189). 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. A working group memo described how the foundation of the FINEEC would lend more credibility to external evaluation (Opetus- ja kulttuuriministeriö 2012). The foundation of the FINEEC was the culmination of a development that had started in the 1990s as the post-deregulation reorganization of evaluation functions. However, the basic principles of evaluation remained the same: the main task was to assist in the development of local quality processes and provide external evaluation data, not to use data to control the education providers.

Formation of local plurality in evaluation

Local quality evaluation constitutes a major element of the Finnish quality evaluation system of education, alongside evaluation at the national level and participation in international large-scale assessments. The Finnish evaluation system relies heavily on the principle that each of the education providers, for the most part the municipalities, evaluates their education provision autonomously. In addition to the education providers schools are also expected to evaluate their operations. This emphasis on evaluation at the local level, including both municipal and school-level self-evaluations, with the absence of school inspection and national testing, constitutes where Finland differs from many other education systems.

The historical roots of local-level evaluation date back to the simultaneous development of the folk education system and municipal administration. The separation of municipalities from congregations was sealed judicially in 1865; the Folk School Act came into effect in 1866. The Act on Municipal Administration imposed a range of tasks on newly established municipalities. In this new legislation 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 schoolkeeping' (Decree 12/1866). The Folk School Act saw towns and the countryside as separate entities. Typically, towns had a municipal board for all folk schools,

whereas sparsely populated rural municipalities had a single board for each school. Their societal relevance was considerable. According to Halila (1949) school board membership was one of the novel municipal positions of trust crucial to the newly established Finnish local administration in the late nineteenth century. The comprehensive school reform of the 1960s and 1970s did not substantially change the position of school boards. The Comprehensive School Act (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 supervision of student welfare (nursing, catering, and transportation, for example) were now included in school boards' duties.

As noted above, decentralization and deregulation substantially altered municipalities' legal position. For example, after 1992 municipalities became free to decide whether to appoint school boards at all. In state and private schools, however, boards remained obligatory by law. Alongside the decentralization of power to and deregulation of municipalities, the New Public Management trends shifted the focus to the outcomes of education from its inputs. The emerging rationale suggested that aims and resources would be prescribed by the state, and the responsibility for achieving aims within the frames of given resources would lie at grassroots level, in municipalities and schools (see e.g. Hirvi 1992, 11–12). Accordingly, schools and municipalities were expected to find a locally specified way of attaining these nationally determined general goals. The new governing rationale was that education providers were not only expected to find a way to attain goals, but to demonstrate and prove their results. The emerging idea was that results must be indicated more clearly than previously (Halinen et al. 1992, 44). Performance evaluation was introduced as a way of indicating results (Halinen et al. 1992, 44; Hirvi 1993, 170). The move towards result-based management was justified as a way to increase local and school-level autonomy. It was also taken as a strong sign of trust in teachers and their professionalism, schools,

and education providers, because no inspection or external sanction-based system was developed (e.g. Hirvi 1992, 12; Laukkanen 1993, 65; Lyytinen 1993, 72; Rask 1993, 102).

Current local quality evaluation in the Finnish context includes both municipal and school-level self-evaluation. Legislation prescribes self-evaluation as an obligation for education providers, but schools are also expected to evaluate themselves. The function and role of self-evaluation, and their historical development have been viewed a little differently for schools and education providers.

The idea of school self-evaluation had already arisen in Finnish policy documents in the mid-1980s (e.g. Kouluhallitus 1986; Lyytinen, Jokinen and Rask 1989) but developed further in the early 1990s, when the idea of creating a local quality evaluation culture in schools and municipalities emerged. School self-evaluation was seen as essential for the quality and outcomes of the entire education system and its development. The school level came to be seen as the most important part of the system, because education occurred in schools, and the outcomes of the education system therefore depended on school-level outcomes (Opetusministeriö 1990, 60). The Ministry of Education also reasoned that schools should evaluate themselves, and that this should be highlighted (Opetusministeriö 1990, 60). In 1994 the national core curriculum (Opetushallitus 1994) raised the idea of school self-evaluation to the curriculum level. It stated that every school should continuously and systematically evaluate their own activities and work. The education providers were charged with supporting schools in their self-evaluation. School-level self-evaluation has been an expected feature of the core curriculum ever since. A change has been seen in its closer integration with a local evaluation policy which entails not only education but other public services provided by municipalities.

