Chapter 11
Educational Assessment in Finland

Mari-Pauliina Vainikainen and Heidi Harju-Luukkainen

The Finnish education has received a lot of attention after decades of relatively high performance in international student assessments. Even though the Finnish education system has received a lot of interest, very little attention has been paid to the model of the Finnish educational assessment system and the lack of standardised measurement and control. Thus, these factors in large are contributing to the overall functioning of the system. In this chapter we provide a historical overview of the development of the assessment model in Finland and further give a description of its current form. We also give an example of the Finnish PISA 2012 oversampling and its results. Finally, we make some critical suggestions on how the system could be improved without adding unnecessary controlling elements to it.

Introduction

The Finnish education system, still after almost two decades, continues to receive international attention due to the relatively high performance level of its students in international educational assessment studies like the Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS) (e.g. OECD 2016). The education system in Finland differs from the high-performing countries in several aspects. An important difference is that the results of Finland are high despite average economic investments into education (OECD 2018).

M.-P. Vainikainen (✉)
Tampere University, Tampere, Finland
e-mail: mari-pauliina.vainikainen@tuni.fi

H. Harju-Luukkainen
Nord University, Levanger, Norway

H. Harju-Luukkainen et al. (eds.), *Monitoring Student Achievement in the 21st Century*, https://doi.org/10.1007/978-3-030-38969-7_11

Finnish schools also seem to be exceptionally equitable in terms of the low level of segregation both by the distribution of socio-economic status of pupils and by their performance levels (Willms 2010). Also the school level differences are among the smallest ones in the world (OECD 2016).

When it comes to the basic principle of the education system, there are a lot of commonalities between the different Nordic countries (Garvis and Eriksen Ødegaard 2017). Therefore, Finland is implementing alongside with the rest of the Nordic countries a so-called Nordic model (Antikainen 2006; Telhaug et al. 2006) or sometimes even defined as a 'Nordic dialogue' (Garvis and Eriksen Ødegaard 2017). According to this model, all students should have equal opportunities and possibilities regardless of their socio-economic status or residential area. Children are also attending their local school, without any tracking, until they are at the end of their compulsory education. This model was introduced in Finland in the Basic Education Act of 1968, and it was gradually implemented from 1972 to 1976.

In Finland, children start their school the year they turn seven. In most cases, children are enrolled in their local school, but in some cases, there is a possibility to apply to a school with a specialised language, music or other programmes during the lower grades. The emphasis on children attending the local public school was strengthened in the legislation reform in 2011 (Thuneberg et al. 2013), and the statistics from the following year show that 96% of the comprehensive schools were run by municipalities (the Official Statistics of Finland, www.stat.fi). Therefore, there are almost no private actors among the Finnish schools. Some adjustments have been made to the education system since the 1970s. For instance, the ability grouping widely practiced during the first decade was officially abolished in the mid-1980s, but otherwise the structure remains unchanged. Different to many other countries, the first official point of tracking occurs only after the ninth year of schooling. At this point, students have to choose between academic and vocational tracks of upper secondary school.

Even though the Finnish education system has received a lot of interest, very little attention has been paid to the model of the Finnish educational assessment system and the lack of standardised measurement and control (see Vainikainen et al. 2017). Thus, these factors in large are contributing to the overall functioning of the system. The aim of this chapter is firstly to provide a historical overview of the development of the assessment model in Finland and further to give a description of its current form. We also give an example of the Finnish PISA 2012 oversampling and its results. Finally, we make some critical suggestions on how the system could be improved without adding unnecessary controlling elements to it. After all, the freedom of municipalities, schools and teachers in organising the education according to their best understanding and implementing different aspects of the curricula might have a higher role in shaping the good educational outcomes of the Finnish youth than previously recognised.

