

Chapter 14

Educational Assessment in Iceland



Meyvant Þórólfsson

Introduction

Iceland is a thinly populated island in the North Atlantic Ocean close to the Arctic Circle. The population is about 1/3 of a million dispersed along its coasts where the south-west part is most densely populated with the city of Reykjavik at its centre. Covering an area of 103,100 km², the island is geologically located on the Mid-Atlantic Ridge that separates the Eurasian plate and the North American plate. Thus, Iceland is among the most active volcanic areas of the world. Glaciers cover one-tenth of the island along with lots of rivers, waterfalls, geysers and fiords. The climate is warmer than the northern latitude indicates because of the warm Gulf Stream, and the bright summer nights; nevertheless, the climate is described as windy, cloudy and unstable.

For ages Iceland was a poor Danish colony, but in 1918 it received its first recognition as an independent state and in 1944 full independence was announced and The Republic of Iceland was founded. It developed rapidly due to beneficial factors, for example, the entry into the so-called Marshall plan after the Second World War, and consequently important economic, technological and scientific advances connected to fishing industries and eventually technology and innovation in other industrial sectors.

The legacy of literature has occupied the lives of Icelanders since long before Gutenberg introduced the printing press in the 1400s. Due to a variety of circumstances, nearly all the ancient literature was written in the vernacular language although most scholars and many intellectual farmers knew Latin by heart. For centuries, literacy has been considered high in Iceland and according to researchers such as Gíslason (1977) and Proppé (1983), specific circumstances have sustained the literary tradition, namely the climate and the northern location of the island with

M. Þórólfsson (✉)
University of Iceland, Reykjavik, Iceland
e-mail: meyvant@hi.is

its long and dark winter periods. People had few other choices than staying indoors where families assembled during long winter nights on evening wakes (i. kvöldvökur). Such wakes were a significant cultural tradition dating since the middle ages and lasting until the first half of the twentieth century. They involved various kinds of intellectual activities such as loud reading of the ancient sagas, poetry, rhymes, Bible reading and telling ghost stories.

During the middle ages, monasteries provided schooling and some priests and farmers also had schools in their homes (Guttormsson 1981; Proppé 1983). Cathedral schools became grammar schools after the Protestant Reformation in 1550, though not intended for the public but an elite preparing for priesthood or judicial practice. Public education did not receive much attention until the eighteenth century when the first law addressing public education came into act in 1880 and in 1907 the state agreed on providing public schools for children aged 10–14 years. By 1900, the first secondary public schools were also founded. Eventually, the basic policy was confirmed that all children should have equal opportunities to acquire basic education without any discrimination. At the compulsory level (elementary and lower-secondary, ages 5–16), education has been totally free for more than a century and now pupils are provided with all learning materials and resources.

Since the establishment of public schooling, assessing pupils and their learning has been organised and carried out by schools as part of the implemented curriculum and concurrently by educational authorities as part of the intended curriculum. Before 1970, the words evaluation (i. mat) and assessment (i. námsmat) were not found in educational discourse in Iceland. Instead, educators talked about tests (i. próf) and grades (i. einkunnir). Now the word assessment (i. námsmat) is almost exclusively used. Furthermore, discourse about different purposes of assessment emerges increasingly, i.e. assessment of learning, assessment for learning and assessment as learning.

Assessment, Testing and Grading

Influenced by the spirit of pietism during the eighteenth century and first half of the nineteenth century, priests were responsible for education and judging pupils' learning. Pupils received marks based on how well they had learned their lessons (Guttormsson 2008). After 1860, schools began to be established and gradually teachers became responsible for public education together with priests. Each day, marks were written in protocols and kept as data about learning and learning processes (Proppé 1983).

As the twentieth century commenced, most schools changed their daily grading to weekly grading, but most importantly the grades were still based upon subjective judgements of teachers and priests (Proppé 1983). The first signs of summative assessment appeared through a debate about yearly spring examinations. According to the Law on the Education of Children from 1907, pupils were to be tested 'orally' each spring in the traditional subjects, such as reading, religion, arithmetic, history,

zoology, crafts and physical education. According to regulations based on the 1907 law, grading was supposed to embody a number scale from 1 (bad) to 8 (excellent) (Proppé 1983).