In 1998, a few years after schools were obliged to undertake continuous self-evaluation, the Basic Education Act required education providers to evaluate the education they provided. Thus, it was not schools but providers who were given the main responsibility for local evaluation. Nevertheless, schools should also evaluate their operations, but increasingly as part of local evaluation and in

relation to municipal evaluation policies. The 1998 Basic Education Act stated that the aim of self-evaluation was to contribute to the development of education and improve learning conditions. As in the national external evaluation presented in the previous subchapter, quality evaluation was not introduced primarily as a control measure but to develop and improve education. Special emphasis is placed on the function of evaluation not as a means of control or sanction, but of developing education continuously in line with the targets set by the curriculum. For example, the current curriculum 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’ (Opetushallitus 2014).

Despite the emphasis on local quality evaluation and the legal obligation to provide it, there is a consensus that no compulsory or predetermined model for local quality evaluation should be implemented (Opetusministeriö 2010; Pitkänen 2019). Since the autonomy of municipalities as education providers was enhanced by the deregulation of the 1990s, they have been allowed to set their own local education policies. In line with this trend as education providers, municipalities determine evaluation issues autonomously. Each of the education providers is free to choose its methods, the focus of evaluations, and how they publish results.

Even if there is no compulsory or predetermined model for quality evaluation in schools and municipalities, since the 1990s much effort has been expended on supporting the creation and mobilization of the quality evaluation culture in Finnish schools. For example, various projects for school self-evaluation have been organized, and national agencies have supported schools by producing inspirational material and models for self-evaluation (e.g. Opetushallitus 1995; 1998; Oppi- ja Laatu 1996; 1998; 2003; Räisänen and Rönholm 2006), and the Ministry’s Quality Criterion for basic education (Opetus- ja kulttuuriministeriö 2012). There appears to be a kind of hegemonical thinking (e.g. Räisänen and Rönholm 2006) within these recommendations

concerning the quality of education to be captured and developed based on universal evaluation models like the EFQM[<http://www.efqm.org/index.php/knowledge-base/what-is-the-efqm-excellence-model-2/>]. It is argued that these models are applicable to any organization, whether public or private, in any service field. Such models seem to be receiving increasing favour and attention from education authorities, education providers, and schools (see e.g. Karvonen 2010). The notion of an evaluation culture in schools embeds the idea that self-evaluation constitutes an inevitable, continuous, and normal part of schools' work. A well-performing school, it is argued, is one which constantly practises self-evaluation. The task of the school in its overall functioning is therefore not only to assess pupils' attainments but to succeed in implementing the curriculum and the general management of everyday school work.

Since the 1990s the evaluation culture in schools and municipalities has been promoted as essential for school development, quality, and outcomes. There has been great emphasis on creating this culture. However, it seems that quality evaluation has still not reached the level intended by education policy. A recent report of the Finnish Education Evaluation Centre describes the situation:

The outcomes showed many providers lacking a functioning self-evaluation system or a systematic assessment culture as part of their quality assurance. The evaluations of some providers, therefore, do not meet the criteria and expectations based on trust, which the 1998 reform of the educational administration would require. At the time, external control and supervision were reduced, for example through eliminating the educational inspection system and moving to a self-regulated quality management by the providers: a system based on self-evaluation (Harjunen et al. 2017).

Despite persistent attempts to create an evaluation culture in schools and municipalities, Finnish education providers and schools still seem only to be moving towards it. The Finnish Education Evaluation Centre has recommended, a national strategy for quality assurance and evaluation

should be drawn up. This would outline quality assurance and continuous steering from early education to upper secondary education (Väättäinen (ed.) 2019, 26). If such a strategy were implemented nationally, the period of autonomous local evaluation might see a turn towards a more controlled evaluation system in the shape of a national model for local evaluation. However, it is unclear what political support there would be for these ideas given the different historical path dependencies.

Discussion

Simola et al. (2017) suggest that Finland has been able to buffer international influences to ensure that quality and evaluation policies can develop against the global mainstream. Indeed, Finnish peculiarities in quality evaluation can be understood as historical trajectories embedded in Finnish society and history. The *local supervision of folk schools through school boards, the inspection of folk education, and participation in comparative learning studies* have evolved into *local self-evaluation, national external evaluation, and international testing* (Varjo et al. 2016). In examining this development from the perspective of historical path dependencies, certain factors contribute to a refined understanding of this buffering effect.