Introduction to the National and International Assessment Context and Its History in Finland

The Finnish educational assessment model has evolved in several stages until it has reached its current, relatively noncontrolling structure (Varjo et al. 2016). The early decades of the comprehensive education system were characterised by a strict control of particularly *inputs* and to some extent also *outputs* (cf. OECD 2015). The inputs were regulated through a detailed national curriculum with state-level obligatory in-service teacher training of obligatory contents of it and pre-examination of textbooks. Outputs were controlled by an active school inspection system that held schools accountable for achievement, the same way it is nowadays done in many other countries (Gustafsson et al. 2015). However, the attempt of introducing standardised national exams in major school subjects in the 1970s failed, and therefore teacher-given school grades became the primary measure of achievement. At that time, grades followed a normal distribution within each class, leading to a situation, in which between-school or class differences were not recognised. Also standardised test was provided for teachers only to facilitate the grading process.

During curriculum reform in 1985, the obligatory national curriculum was replaced by a *National Framework Curriculum*, which gave the municipalities as organisers of education more freedom for local decision-making. It also gave them wider possibilities to assess the outputs of the system locally even though at this point, the national school inspection system was still active. Besides that, there was no national assessment system, and the country's participation in the International Educational Assessment (IEA) surveys was irregular. As the recommendations given in the National Framework Curriculum on grading were relatively unspecific, too, grading was still largely done at class level on a normative scale even though the aim was to adopt a criterion-referenced model and not to rank students.

The next curriculum reform took place in 1994, taking local-level decision-making on the next stage. The *national core curriculum* contained only the obligatory core, and municipalities and/or schools had to write their own curricula based on it. Organisers were also expected to assess the outcomes of the education they provided. Further, also the school inspection system and controlling of the learning materials had been ceased before the new core curriculum was introduced. Therefore, at this time Finland experienced its first period of decentralisation of monitoring mechanisms for educational outputs, even though practices for evaluating educational outcomes (see National Board of Education 1999, for the revised English version) had not yet been formally introduced. This, however, did not last long as the final version of the assessment practices for evaluating educational outcomes was published in 1995 and the more comprehensive version in 1998 simultaneously with the educational legislation reform. Educational assessment was now also defined as the organiser's responsibility in the legislation. In practice, these local assessments were largely based on self-evaluation.

The national monitoring model for evaluating educational outcomes was introduced (National Board of Education 1999). It was designed to provide some indicators of performance trends and further needs for the national policy development work. On an international scale, this model was light, and accountability was not a part of it on any level. On the contrary, the introduced approach made it almost impossible to make any conclusions on individual school level, and this new type of data was solely for national monitoring purposes.

In this model the educational outcomes, were divided into three main categories that all comprised several subcategories. The first category, *efficiency*, measured the functioning of the educational system, whereas the second category *effectiveness* was about student-level outcomes. The third category was *economy* for successful allocation of resources. From the perspective of educational assessment, the second category was the most interesting one as it posed requirements of conducting external assessments. In practice, this led to two kinds of applications. First, sample-based *curricular assessments* were developed to measure learning outcomes in most important school subjects. Second, *national thematic assessments* were developed to provide information on a wider scope of educational outcomes to complement the information obtained from international large-scale assessments that were now emerging: as the first cycle of PISA in 2000 and by rejoining the IEA assessments. Curricular sample-based assessments are still in 2018 not implemented regularly at predefined grade levels in major subjects, but the subjects and grade levels to be assessed are instead specified for a few years at a time in a *plan for educational assessment*. The current organisation of these assessments is described below. In the 1990s, the National Board of Education that was also responsible for curriculum development had a unit for implementing them.

In the mid-2000s, Finland had the second period without a clear national structure for educational assessment. For the whole decade, the National Board of Education implemented curricular sample-based assessments, and Finland participated in PISA (and more irregularly in TIMMS and PIRLS, but not between 1999 and 2011). However, the national coordination of assessments was restructured several times (see Varjo et al. 2016), which led to a situation where many thematic assessments – including the national learning to learn assessment programme – were ceased even though no formal decisions were made of not having them. At the same time, the new core curriculum of 2004 continued to give organisers of education a lot of freedom and responsibility to define their own assessment practices to fulfil the requirements stated in the 1998 legislation. Both general and subject-specific assessment criteria were specified and harmonised with curricular goals, and descriptions of 'good performance' were given to facilitate both formative and summative assessments. Themes introduced during the previous decades about cross-curricular or transversal competences were to some extent included in the core curriculum as general goals, but relatively little was said about them in assessment criteria or descriptions about subject-specific goals.