By the end of the second decade of the twentieth century, ideas of psychometric methods and written tests began to emerge. In 1920, a distinguished Icelandic scholar, Steingrímur Arason, came home from his studies at Columbia University, New York, influenced by Edward Lee Thorndike's educational psychology. Arason's introduction of quantitative measurements and written tests induced immense controversy, which was no surprise because the battle between progressive thinkers and traditionalists was striking at that time. One of the largest compulsory schools in the country, Austurbæjarskóli, publicly declared itself as a progressivist school. Its principal, Sigurður Torlacius, had received his education in Europe and became familiar with 'The New School' and thinkers like Maria Montessori in Italy, Ovide Decroly in Brussels, John Dewey in America and members of the 'active school' (g. *Arbeitschule*) in Germany. Torlacius maintained that testing and grading practices were misleading because they focused on trivial skills instead of other important competences. He expressed his school's policy about testing and grading this way:

Instead of the spring examinations we need process evaluation by school specialists ... Instead of the grades, we should mainly show what the children have done ... besides that, we should have personal communications between the teacher and the home, through which information about the children can be given both ways. (Thorlacius 1932, p. 23)

But his suggestions did not receive much support, so as Ellen C. Lagemann (1989) put it, 'Thorndike won and Dewey lost'. As the British philosopher of education R. S. Peters identified (as cited in Walker and Soltis 2009, p. 14), such high-sounding aims were commitments to certain values, but their role in everyday activities of teachers turned out to be insignificant.

Steingrímur Arason managed to convince most eminent scholars that the 'new testing methods' would secure reliable judgements about learning and one of them added that most importantly they would secure fairness and equity (cf. Hjörvar 1921). Still, there were those who feared psychometric tests and conceived them as dogmatic and that trying to measure 'cultural and social dimensions with quantitative measurements' out of context was unwise (Proppé 1983, p. 267). But ultimately the general agreement was that educational authorities should provide centralised written examinations because the old methods were considered too subjective and useless for comparison. Arason himself argued that it was time to provide opportunities for comparison between and within schools. The new methods came into use in most schools in the 1920s and in 1929 the first national tests were introduced. In the coming years and decades, centralised testing, though not standardised, earned its place as the mainstream way of assessing learning.

In 1946, the first law for one unified school system in Iceland was passed, *The Education Act of 1946*. Included were centralised examinations compulsory for grades 4, 6 and 8 and centralised entrance examination for grammar schools after grade 9, the National Examination (i. *Landspróf*). It was optional and at first very few students of each cohort passed it, 7% in 1950, 17% in 1965 and 25% in 1975.

Gradually, scholars began to worry about the negative influences that the ‘Landspróf’ had on the whole school system. Though it was originally meant as egalitarian means to secure equal rights for everyone, it gradually involved constricting effects with its emphasis on mere knowledge in traditional subjects. The focus was solely on book learning in subjects such as Icelandic, English, Geography, Mathematics and Physics. An eminent school administrator and educational advisor argued that schools should normally be organised bottom-up. But he asked if the academic emphasis and influence of the ‘landspróf’ had turned things upside down: ‘Are the schools not shaped top-down instead? Do learning conditions and organisation of secondary schools not indeed control what is done in primary and lower-secondary schools?’ (Gunnarsson 1963)

According to a new Act on the Comprehensive Primary School that came into action in 1974, the assessment discourse finally took new directions. As the following paragraph indicates, the discourse about assessment was gaining a different momentum:

Assessment of learning should not only be practiced at the end of a learning unit, rather it should be among the continuous activities of the school practice, entirely integrated with learning and teaching. The main purpose of assessment of learning is the motivation of students and learning assistance. (Law on the Comprehensive Primary School 1974)

And a pamphlet from the Ministry denoted:

Assessment has received increasing attention worldwide. At the same time focus on the nature and needs of individual students has increased and the learning process receives no less attention than the product of learning. (ME 1979, p. 3)

For two decades from 1970 to 1990, the pendulum swung ‘nervously’ from left to right, featuring an amalgamation of ideas rooted in cognitive and moral psychology (Jean Piaget and Lawrence Kohlberg), on the one hand, and, on the other hand, rational ideas rooted in behaviourist psychology (Ralph Tyler, Benjamin Bloom and Hilda Taba). For a whole decade, there lasted a sharp debate about public education. Finally, a new national curriculum was issued in 1989, featuring an intense learner-centred ideology and familiar pedagogical ideas from the progressive era. Thus, the 1989 curriculum featured what was then labelled as ‘the new progressivism’ (cf. Ravitch 1983). It was open-ended, advocating that boundaries between traditional subjects should be ‘blotted out’ (MEC 1989, p. 32) and that teaching, learning and assessment should reflect the idea of a ‘whole child development’.