First, international cooperation and participation have failed to politicize the national education development discourse. International comparison has been seen as important, but it has been compartmentalized as a tool mainly used to check how Finland was performing in relation to other countries as an alternative to using OECD or IEA national development data. Policy texts on international comparison also strike a tone of competition with other nations. Such a discourse is completely absent when the situation within the country is discussed. Describing Finnish education providers or schools in competitive terms is unheard of in both past and current political education debate.

Second, the decentralization of governance has forced the national level to accept that there are few ways the central government can use to control quality at local level. Since the Folk School Act of 1866 the local provision of basic education has been assumed. Green et al. (1999) maintain that after a period of centralization from the 1930s ‘traditional Nordic localism’ re-emerged during the 1980s. The idea of evaluation as a way of developing and improving education is now cemented in the Finnish quality evaluation discourse. The idea of evaluation as development has also been strong. It is also convenient given that the national level has few control mechanisms. The formation of the Finnish Education Evaluation Centre has the potential to change this trend in the longer term.

Third, there are more than 300 basic education providers and over 2300 schools in Finland. Each education provider has autonomy in deciding on evaluation. Each can decide the focus and methods of its evaluation. Despite the strong emphasis on a strong local quality evaluation culture and attempts to mobilize it, there appears to be great variance between education providers’ and schools’ implementation. There may therefore be hundreds of local evaluation systems and cultures in Finland, but it seems that the customer-oriented quality approach and the use of organizational quality evaluation models, which are arguably universally applicable, have been strengthening. This prompts a question about the culture of quality evaluation – if it merely entails the adoption of the quality models thought to be universally applicable, independent of the school context, will it lead to confusion concerning what constitutes quality in basic education? This discussion should begin with the political debate about what quality means (see Pitkänen 2019).

Finally, it has been emphasized that evaluations conducted at the local, regional, and national levels should support each other by disseminating results from the municipalities to the national level, and from the national level to the municipalities. The idea has been that national and municipal evaluation and school-level self-evaluation should be complementary (Laukkanen 1995a, 70; Salmio and Vainio 1995, 8), and data should circulate between these levels. However, great

emphasis has been placed on local evaluation. It has come to be understood as the cornerstone of the Finnish evaluation system (Rask 1993a, 144; Halinen 1995, 99).

Ultimately, the buffering of international accountability-based testing and swimming against the global quality evaluation flow is built on 1) the compartmentalization of international tests, 2) the fact that national coordination began to see a deregulated system as a necessity and virtue, and was long fragmented in different evaluation functions, and 3) the important role the local level has played historically in upholding and evaluating the quality of education.

Acknowledgements

References

- Anderson-Levitt, K. (2012). "Complicating the concept of culture." *Comparative Education* 48(4): 441–454. <https://doi.org/10.1080/03050068.2011.634285>
- Barber, M. (2014). "Data Work." In T. Fenwick, E. Mangez, and J. Ozga (eds) *Governing Knowledge: Comparison, Knowledge-Based Technologies and Expertise in the Regulation of Education. World Yearbook of Education 2014*, 75–86. (London: Routledge).
- Decree (1010/1990). *Asetus lääninhallitusasetuksen muuttamisesta*. [Decree on changing the act on provincial governments.] <https://www.finlex.fi/fi/laki/alkup/1990/19901010>
- Decree (12/1866). *Kansakouluasetus*. [Folk School Act].
- European Commission (2015). *Assuring Quality in Education: Policies and Approaches to School Evaluation in Europe. Eurydice Report*. (Luxembourg: Publications Office of the European Union). <https://publications.europa.eu/en/publication-detail/-/publication/4a2443a7-7bac-11e5-9fae-01aa75ed71a1/language-en>
- Government (2011). *Programme of Prime Minister Jyrki Katainen's Government. 22 June 2011*.
- Green, A. (1997). *Education, Globalization and the Nation State*. (London: MacMillan Press).

Green, A. (1999). "Education and globalization in Europe and East Asia: Convergent and divergent trends." *Journal of Education Policy* 14(1): 55–71. <https://doi.org/10.1080/026809399286495>

Green, A., Wolf, A., and Leney, T. (1999). *Convergence and Divergence in European Education and Training Systems*. (London: Institute of Education, University of London).

Halila, A. (1949). *Suomen kansakoululaitoksen historia. Osa 2: Kansakouluasetuksesta piirijakoon* [The History of the Finnish Folk Education System. Part 2: From Folk Education Act to the school district division]. *Suomalaista tiedettä* No 3,2. (Helsinki: Suomen tiedeakatemia).