National and International Assessment in Finland Today

National Assessments

Until recently, the Finnish National Board of Education used to conduct national assessments on students' learning achievement. Since May 2014, this is now a duty of the Finnish Education Evaluation Centre (FINEEC). Among other things, the FINEEC is responsible for evaluating learning outcomes with respect to the distribution of lesson hours and the national core curriculum targets stipulated in the Basic Education Act (628/1998). The assessment of learning outcomes is based on sampling. Typical sample sizes comprise 5–10% of the age group, which means that each assessment involves about 4000–6000 students (Jakku-Sihvonen 2013, 24). The assessed schools represent around 15% of all the schools that give basic education in Finland (Ouakrim-Soivio, 2013, 21; Harju-Luukkainen et al. 2016a).

According to Harju-Luukkainen et al. (2016a), the assessment of learning outcomes can be viewed from many perspectives, and it has got different purposes for different target groups. National assessments provide valuable information for the highest educational authorities. In Finland, basic education is expected to secure equal educational opportunities for all students. Therefore, the equity of learning outcomes is studied from several perspectives, for example, those of students' gender, region, type of municipality and socio-economic background as well as language spoken at school. In principle, reaching the objectives for equal learning opportunities as defined in the national core curricula should lead to educational equity so that there would be no statistically significant differences between the learning outcomes of boys and girls, for example, or between different regions in Finland.

Secondly, from the school's perspective, the national assessments of learning outcomes provide benchmarks for schools to evaluate their own success in reaching their objectives of teaching and learning in different subjects. Schools selected to an assessment receive feedback in the form of reference data on the results and learning-related perceptions of their own students. Because there are no national examinations at the end of basic education, many schools welcome this opportunity to compare their own results and grading practices to the national benchmarks and use the assessment as a tool to develop their instruction in different subjects (Ouakrim-Soivio and Kuusela 2012, 13; Harju-Luukkainen et al. 2016a).

Thirdly, teachers assess each student based on student performance. At the end of basic education (i.e. grade 9 in the comprehensive school), most of the students are 15-year-olds and about to finish their compulsory education. Grading is obligatory at the final phases of basic education (grades 8 & 9), but most schools begin to use numerical grades already at earlier grade levels. In Finland, the national core curriculum for basic education determines the learning objectives for each school subject. Also grading guidelines are given but with specific description for good

competence only, which equals the grade 8 on the student assessment scale ranging from 4 to 10, where 4 means failed and 10 is the highest grade. This good competence level serves as a baseline for assessment, and it should help ensure an objective evaluation for all students attending basic education. Objective evaluation at this point is of great importance; the grades obtained in different subjects at the end of compulsory education will largely determine the next steps in the student's educational path. In sum in Finland today, student assessments at different levels (national, school or individual level) all strive for the same goal: higher equality in education (Harju-Luukkainen et al. 2016a).