The old criticism against centralised examinations thus continued in the 1980s and 1890s, not least because they had been conducted as norm-referenced from 1977 to 1983. These centrally governed examinations received the term ‘Samræmd próf’ and later on ‘Samræmd könnunarpróf’, where ‘samræmd’ means ‘coordinated’ or ‘centralised’. A system of relative grading was developed where the top 7% received A, 24% B, 38% C, 24% D and 7% received E. Because of entry requirements for secondary school, almost one-third of the student population received the message that they were not qualified for secondary education. The norm-referenced testing system was widely rejected by educators and was abolished, but it has in part

prevailed, although its interpretation and application have changed and the purpose is increasingly formative.

Despite a short back-to-basics period at the beginning of the new century focusing on detailed learning objectives and more centralised tests (MESC 1999), there have been no entry requirements for secondary schooling since 2002. Formative examinations (i. könnunarpróf) are first and foremost meant as supporting tools for teachers and their students providing information about strengths and weaknesses. In 2007, such formative examinations were finally presented by educational authorities as the only official assessment instruments to be used and since then no centralised achievement examinations have been used as high-stakes summative judgement about learning outcomes in Icelandic compulsory schools.

There was an increasing demand that enacted curricula in schools should receive increased attention with respect to assessment practices. Furthermore, it was suggested that teachers and schools should be responsible for both summative and formative assessment. Therefore, teachers should be provided with professional support to develop their assessment practices. Consequently, situated classroom assessment received increased attention and new conceptions began to emerge, such as authentic assessment, performance-based assessment, self-assessment, intrinsic motivation, metacognition, and last but not least, an ‘old wine in new bottle’ ‘feedback’. Additionally, new assessment tools were introduced, such as rubrics, rating scales and portfolios.

But surprisingly, a quite different perspective caught the attention of education authorities at the turn of the century. As the emphasis on classroom assessment was gaining momentum, Icelandic authorities decided to take part in large-scale international studies of achievement such as IEA’s first TIMSS study in 1995 and later OECD’s PISA programme. Generally, the results of these studies of achievement have indicated a declining trend regarding achievement of Icelandic students in literacy, mathematics and science. Furthermore, reports imply that there has been a fall in the number of Icelandic students at higher proficiency levels of PISA and a rise in the number of students at lower proficiency levels.

Since the current national curriculum came into force in 2011 (MESC 2014), teachers have become increasingly responsible for assessment:

Emphasis should be on formative assessment where pupils regularly consider their education with their teachers in order to attain their own educational goals and decide where to head. Criteria, on which the assessment is based, have to be absolutely clear to pupils. (MESC 2014, p. 26)

Furthermore, teachers have to cultivate a system of assessment criteria related to a scale (A, B+, B, C+, C, D) where A means exceptional competence, B stands for good competence, C for passable competence, and D for competence that does not reach the standard described in C. Most pupils are expected to have reached B or above by the end of compulsory education. And teachers are still reminded of their responsibility:

In the final assessment it is of fundamental importance that teachers ... make sure that the assessment is based on reliable data and that they use a variety of methods to acquire data,

in order to give pupils, their parents and the school as clear information as possible on the pupils' status. Thus teachers can gain better insight into the studies of each pupil. For an accurate conclusion, such as from conversations or on-site inspection, it may be relevant for teachers to cooperate when they consider the data that the assessment is based on and to use precise criteria. (MESC 2014, p. 92)

The importance of teacher collaboration as maintained in the last sentence above was certainly relevant and appropriate. It entails what has been called 'moderation', that is, systematic collaboration in organising learning, and benchmarking judgements about student achievement. Research indicates that sharing common knowledge about learning outcomes and levels of achievement enhances reliability, validity and fairness regarding achievement decisions (cf. Little et al. 2003).

Relevant Research Findings

Research findings confirm that since the current national curriculum came into force, teachers and schools do need professional support when assessing student learning, both regarding theoretical issues and praxis. According to some recent findings many interrelated issues are worthy of note. Four of them are reviewed here.