Halinen, I. (1995). "Itsearviointia Eläintarhan ala-asteella" [Self-evaluation at the Eläintarha primary school]. In B. Kilpinen, K. Salmio, L. Vainio, and A. Vanne (eds) *Itsearviointin teoriaa ja käytäntöä [Theory and practice of self-evaluation]*, 99–110. (Helsinki: Opetushallitus).

Halinen, I., Hämäläinen, K., and Laukkanen, R. (1992). "Koulun tuloksellisuuden kansainvälisiä virtauksia." [International trends in school performance.] In R. Laukkanen, K. Salmio, and R. Svedlin (eds) *Koulun itsearviointi [School self-evaluation]*, 29–47. (Helsinki: Opetushallitus).

Hallituksen esitys (44/1967). *Hallituksen esitys eduskunnalle laiksi koulujärjestelmän perusteista*. [The Government Bill for the legislative basis of the school system].

<https://www.edilex.fi/he/viitekortti/19670044>

Hämäläinen, K. and Jakku-Sihvonen, R. (2000). *Laatua koulutuksen laatupolitiikkaan [Quality for the Education Quality Policy]*. (Helsinki: Opetushallitus).

https://oph.fi/download/48949_laatu.fi.pdf

Hanska, J. and Lahtinen, A. (2010). "Keskiajalta 1500-luvun lopulle [From the Middle Ages to the end of the sixteenth century]." In J. Hanska and K. Vainio-Korhonen (eds) *Huoneentaulun maailma. Kasvatus ja koulutus Suomessa keskiajalta 1860-luvulle*, 17–111. (Helsinki: Suomalaisen Kirjallisuuden Seura).

Harjunen, E., Hietala, R., Lepola, L., Räisänen, A., and Korpi, A. (2017). *Arvioinnilla luottamusta. Perusopetuksen ja lukiokoulutuksen järjestäjien laadunhallinta- ja itsearviointikäytänteet. [Trust through Evaluation: Quality Management and the Self-evaluation Practices of Comprehensive and General Upper Secondary Education Providers]* (Helsinki: Kansallinen koulutuksen arviointikeskus). https://karvi.fi/app/uploads/2017/01/KARVI_0417.pdf

Hirvi, V. (1992). ”Koulun itsearviointia kehittämään [Onwards to developing school self-evaluation]”. In R. Laukkanen, K. Salmio, and R. Svedlin (eds) *Koulun itsearviointi [School self-evaluation]*, 11–13. (Helsinki: Opetushallitus).

Hirvi, V. (1993). ”Näköala koulun kehittämiseen [A perspective for school development]”. In K. Hämäläinen, R. Laukkanen, and A. Mikkola (eds) *Koulun tuloksellisuuden arviointi [Evaluation of School Performance]*, 169–173. (Helsinki: Opetushallitus ja Vantaan täydennyskoulutuslaitos).

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 R. Jakku-Sihvonen and S. Heinonen (eds) *Johdatus koulutuksen uudistuvaan arviointikulttuuriin*, 9–21. (Helsinki: Opetushallitus).

Joutsivuo, Timo (2010). ”Papeiksi ja virkamiehiksi [Becoming Priests and Civil Servants.]” In *Huoneentaulun maailma*. In Jussi Hanska and Kirsi Vainio-Korhonen, *Kasvatus ja koulutus Suomessa keskiajalta 1860-luvulle*, 112–183. (Helsinki: SKS).

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).

Kangasniemi, E. (2004). ”Matematiikalla se alkoi [It all started with mathematics].” In K. Leimu (ed.) *Kansainväliset IEA-tutkimukset Suomi-kuvaa luomassa*, 99–121. (Jyväskylä: Koulutuksen tutkimuslaitos).

Karvonen, J. (2010). *Laadunhallinta opetustoimessa ja varhaiskasvatuksessa* [Quality Management in Teaching and Early Childhood Education]. (Helsinki: Suomen Kuntaliitto).

Kauko, J. (2019, forthcoming). ”The Finnish Comprehensive School: Conflicts, Compromises, and Institutional Robustness.” In Paul t’ Hart and Mallory E. Compton (eds) *Great Policy Successes: How Governments Get it Right in a Big Way at Least Some of the Time*. (Oxford: Oxford University Press).

Kauko, J., Rinne, R., and Takala, T. (2018). ”Conclusion.” In J. Kauko, R. Rinne, and T. Takala (eds) *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia*, 180–190. (London: Routledge).