International Assessments

Finland has been participating in PISA assessments since the first cycle in 2000. In addition, Finland has taken part to IEA assessments (PIRLS, TIMSS) in the recent cycles, but earlier the country's participation in these assessments has been too irregular to monitor trends based on these data. Thus, at basic education level, national discussions about performance trends in international assessments are largely based on PISA. When the results of the PISA 2000 cycle were first published, they provoked surprisingly little public discussion. The high ranking was not expected, and the reception of the news was almost sceptical. Yet, it most likely changed the course of educational political discussion as there had been voices claiming that the comprehensive education system does not support the optimal development of students with higher academic goals. The results showed that high-performing students did not do any worse than their counterparts in other countries, whereas the weakest students clearly outperformed their comparison groups anywhere else (Hautamäki et al. 2009). Thus, the results may have contributed to the basic education legislation changes that have strengthened the main principles of the Nordic educational ideas even further (see Thuneberg et al. 2013). The first PISA results were used as evidence for that the structure of the education system did not need extensive reforms. Accordingly, the declining trend observed both in international and national assessment studies since 2006 (Hautamäki et al. 2013; Vettenranta et al. 2016) has been taken seriously, and programmes have been launched to turn the trend again. These programmes have included thematic assessments that go deeper into the details of the national features of the education system (e.g. support structures) and additional funding for municipalities and schools to improve their practices. The data of international assessments are also utilised in this purpose through more detailed analyses.

An Example: Oversampling PISA 2012 in Finland

Like many other European societies, Finland has faced many changes during the past decade. One major change has been the increasing number of students speaking languages other than the national ones. Due to the increase in the Finnish migrant population, students with a migrant background were oversampled for the first time in PISA 2012. In this, Finland was the second country to conduct an oversampling of one of its student population, after Denmark. Oversampling means that more students were selected for testing than would be their true proportion in the population. The oversampling made it possible to gain more representative data on students with migrant backgrounds. The Finnish PISA data on migrant students consisted of 691 first-generation and 603 second-generation students, most of whom lived in the metropolitan areas. The rest of the data comprised a total of 7535 students across the country (see Harju-Luukkainen et al. 2014). In the following, we will give an overview of these results (see more Harju-Luukkainen and McElvany 2018).

A first report on these national findings was published in 2014. The PISA 2012 migrant data was analysed and the reports published just before the migration crisis hit Europe in 2015. During this time, more than a million migrants and refugees crossed Europe, coming mainly from Syria, Afghanistan, Iraq, Kosovo and Albania to Finland among other countries. Therefore, the oversampling of students with a migrant background in Finland gave an important insight on how Finland had managed in educating their language minority students.

Even though Finland is known for its good educational outcome, the results were not that positive. According to Harju-Luukkainen et al. (2014), the results of students with migrant backgrounds were alarming compared to their nonmigrant peers in Finland. As shown in Fig. 11.1, students with migrant backgrounds performed poorer in PISA 2012 across all assessment domains compared to their nonmigrant peers. The differences were statistically significant. In mathematics, for instance, nonmigrant students received a mean score of 522 points; first- and second-generation students with migrant backgrounds received 425 points and 449 points, respectively. The definition of who is counted as first- and second-generation migrants can vary in different studies. However, according to the OECD and PISA 2012 assessment, first-generation migrant students are those who have immigrated to Finland during their lifetime. Second-generation students in turn have been born and raised in Finland, but both of the parents were born outside Finland. According to the OECD (2014, p. 16), 41 points scored equals approximately 1 year of schooling. First-generation students were therefore lagging behind by more than 2 school years and second-generation students by almost 2 school years. The results were more or less similar in scientific literacy and reading literacy as well. Similar