First, teachers seem to face difficulties when the issue is assessing the process of learning rather than assessing what has been taught (Sigþórsson 2008; Þórólfsson et al. 2011). In other words content coverage and assessment of what has been taught seem to receive more approval than assessing learning and what has been learned. As an example a majority of participants in Sigþórsson's study (2008) admitted they were typical transmitters of knowledge relying on school books and other written resources and accordingly assessed students' knowledge and skills. Science teachers in the same study observed that proper assessment of learning was problematic; most participants were convinced that they would practice different teaching and assessments if the system allowed it, and they...

... justified their way of teaching and how it differed from what they preferred primarily by the quantity of content that they had to cover and how it required teaching methods that enabled them to cover more content in a shorter time. (Sigþórsson 2008, p. 145)

Most intriguing was the fact that the science teachers maintained that there was not enough time and resources for hands-on learning and experiments (Sigþórsson 2008); class schedules did not allow such methods, which relates to the second issue.

The second issue concerns arranging proper conditions to assess complex and wide ranging competences such as critical thinking, problem solving, collaboration, and applying knowledge to new contexts:

The change means that now wide ranging competences need to be assessed, and how the pupil uses knowledge and skills, not merely how good he or she is at reciting facts and remembering things by heart. A lower-secondary school principal described the changes in this way: 'It's like changing a flat tire, you need to be able to execute it, not just recite orally how to do it.' (MESC 2016)

When teachers and administrators were interviewed about assessing how pupils applied knowledge and skills, there was an agreement that informal and authentic assessment was needed, though not always easy to implement:

We are not saying that they need to learn directly about Europe, Asia and for example rivers in Russia. Instead they need to show that they are able to read geographical maps and understand figures, graphs and tables about climate, vegetation, and such things. Thus assessment is more you know, we try to work with knowledge in class and the assessment is more about how they apply what they have hopefully learned previously. (Pétursdóttir 2018, interview with social science teacher)

The third issue of concern has to do with knowledge and skills regarding formative assessment. According to specialists such assessment is certainly not an easy job (Black and Wiliam 1998b; Leahy et al. 2005; Heritage 2010). Some teachers contend (Sigþórsson 2008) that it mainly involves regular testing during an ongoing course of instruction for the purpose of improving instruction, which is in fact a valid purpose. But formative assessment embodies a great deal of more complex teacher–student interactions and also student–student and teacher–teacher interactions. It features a process that takes place during learning and instruction where both students and teachers are active participants, ‘sharing learning goals and understanding how their learning is progressing, what next steps they need to take, and how to take them’ (Heritage 2010). Furthermore, it has to do with metacognition and pupils’ awareness and understanding of their own thinking.

Two Icelandic studies (Pálsdóttir 2006; Þórólfsson et al. 2011) suggest that formative assessment appears as more rhetorical than real praxis. Pálsdóttir’s study (2006) indicates that many schools lack clear strategies regarding assessment, especially formative assessment. Participants stated that in their schools there was a lot of discussion and work being done to develop assessment, and ‘self-assessment, portfolio assessment, and peer-assessment were considered useful assessment methods’ (p. 105) but they did not sense real emphasis on using them. Þórólfsson et al. (2011) found that discourse indicated focus on performance-based assessments, portfolios and authentic assessment, but ‘real practice seems to endorse an academic school curriculum to a considerable extent, setting standards for students and using tests as a motivation for pupils to learn the curriculum and teachers to teach it’ (p. 120).

The fourth issue concerns the transition from statistics and number grades to qualitative evaluation and letter grades. A key concept reflecting this transition is ‘competence’ referring to a wide range of cognitive, physical and attitudinal abilities that are supposed to be ‘evaluated’ by teachers not ‘measured’. Consequently, in addition to knowledge and practical skills, abilities such as solving problems and organising and interpreting information are to be assessed. Studies (Pétursdóttir 2018; Þórólfsson 2017) indicate that the time lag until the new system will gain full execution may become substantial. Of those responsible for the new system in their schools (mostly administrators) in school year 2016–2017, almost two-thirds agreed or strongly agreed that their schools were insufficiently prepared for matching assessed learning outcomes with the criteria based on letter grades as stipulated in the national curriculum (Þórólfsson 2017).

Discussion

In conclusion, this historical overview demonstrates that assessment in education is an enormous issue encompassing numerous important problems and questions that educators need to consider according to context. What is the purpose of the assessment? What should be assessed? How? By whom? When? Where? How will the results (data) be interpreted? How will the results be presented and used and for what purposes?