Kettunen, P., Jalava, M., Simola, H., and Varjo, J. (2012). ”Tasa-arvon ihanteesta erinomaisuuden eetokseen” [From the ideal of equality to an ethos of excellence]. In P. Kettunen and H. Simola (eds) *Tiedon ja osaamisen Suomi Kasvatus ja koulutus Suomessa 1960-luvulta 2000-luvulle*, 25–62. (Helsinki: SKS).

Konttinen, R. 1995. ”Arvostelusta näyttöön – koulutuksen arvioinnin kehityspiirteitä Suomessa” [From assessment to evidence – The development of the evaluation of education in Finland]. In S. Takala (ed.) *Arviointi ja koulutuksen laadun kehittäminen*, 9–22. (Jyväskylä: Kasvatustieteiden tutkimuslaitos).

Kouluhallitus (1986). *Peruskoulun ja lukion tuloksellisuus. Ehdotus tuloksellisuuden arviointijärjestelmäksi kouluhallituksessa sekä peruskoulun tilannekartoitus vuodelta 1986*. [Performance of Comprehensive and Upper Secondary Schools.] Osa 1. Kouluhallituksen alaisen koulutoimen ja vapaan sivistystyön tuloksellisuutta käsittelevä työryhmä. (Helsinki: Kouluhallitus).

Laukkanen, R. (1993). "Arvioinnin valtakunnallinen merkitys." [The national significance of evaluation] In K. Hämäläinen, R. Laukkanen, and A. Mikkola (eds) *Koulun tuloksellisuuden arviointi [Evaluation of School Performance]*, 54–66. (Helsinki: Opetushallitus and Helsingin yliopiston Vantaan täydennyskoulutuslaitos).

Laukkanen, R. (1995). "Itsearviointi osana arvioinnin kehittämistä" [Self-evaluation as part of evaluation development]. In B. Kilpinen, K. Salmio, L. Vainio, and Vanne, A. (eds) *Itsearviointin teoriaa ja käytäntöä [Theory and Practice of Self-evaluation]*, 69–74. (Helsinki: Opetushallitus).

Laukkanen, R. (1998). "Accountability and evaluation: Decision-making structures and the utilization of evaluation in Finland." *Scandinavian Journal of Educational Research* 42(2): 123–133. <https://doi.org/10.1080/0031383980420202>

Laukkanen, R. (ed.) (1994) *Koulutus kansainvälisessä vertailussa [Education in International Comparison]*. Helsinki: (Helsinki: Opetushallitus ja Opetushallitus).

Laukkanen, R. (ed.) (1997) *OECD-maiden koulutuspolitiikan analyysi [Analysis of OECD Countries' Education Policy]*. (Helsinki: Opetushallitus ja OECD).

Law (365/1995). *Kuntalaki 365/1995 [Law on Municipalities]*.

<https://www.finlex.fi/fi/laki/alkup/1995/19950365>

Law (476/1983). *Peruskoululaki [Comprehensive School Act]*.

<https://www.finlex.fi/fi/laki/alkup/1983/19830476>

Law (628/1998). *Perusopetuslaki. [Basic Education Act]*.

<https://www.finlex.fi/fi/laki/ajantasa/1998/19980628>

Law (707/1992). *Laki peruskoululain muuttamisesta 707/1992 [Law on the Amendment of the Comprehensive School Act]*. <https://www.finlex.fi/fi/laki/alkup/1992/19920707>

Leeuw, F. L. (2002). Evaluation in Europe 2000: Challenges to a growth industry. *Evaluation*, 8(1): 5–12. <https://doi.org/10.1177/2F1358902002008001743>

Leimu, K. (2004). ”Esipuhe” [Foreword]. In K. Leimu (ed.) *Kansainväliset IEA-tutkimukset Suomikuva luomassa* [International IEA Studies Creating the Image of Finland], 3–5. (Jyväskylä: Koulutuksen tutkimuslaitos).

Leino-Kaukiainen, P. and Heikkinen, A. (2011). ”Yhteiskunta ja koulutus” [Society and education]. In A. Heikkinen and P. Leino-Kaukiainen (eds) *Valistus ja koulunpenkki. Kasvatus ja koulutus Suomessa 1860-luvulta 1960-luvulle* [The Enlightenment and the School Bench: Training and Education in Finland from the 1860s to the 1960s], 16–33. (Helsinki: SKS).