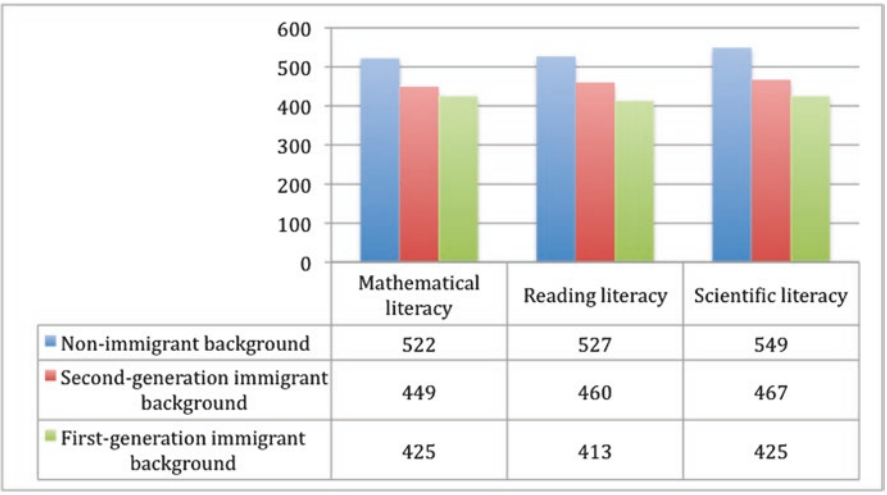

Fig. 11.1 Mathematic, reading and scientific literacy of different student groups in PISA 2012. (Source: Harju-Luukkainen et al. 2014, p. 25)

Table 11.1 Percentage of students on different performance levels in PISA 2012 (Harju-Luukkainen et al. 2014, p. 27)

Performance level	First generation	Second generation	Nonmigrant
6	0.7	0.4	3.6
5	2.3	2.4	12
4	8.3	10.4	23.7
3	14.9	22.1	29.2
2	22.2	26.6	20.4
1	26.2	24.2	8.3
Below level 1	25.3	13.9	2.7
	100%	**100%**	**100%**

observations have been done in all other PISA cycles even though the migrant sample has been smaller.

In the data especially the low percentage of high performers as well as high level of low performers among the migrant students was observed. According to Harju-Luukkainen et al. (2014) on the highest performance levels (levels 5 and 6), there were almost no students with a migrant background (varying between 0.4% and 2.4%) and a very small difference between the proportions of first- and second-generation students at the highest performance levels. Further, 51.5% of first-generation migrant students were at the lowest performance levels (level 1 and below), as were 38.1% second-generation migrants. According to Harju-Luukkainen and McElvany (2018), it is troubling when the second-generation migrants have taken part of the entire Finnish education system, the performance is still on a very

low level, and the difference between the first- and second-generation migrants is relatively small (Table 11.1).

The migration population in Finland is still relatively small and heterogenous. For this reason, there have only been a few studies looking into the reasons behind these differences. According to Harju-Luukkainen et al. (2014), mathematical literacy performance in all native student groups in Finland was explained by such variables as self-concept for mathematics, confidence in mathematics performance and anxiety for mathematics (see also Harju-Luukkainen et al. 2016b). However, the explanatory power of these variables was weaker for students with a migrant background than for other students. Therefore, there is most likely a wider range of underlying factors for this minority group that are either unknown or at least beyond the scope of PISA assessments. Harju-Luukkainen et al.'s (2015) investigation of resilient second-generation migrants students' educational outcomes in mathematics found that the factors connected to good educational outcomes were (1) the family's language choices, (2) high ESCS (student's socio-economic and cultural status index), (3) cultural closeness, (4) teacher's support and individualisation of teaching materials, (5) low truancy and intact learning continuums and (6) strong self-concept in mathematics (see also Harju-Luukkainen et al. 2017). How well a student masters the language of instruction seems therefore to be one of the most important factors, which is something that Kuukka and Metsämuuronen (2016) and Saario (2012) also emphasise. The study conducted by the FEEC (Kuukka and Metsämuuronen 2016) revealed that migrant pupils' Finnish language skills were good, already in the upper grades of comprehensive school. However, the concept of text skills of various subjects requires more from the pupils than is required by the criterion of a proficiency scale. Therefore, it is crucial to ask if the different assessments capture the true level of migrant students' competencies and skills.