Central professionals that these questions weigh on are teachers, who need to be well informed regarding assessment, both theoretically and empirically. Teachers need to be familiar with research and theories and be prepared to discuss with parents, students, colleagues and other professionals about the different purposes of assessment and methodology. Furthermore, they are obliged to possess knowledge of basic concepts such as validity, reliability, criteria, relative grading, and norm-referenced versus domain-referenced evaluation systems. According to law and the current national curriculum, Icelandic teachers are most responsible for reliable and valid assessment so it concerns their professional identity.

As explained above, the pendulum has swung regularly from an emphasis on measuring learning outcomes (products) to assessing the process of learning. Education and assessment have in fact reflected an amalgamation of different ideologies. Michael Schiro (2008) identified four such ideologies, scholar academic ideology, social efficiency ideology, learner-centred ideology, and social reconstruction ideology. The emphasis on measuring learning outcomes relates more to the first two and an emphasis on learning and assessment as process relates more to the last two. But as Schiro (2008) indicated, all such ideologies represent ideals abstracted from reality, not reality itself. Hence, we may experience ideas that seem real parts of the enacted curriculum, but when observed closer turn out to be more rhetorical. According to recent research in Iceland this seems to apply to formative assessment in some instances.

International comparative studies of achievement such as TIMSS and PIRLS organised by the International Association for Educational Achievement (IEA) and PISA organised by OECD have an interesting role regarding such ideologies. As stated by Schiro (2008) the social efficiency ideology aims at providing knowledge that promotes the ability to function in society, viewing learning and teaching as a process by which behaviour is shaped, and assessment as a means to confirm how well they are prepared (shaped) to function as citizens. Learners are like raw materials to be shaped according to particular objectives.

By and large, PISA embodies similar ideology, that is, social efficiency. It examines not just what students know in science, reading and mathematics, but also what they can do in real life with what they have learned. Iceland has taken part in PISA since it started in 2000. Therefore, it must be essential to observe its role and influences, because PISA is not a typical academic research enterprise: 'It is meant to provide results to be used in the shaping of future policies ... PISA concepts, ideology, values and not least the results and the rankings, shape international educational

policies and also influence national policies in most of the participating countries' (Sjøberg 2007, p. 203). Svein Sjøberg (2007, 2018) has drawn attention to some debatable features of PISA, for instance, how results are statistically reported as simple ranking in league tables, drawing attention away from more significant factors and data. Sjøberg has also identified that a written test in science can hardly measure locally situated competencies, for example, those acquired on excursions, through inquiry learning, or in experimental work. His criticism also sheds light on problems related to reliability and validity:

... young learners in different countries and cultures may vary in the way they behave in the PISA test situation. I claim that in many modern societies, several students are unwilling to give their best performance if they find the PISA items long, unreadable, unrealistic and boring, in particular if bad test results have no negative consequence for them. (Sjøberg 2007, p. 203)

Finally, I want to re-emphasise the significance of teacher moderation. Systematic collaboration in organising learning and benchmarking judgements about student achievement is of most importance according to the current national curriculum. Networking teachers is bound to be beneficial, whether the issue is education ideologies, assessment policies, interpreting and using PISA data, or discussing assessment criteria related to wide-ranging learning outcomes and a new marking system featuring letters (A, B+, B, C+, C og D).

References

- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–73.
- Black, P. og Wiliam, D. (1998b). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–149.
- Gíslason, M. (1977). *Kvällsvaka: en isländsk kulturtradition belyst genom studier i bond-ebefolkningens vardagsliv och miljö under hälften af 1800-tallet och början af 1900-tallet*. Uppsala: Studia Ethnologica Uppsaliensia.
- Gunnarsson, K. J. (1963, November 13). Skólamál [School issues]. *Morgunblaðið*, pp. 15–16.
- Guttormsson, L. (1981). Läskundighet och folkbildning före folkskoleväsenet. *Nordisk kulturhistoria, Nordiska historikerötet i Jyväskylä i 1981*. Jyväskylä: Yliopisto.
- Guttormsson, L. (2008). Fræðsluhéðin: kirkjuleg heimafræðsla. In L. Guttormsson (Ed.), *Alþýðufræðsla á Íslandi 1880–2007* (Public education in Iceland 1880–2007) (pp. 21–35). Reykjavík: Háskólaútgáfan.
- Harlen, W. (2008). Teachers' summative practices and assessment for learning – Tensions and synergies. In W. Harlen (Ed.), *Student assessment and testing* (pp. 292–308). London: Sage.
- Heritage, M. (2010). *Formative assessment: Making it happen in the classroom*. Thousand Oaks: Corwin.
- Hjörvar, H. (1921). Nýjar prófaaðferðir (New examination methods). *Skólablaðið*, March 1921, pp. 30–32, and April 1921, pp. 45–46.
- Lagemann, E. C. (1989). The plural worlds of educational research. *History of Education Quarterly*, 29(2), 183–214.
- Law on the Comprehensive Primary School*, 1974. Lög um grunnskóla nr. 63, 1974. Reykjavík: The Althing.

- Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment that keeps learning on track minute-by-minute, day-by-day. *Educational Leadership*, 63(3), 19–24.
- Little, J. W., et al. (2003). Looking at student work for teacher learning, teacher community and school reform. *Phi Delta Kappan*, 85(3), 185–192.
- ME (Ministry of Education, Menntamálaráðuneytið). (1979). *Um námsmat (Bæklingur nr. 29)* (About assessment (Pamphlet nr. 29)). Reykjavík: Author.
- MEC (Ministry of Education and Culture, Menntamálaráðuneytið). (1989). *Aðalnámsskrá grunnskóla* (National curriculum for compulsory schools). Reykjavík: Author.
- MERC (Ministry of Education, Science and Culture, Menntamálaráðuneytið). (2016). *Nýir tímar: Aukin áhersla á hæfni og læsi í grunnskólum*. Reykjavík: Author.
- MESC (Ministry of Education, Science and Culture, Menntamálaráðuneytið). (1999). *Aðalnámsskrá grunnskóla* (National curriculum for compulsory schools). Reykjavík: Author.
- MESC (Ministry of Education, Science and Culture, Menntamálaráðuneytið). (2014). *National curriculum for compulsory schools* in English. Reykjavík: Author.
- Pálsdóttir, E. I. (2006). *Námsmat í höndum kennara* (Assessment in the hands of teachers). MED-dissertation. Akureyri: University of Akureyri.
- Pétursdóttir, D. R. (2018). *Námsmat á skilum skólastiga: Tilgangur, fyrirkomulag og nýtt matskerfi* (Assessment and the transition from compulsory education to secondary education: Purpose, nature and organisation of assessment between the two school levels). MED-dissertation. Reykjavík: University of Iceland.
- Proppé, Ó. J. (1983). *A dialectical perspective on evaluation as evolution: A critical view of assessment in Icelandic schools*. Reykjavík: Bóksala stúdenta.
- Ravitch, D. (1983). *The troubled crusade*. New York: Basic Books.
- Schiro, M. (2008). *Curriculum theory: Conflicting visions and enduring concerns*. Los Angeles: SAGE.
- Sigþórsson, R. (2008). *Mat í þágu náms eða nám í þágu mats: Samræmd próf, kennsluhugmyndir kennara, kennsla og nám í náttúrufræði og íslensku í fjórum íslenskum grunnskólum* (Assessment for learning or learning for assessment). PhD dissertation. Reykjavík: Kennaraháskóli Íslands.
- Sjøberg, S. (2007). PISA and “real life challenges”: Mission impossible? In S. T. Hopmann, G. Brinek, & M. Retzl (Eds.), *PISA zufolge PISA* (PISA according to PISA) (pp. 203–224). Berlin: LIT Verlag.
- Sjøberg, S. (2018). The power and paradoxes of PISA: Should inquiry-based science education be sacrificed to climb on the rankings? *NorDiNa*, 14(2), 186–202.
- Thorlacius, S. (1932). *Skólamál: Einkunnagjafir og fleira* (School issues: grading etc.) (pp. 1–24). Reykjavík: Prentsmiðjan Acta.
- Þórólfsson, M. Þ. (2017). *The new assessment system* (Report on changes of assessment practices in Icelandic compulsory schools (Draft)). Reykjavík: Bóksala stúdenta.
- Þórólfsson, M., Sigurgeirsson, I., & Karlsdóttir, J. (2011). Námsmat í náttúrufræði: Hvað má lesa úr skólanámsskrám grunnskóla? (Assessment in science education: what do school curricula emphasise?). *Uppeldi og menntun*, 20(1), 99–121.
- Walker, D. F., & Soltis, J. F. (2009). *Curriculum and aims*. New York: Teachers College Press.