Lyytinen, H. K. (1993). ”Itsearviointin kehitysnäkymiä.” [The development of views on self-evaluation] In K. Salmio (ed.) *Tuloksellisuuden arviointi koulun kehittämisen välineenä* [Evaluation of Performance as a Tool for School Development], 71–89. (Helsinki: Opetushallitus).

Lyytinen, H. K. and Rautama, P. (2002). ”Tarkastustoiminnasta valtion alueelliseen sivistyshallintoon” [From inspections to the state provincial educational administration. In T. Holopainen, E. Lukkarinen, and P. Rautama (ed.) *Paluu kansalliseen koulutuspolitiikkaan: Kirjoituksia Suomalaisesta koulutuksesta ja koulutuspolitiikasta* [The Return to a National Education Policy: The Literature on Finnish Education and Education Policy], 202–213. (Helsinki: Valtion alueellisen sivistyshallinnon virkamiehet VSV ry.).

Lyytinen, H. K., Jokinen, H., and Rask, S. (1989). *Koulun työkäytäntöjen arviointi. Prosessiarvioinnin näkökulma* [Evaluating School Work Practices: A View of Process Evaluation]. (Helsinki: Kouluhallitus).

Marginson, S. and Rhoades, G. (2002). “Beyond national states, markets, and systems of higher education: A glonacal agency heuristic.” *Higher Education* 43(3), 281–309.

<https://doi.org/10.1023/A:1014699605875>

Martin, Michael O., Mullis, Ina V. S., Foy, Pierre, and Martin Hooper (2016). *TIMSS 2015 International Results in Science*. (Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College).

Martin, Michael O., Mullis, Ina V.S., Gonzales, Eugenio J., Gregory, Kelvin D., Smith, Teresa A., Chrostowski, Steven J., Garden, Robert A., and Kathleen M. O'Connor (2000). *TIMSS 1999 International Science Report. Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. (Boston, MA: International Study Center Lynch School of Education Boston College).

Mullis, Ina V. S., Martin, Michael O., Foy, Pierre, and Alka Arora (2012). *TIMSS 2011 International Results in Mathematics*. (Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College).

Mullis, Ina V. S., Martin, Michael O., Foy, Pierre, and Kathleen T. Drucker (2012). *PIRLS 2011 International Results in Reading*. (Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College).

Mullis, Ina V. S., Martin, Michael O., Foy, Pierre, and Martin Hooper (2016). *TIMSS 2015 International Results in Mathematics*. (Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College).

Mullis, Ina V. S., Martin, Michael O., Foy, Pierre, and Martin Hooper (2017). *PIRLS 2016 International Results in Reading*. (Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College).

Mullis, Ina V. S., Martin, Michael O., Gonzalez, Eugenio J., Gregory, Kelvin D., Garden, Robert A., O'Connor, Kathleen M., Chrostowski, Steven J., and Teresa A. Smith (2000). *TIMSS 1999 International Mathematics Report. Findings from IEA's Repeat of the Third International*

Mathematics and Science Study at the Eighth Grade. (Boston, MA: International Study Center Lynch School of Education Boston College).

Nikander, E. (2011). ”Kirkollisen ja maallisen kouluhallinnon eroaminen” [The separation of ecclesiastical and secular education administration]. In A. Heikkinen and P. Leino-Kaukiainen (eds) *Valistus ja koulunpenkki. Kasvatus ja koulutus Suomessa 1860-luvulta 1960-luvulle* [The Enlightenment and the School Bench: Training and Education in Finland from the 1860s to the 1960s], 140–155. (Helsinki: SKS).

OECD (2002). *Reading for Change: Performance and Engagement across Countries. Results from PISA 2000.* (Paris: OECD).

OECD (2004). *Learning for Tomorrow’s World. First Results from PISA 2003.* (Paris: OECD).

OECD (2006). *PISA 2006. Science Competencies for Tomorrow’s World: Volume 1: Analysis.* (Paris: OECD).

OECD (2010). *PISA 2009 Results: Learning Trends Changes in Student Performance Since 2000 (Volume V).* (Paris: OECD).

OECD (2014). *PISA 2012 Results in Focus. What 15-year-olds know and what they can do with what they know.* (Paris: OECD).

OECD (2016). *Pisa 2015 Results. Excellence and Equity in Education.* (Paris: OECD Publishing).

OECD. (1973). *A Framework for Educational Indicators to Guide Government Decisions.* (Paris: OECD).