The degree to which these family-related attributes have an impact on students' educational outcomes varies not only from country to country (OECD 2010) but also within countries (Harju-Luukkainen and McElvany 2018). In Finland, there are to be found differences between the different student groups and how their family-related attributes affect the students' educational outcome. According to PISA 2012, the ESCS index (student's socio-economic and cultural status index) explained 11% of the variance between first-generation students, 7% of the variance between second-generation students and 8% of the variance between students without migrant backgrounds (Harju-Luukkainen et al. 2014). Further, the assessment conducted by the FEEC revealed that variables which could best explain students' low learning outcomes from different language groups were related to students' socio-economic background. In all, the connection between the socio-economic background and learning outcomes was significant (Kuukka and Metsämuuronen 2016). In these abovementioned studies (as well as many other studies), the ESCS has not been controlled. In a study conducted by Kilpi-Jakonen (2012, p. 167), the study revealed that differences between migrant and nonmigrant student groups in Finland are relatively small after controlling for parental resources. Kilpi-Jakonen (2011; 2012) concluded that parental education and parental income have smaller and larger effects, respectively, for children of migrants than for nonmigrants. This

leads to a disadvantaged group with migrant parents who have high education levels but low incomes (Harju-Luukkainen and McElvany 2018). According to Kalalahti et al. (2017), the youth with migrant background in Finland, especially boys, share a 'paradox of immigrant schooling' which refers to the positive attitude towards education, but at the same time, they face difficulties in learning and studying. Overall, according to Kilpi-Jakonen (2012), children of migrants can be seen to benefit from the relatively equal Finnish education system while remaining disadvantaged by their parents' difficulties in the labour market (see more Harju-Luukkainen and McElvany 2018; Karppinen 2008; Kilpi-Jakonen 2011).

Critical Discussion of the Country's Assessment Policies, Practices and Results

Finland has developed its assessment policies and streamlined many of the practices during the history of its basic education. A further developmental object is to find a balance in how the results of the national assessments as well international assessments are used on national level.

However, still a national central organ that oversees both the national and international assessment policies, practices as well as results does not exist. This has lead in some cases to a situation, where participation in some important international assessment has been irregular and there has been problems with the capability of national sample-based assessments to produce enough comparable data for monitoring of trends. In the Finnish context, it still seems functional not to have extensive standardised examinations as the lack of them gives more freedom to schools and teachers to implement the curriculum in a purposeful way. However, the declining trend and the increasing regional differences (Vettenranta et al. 2016) call for a slightly more detailed monitoring system that could be realised within the current educational assessment model by securing sufficient coverage of school subjects, age groups, at-risk subpopulations and geographical areas.

Lack of resources is also a risk for Finnish assessment context. The government of Finland should direct enough of funding towards the analysis of already collected datasets in order to reveal, for instance, possible negative trajectories behind educational outcomes of different student groups. As the situation is now, only basic reporting and data collection can be done with the governmental funding. This might lead to a situation that education policy decisions, which are solely done on the basis of national reports, can be weakly justified.

References

Antikainen, A. (2006). In search of the Nordic model in education. *Scandinavian Journal of Educational Research, 50*(*3*), 229–243.

Garvis, S., & Eriksen Ødegaard, E. (2017). *Nordic dialogues on children and families. Evolving families*. Oxon: Routldege.

Gustafsson, J. E., Ehren, M. C. M., Conyngham, G., McNamara, G., Altrichter, H., & O'Hara, J. (2015). From inspection to quality: Ways in which school inspection influences change in schools. *Studies in Educational Evaluation, 47*, 47–57.

Harju-Luukkainen, H., & McElvany, N. (2018). Immigrant student achievement and educational policy in Finland. In L. Volante, D. Klinger, & Ö. Bilgili (Eds.), *Immigrant student achievement and education policy: Cross-cultural approaches* (pp. 87–102). Cham: Springer.

Harju-Luukkainen, H., Nissinen, K., Sulkunen, S., & Suni, M. (2014). *Avaimet osaamiseen ja tulevaisuuteen: Selvitys maahanmuuttajataustaisten nuorten osaamisen tasosta ja siihen liittyvistä taustatekijöistä PISA 2012 – tutkimuksessa* [Keys to competence and future: A report on PISA 2012 results and related underlying factors for students with an immigrant background]. Jyväskylä: Finnish Institute for Educational Research.

Harju-Luukkainen, H., Nissinen, K., & Mirja, T. (2015). Matematiikka ja maahanmuuttajataustaiset nuoret [Mathematic literacy and immigrant students]. In J. Välijärvi & P. Kupari (Eds.), *Millä eväillä tulevaisuuden peruskoulu nousuun?* (pp. 108–123). Helsinki: Ministry of Education and Culture.

Harju-Luukkainen, H., Vettenranta, J., Oukrim-Soivio, N., & Bernelius, V. (2016a). Differences between PISA reading literacy scores and grading for mother tongue and literature at school: A geostatistical analysis of the Finnish PISA 2009 data. *Education Inquiry, 7*(4), 463–479. https://doi.org/10.3402/edui.v7.29413.

Harju-Luukkainen, H., Sulkunen, S., & Stolt, S. (2016b). Uppfattning av lärandestrategier hos elever som för det mesta talar finska hemma, men går i svenskspråkig skola [Understanding of reading related strategies among students who speak mostly Finnish at home but attend a Swedish language school]. *Kognition och Pädagogik, 99*, 78–89.

Harju-Luukkainen, H. Tarnanen, M., & Nissinen, K. (2017). Monikieliset oppilaat koulussa: eri kieliryhmien sisäinen ja ulkoinen motivaatio sekä sen yhteys matematiikan osaamiseen PISA 2012–tutkimuksessa [Multilingual students in school: The inner and outer motivation of students with an immigrant background and its connection to mathematical performance in PISA 2012]. In A. Huhta & R. Hildén (Eds.), *Kielitaidon arviointitutkimus 2000-luvun suomessa* [Language assessments in 20th century Finland] (pp. 167–183). Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA ry.

Hautamäki, J., Hautamäki, A., & Kupiainen, S. (2009). Educational equity account in Nordic countries. In T. Matti (Ed.), *Northern lights on PISA 2006: Differences and similarities in the Nordic countries* (TemaNord 2009:547) (pp. 157–167). Copenhagen: Nordic Council of Ministers.

Hautamäki, J., Kupiainen, S., Marjanen, J., Vainikainen, M. –P., & Hotulainen, R. (2013). *Oppimaan oppiminen peruskoulun päättövaiheessa: Tilanne vuonna 2012 ja muutos vuodesta 2001* [Learning to learn at the end of basic education: Situation in 2012 and change from 2001] (University of Helsinki. Department of Teacher Education Research Report 347). Helsinki: Unigrafia.

Jakku-Sihvonen, R. (2013). Oppimistulosten arviointijärjestelmistä ja niiden kehittämishaasteista [On assessment systems for learning outcomes and developmental challenges of these systems]. In Räisänen, A. (Ed.). *Oppimisen arvioinnin kontekstit ja käytännöt* [Contexts and practices for the assessment of learning] (Raportit ja selvitykset 2013:3, pp. 13–36). Helsinki: Opetushallitus.

Kalalahti, M., Varjo, J., & Jahnukainen, M. (2017). Immigrant-origin youth and the indecisiveness of choice for upper secondary education in Finland. *Journal of Youth Studies., 20*, 1242. https://doi.org/10.1080/13676261.2017.1321108.

Karppinen, K. (2008). Koulumenestys, koulutukseen valikoituminen, tutkinnon suorittaminen ja työelämään siirtyminen [School success, selection of education, graduation and transition to work]. In J. Kuusela, A. Etelälahti, Å. Hagman, R. Hievanen, K. Karppinen, L. Nissila, U. Rönnberg, & M. Siniharju (Eds.), *Maahanmuuttajaoppilaat ja koulutus Tutkimus oppimistuloksista, koulutusvalinnoista ja työllistämisestä* (pp. 135–186). Helsinki: Opetushallitus.

Kilpi-Jakonen, E. (2011). Continuation to upper secondary education in Finland: Children of immigrants and the majority compared. *Acta Sociologica, 54*(1), 77–106. https://doi.org/10.1177/0001699310392604.

Kilpi-Jakonen, E. (2012). Does Finnish educational equality extend to children of immigrants? Examining national origin, gender and the relative importance of parental resources. *Nordic Journal of Migration Research, 2*(2), 167–181. https://doi.org/10.2478/v10202-011-0039-4.