Olkinuora, E. and Perkki, L. (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). *Kansallisen koulutuksen arviointineuvoston perustaminen* [Foundation of Finnish Education Evaluation Council]. Opetus- ja kulttuuriministeriön työryhmämuistioita ja selvityksiä 2012:19. (Helsinki: Opetus- ja kulttuuriministeriö).
<http://urn.fi/URN:ISBN:978-952-263-167-1>

Opetus- ja kulttuuriministeriö (2012). *Koulutuksen arviointisuunnitelma vuosille 2012–2015* [Education Evaluation Plan for 2012–2015]. Opetus- ja kulttuuriministeriön julkaisuja 2012:14. (Helsinki: Opetus- ja kulttuuriministeriö). <http://urn.fi/URN:ISBN:978-952-263-117-6>

Opetus- ja kulttuuriministeriö (2012). *Perusopetuksen laatukriteerit: Perusopetuksen, perusopetuksen aamu- ja iltapäivätoiminnan sekä koulun kerhotoiminnan laatukriteerit* [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: Opetus- ja kulttuuriministeriö). <http://urn.fi/URN:ISBN:978-952-263-170-1>

Opetushallitus (1994). *Peruskoulun opetussuunnitelman perusteet* [Core Curriculum for the Comprehensive School]. (Helsinki: Opetushallitus).

Opetushallitus (1995). *Koulutuksen tuloksellisuuden arviointimalli* [Evaluation Model for School Performance]. (Helsinki: Opetushallitus).

Opetushallitus (1998). *Koulutuksen tuloksellisuuden arviointimalli* [Evaluation Model for School Performance]. (Helsinki: Opetushallitus).

Opetushallitus (2004). *Perusopetuksen opetussuunnitelman perusteet*. [Core Curriculum for Basic Education.] (Helsinki: Opetushallitus). https://www.oph.fi/download/139848_pops_web.pdf

Opetushallitus (2014). *Perusopetuksen opetussuunnitelman perusteet* [Core Curriculum for Basic Education]. (Helsinki: Opetushallitus).
https://www.oph.fi/download/163777_perusopetuksen_opetussuunnitelman_perusteet_2014.pdf

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: 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: Opetusministeriö).

Opetusministeriö (1997). *Koulutuksen arviointistrategia* [Evaluation Strategy for Education]. (Helsinki: Opetusministeriö).

Opetusministeriö (2004). *Koulutus ja tutkimus 2003–2008. Kehittämissuunnitelma* [Education and Research 2003–2008. Development Plan]. (Helsinki: Opetusministeriö).

<http://urn.fi/URN:ISBN:952-442-690-0>

Opetusministeriö (2010). *Perusopetuksen laatukriteerit* [Quality Criteria for Basic Education]. Opetusministeriön julkaisuja 2010:6. (Helsinki: Opetusministeriö). <http://urn.fi/URN:ISBN:978-952-485-844-1>

Oppi ja laatu -hanke. (1996). *Opetustoimen laadun arviointiperusteet. Virikkeitä sisäiseen arviointiin ja kehittämiseen* [Basis for Evaluating Quality in the Education Sector: Suggestions for Internal Evaluation and Development]. (Helsinki: Efeko).

Oppi ja laatu -hanke. (1998). *Opetustoimen laadun arviointiperusteet. Virikkeitä sisäiseen arviointiin ja kehittämiseen* [Basis for Evaluating Quality in the Education Sector: Suggestions for Internal Evaluation and Development]. (Helsinki: Efektia Oy).

Oppi ja laatu -hanke. (2003). *Opetustoimen laadun arviointiperusteet* [Basis for Evaluating Quality in the Education Sector]. (Helsinki: Efektia Oy).

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: University of Helsinki). <http://urn.fi/URN:ISBN:978-951-51-5238-1>

Power, M. (1994). *The Audit Explosion*. (London: Demos).

Räisänen, A., and Rönholm, H. (2006). *Itsearviointi koulu yhteisöä kehittäväksi. EFQM-arviointimalli yleissivistävässä koulutuksessa* [Self-evaluation in Developing the School Community: the EFQM Evaluation Model in General Education]. (Jyväskylä: Koulutuksen arviointineuvosto).

Rask, S. (1993). "Itsearviointi koulun kehittämisen välineenä. [Self-evaluation as a school development tool]". In K. Salmio (ed.) *Tuloksellisuuden arviointi koulun kehittämisen välineenä* [Performance evaluation as a school development tool], 97–106. (Helsinki: Opetushallitus).

Rautalin, M. (2013). *Domestication of international comparisons: The role of the OECD Programme for International Student Assessment (PISA) in Finnish education policy*. (Tampere: Tampere University Press). <http://urn.fi/URN:ISBN:978-951-44-9278-5>

Sahlberg, P. (2011), "PISA in Finland: An education miracle or an obstacle to change?", *CEPS Journal: Center for Educational Policy Studies Journal*, 1(3): 119–140.

<https://ojs.cepsj.si/index.php/cepsj/article/view/418>

Sahlberg, P. (2016). "The global educational reform movement and its impact on schooling." In K. Mundy, A. Green, B. Lingard, and A. Verger (eds) *The Handbook of Global Education Policy*, 128–144. (West Sussex: Wiley-Blackwell.)

Salmio, K., and Vainio, L. (1995). ”*Johdanto*. [Introduction]” In B. Kilpinen, K. Salmio, L. Vainio, and A. Vanne (eds) *Itsearviointien teoriaa ja käytäntöä* [Self-evaluation Theory and Practice], 7–11. (Helsinki: Opetushallitus).

Schatz, M., Popovic, A., and Dervin, F. (2015). “From PISA to national branding: exploring Finnish education®.” *Discourse: Studies in the Cultural Politics of Education* 26(2), 172–184.
<https://doi.org/10.1080/01596306.2015.1066311>

Seppänen, P., Rinne, R., Kauko, J., and Kosunen, S. (2019). “The use of PISA results in education policy-making in Finland.” In F. Waldow, and G. Steiner-Khamsi (eds) *Understanding PISA’s Attractiveness: Analysis in Comparative Policy Studies*, 137–160. (London: Bloomsbury).

Simola, H., Kauko, J., Varjo, J., Kalalahti, M. and Sahlström, F. (2017). *Dynamics in Education Politics – Understanding and explaining the Finnish Case*. (London: Routledge).

Simola, Hannu (2002). “From exclusion to self-selection: examination of behaviour in Finnish primary and comprehensive schooling from the 1860s to the 1990s”. *History of Education* 31(3): 207–226. <https://doi.org/10.1080/00467600210122595>

Smith, D. (1990). “Assessment, technology and the quality revolution”. In T. Schuler (ed.) *World Yearbook of Education 1990: Assessment and Evaluation*, 41–55. (London: Kogan Page).

Temmes, M., Ahonen, P., and Ojala, P. (2002). *Suomen koulutusjärjestelmän hallinnon arviointi* [An Evaluation of the Governance of the Finnish Education System]. (Helsinki: Opetusministeriö).
<http://urn.fi/URN:ISBN:952-442-974-8>

Väätäinen, 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).
https://karvi.fi/app/uploads/2019/04/KARVI_T0519.pdf

Väljjarvi, J. and Kupari, J. (2015). *PISA. Millä eväillä uuteen nousuun? PISA 2012 tutkimustuloksia* [PISA. What Is Needed for a Resurrection? Research Results from PISA 2012]. Opetus- ja kulttuuriministeriön julkaisuja 2015:6. (Helsinki: Opetus- ja kulttuuriministeriö).

<http://urn.fi/URN:ISBN:978-952-263-334-7>

Valtioneuvosto (1990). *Valtioneuvoston koulutuspoliittinen selonteko eduskunnalle 22.5.1990: Suomen koulutusjärjestelmä, koulutuksen taso ja kehittämissuunnitelmat* [Government report on education policy for Parliament, 22 May 1990: Finnish education system, level of education, and development trajectories]. (Helsinki: Valtioneuvosto).

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: Helsingin yliopisto, käyttäytymistieteiden laitos).

<http://urn.fi/URN:ISBN:978-952-10-3640-8>

Varjo, J., Simola, H., and Rinne, R. (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: Suomen kasvatustieteellinen seura).

Verger, A., Parcerisa, L., and Clara Fontdevila (2019). “The growth and spread of large-scale assessments and test-based accountabilities: a political sociology of global education reforms.” *Educational Review* 71(1), 5–30. <https://doi.org/10.1080/00131911.2019.1522045>

Vuorela, T. (1980). *Opetusministeriön historia II. Taantumuksesta uudistukseen 1825–1868* [The History of the Ministry of Education II: From Recession to Reform 1825–1868]. (Helsinki: Opetusministeriö).