Kuukka, K., & Metsämuuronen, J. (2016). *Perusopetuksen päättövaiheesn suomi toisena kielenä (S2) –oppimäärän oppimistulosten arviointi 2015* [Estimating the learning outcomes of Finnish second language (S2)]. Tampere: Kansallisen koulutuksen arviointikeskus.

National Board of Education. (1999). *A framework for evaluating educational outcomes in Finland* (National Board of Education, Evaluation 8/1999). Helsinki: University Printing House.

OECD. (2010). *PISA 2009 results: Overcoming social background—Equity in learning opportunities and outcomes* (Vol. II). Paris: Author.

OECD. (2014). *PISA 2012 results: What students know and can do – Student performance in mathematics, reading and science* (Vol. I). Paris: OECD.

OECD. (2015). *Education at a glance 2015* (OECD Indicators). Paris: OECD Publishing. https://doi.org/10.1787/eag-2015-en.

OECD. (2016). *PISA 2015 results (volume I): Excellence and equity in education*. Paris: OECD Publishing. https://doi.org/10.1787/9789264266490-en.

OECD. (2018). *Education at a glance 2018* (OECD Indicators). https://doi.org/10.1787/eag-2018-en.

Ouakrim-Soivio, N. (2013). *Toimivatko päättöarvioinnin kriteerit? Oppilaiden saamat arvosanat ja Opetushallituksen oppimistulosten seuranta-arviointi koulujen välisten osaamiserojen mittareina* [Do the national criteria for students' final grades work? Teachers' grading and national assessments as indicators for between-school differences]. Helsinki: Opetushallitus.

Ouakrim-Soivio, N. & Kuusela, J. (2012). *Historian ja yhteiskuntaopin oppimistulokset perusopetuksen päättövaiheessa 2011* [Learning outcomes in history and social studies at the end of basic education 2011] (Koulutuksen seurantaraportit 2012:3). Helsinki: Opetushallitus.

Saario, J. (2012). *Yhteiskuntaopin kieliympäristö ja käsitteet. Toisella kielellä opiskelevan haasteet ja tuen tarpeet* [The language environment and concepts of social work: Challenges and support needs in another language] (Doctoral dissertation). University of Jyväskylä, Finland.

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model in education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research, 50*(*3*), 245–283.

Thuneberg, H., Vainikainen, M.-P., Ahtiainen, R., Lintuvuori, M., Salo, K., & Hautamäki, J. (2013). Education is special for all – The Finnish support model. *Gemeinsam leben, 21*(*2*), 67–78.

Vainikainen, M.-P., Thuneberg, H., Marjanen, J., Hautamäki, J., Kupiainen, S., & Hotulainen, R. (2017). How do Finns know? Educational monitoring without inspection and standard-setting. In S. Blömeke & J.-E. Gustafsson (Eds.), *Standard setting: International state of research and practices in the Nordic countries* (pp. 243–259). Cham: Springer. https://doi.org/10.1007/978-3-319-50856-6_14.

Varjo, J., Simola, H., & Rinne, R. (2016). *Arvioida ja hallita. Perään katsomisesta informaatio-ohjaukseen suomalaisessa koulupolitiikassa* [To evaluate and govern: From "looking after" to steering by information" in Finnish education policy] (Publications of the Finnish Educational Research Association, 70). Jyväskylä: Jyväskylän yliopistopaino.

Vettenranta, J., Välijärvi, J., Ahonen, A., Hautamäki, J., Hiltunen, J., Leino, K., Lähteinen, S., Nissinen, K., Nissinen, V. Puhakka, E., Rautopuro, J. & Vainikainen, M. -P. (2016). *PISA 2015 ensituloksia. Huipulla pudotuksesta huolimatta.* Opetus- ja kulttuuriministeriön julkaisuja 2016:41.

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record, 112*(*4*), 1007–1037.