Nationally co-ordinated tests in Iceland: A critical view  
Rúnar Sigþórsson

Good morning, and thanks to Nordisk Komité for inviting me here. It is both an honour and a pleasure to get the opportunity to address this audience of Nordic parent representatives.

My hometown, Akureyri on the north coast of Iceland which you can see on that picture, started to build as a trading centre in the early decades of the 19<sup>th</sup> century. In its early times it used to be known for substantial Danish influence, resulting from a number of Danish tradesmen who had their businesses there. In order to live up the Danish, so the folklore goes, the upper class of the local inhabitants spoke Danish on Sundays. Today is only Saturday, Sunday is not until tomorrow, and therefore I should really be speaking Icelandic today, but to save you from that I will compromise on speaking English instead and I do hope you don’t mind that.

What I want to speak about today is the centralised, external assessment or evaluation in schools often referred to as Standardised or National testing. In particular I want to put that kind of testing into context with the broader picture of teaching and learning and examine the consequences of such testing. I draw mainly on literature and research from the USA and the UK and take examples from my own research on the National testing system in the Icelandic compulsory school – Grunnskólinn.

The Icelandic context

To brief you shortly on the Icelandic context Grunnskólinn is compulsory for all children from 6 to 16 and is entirely the responsibility of the local communities since 1996. It is divided into ten grades and has two levels: Primary level with grades 1–7 and lower secondary level with grades 8–10. There is, however, little distinction between the levels and most often students finish their compulsory education in the same school. There are however two other possible types of schools: Schools with grades 1–7 only, and schools with grades 8–10 only.

The compulsory school is governed by a legislation (the compulsory school act) and a National Curriculum. The current National curriculum was published 1999. It is highly prescriptive in terms of learning aims and objectives but gives schools more autonomy over the implementation – the ways aims and objectives are met. The curriculum comprises around 900 pages in 12 separate booklets; one providing general guidelines, and the remaining 11 for individual subjects or in some cases clusters of subjects.
The NC states the intended outcomes of learning in three different ways: First there are the final outcomes of each subject, stated as rather broad aims intended to be met at the end of Grunnskólinn. Second there are two sets of intermediate objectives to be met by the end of grade 4 and grade 7. And finally there are learning targets for each year (Menntamálaráðuneytið, 1999).

Assessment is both carried out with a system of National testing and by individual teachers and schools. Schools are required to have formal assessment procedures where results are formally fed back to students and parents two or three times during the school year. At the primary level diverse assessment methods are applied but as students grow older these give way to a more formal quantitative assessment in the form of written exams. When it comes to the lower secondary level, assessment is mainly quantitative, based on written examinations in every subject at least twice during the school year. The results of these exams, in the form of numbered grades, are the backbone of what is fed back to students and parents, although other factors, such as homework and general performance and conduct are usually taken into account to some extent.

The National examinations are held in two subjects, Icelandic and Mathematics in grades 4 and 7, and in six subjects in grade 10. These subjects are: Icelandic, Maths, English, Danish, Social studies and Natural sciences. These examinations are optional, at least in theory, but of course hold the key to enter the upper secondary schools. Virtually every student takes the examination in Icelandic, Mathematics and English, but the participation in other tests varies between 50 and almost a 100%.

The conduct of the tests follows a strict procedure. They are composed, marked and organised by a central government organisation, the Educational Testing Institute, and are held at the same time all over the country under the supervision of officials appointed by the Ministry of Education. However the tests are not standardised – even if they have many of the same characteristics and serve the same purposes. They could therefore perhaps be referred to as semi-standardised, but here I refer to them as Nationally co-coordinated tests or simply as National tests.

The current system of NT in grade 10 has been in effect since 1977, but with some modifications. Most of that time four examinations were mandatory for all students (even though exceptions were possible for special education students). But since 2003 the tests are held in six subjects and are optional. The testing in grades 4 and 7 was added in 1996.

**Standardised tests and the teaching and learning process**

I think it is now time to change the focus from the system to what goes on within the system and ask: What are the purposes of National testing? Who wants it and who needs it and what for? How do National tests fit into our knowledge of effective teaching and learning – and what are their effects on the teaching and learning we want to take place in schools. In what remains of my presentation I will address those questions.
The model

First of all I would like to present you with a model of the teaching and learning process to give you an idea of the context within which the National testing is operated.

The model is built on the premise that what really matters in schools is the nature and outcomes of learning (Fullan, 2001; Hopkins, Ainscow and West, 1994) and that, therefore, has to be the core of the model. I want to make it very clear from the outset that I use outcomes here in a broad sense comprising of course subject knowledge, but also interest, creativity, problem solving skills, values and character assets; in short what we value most in our children and people around us. Needless to say perhaps, these outcomes must be for ALL students so that every individual is challenged with meaningful learning experiences that lead to success and prepare for the future.

The model then has two sides: One relates to students and their conditions for learning and the other relates to teachers. On the student side the first condition is what Carol Ann Tomlinson calls student traits or individual characteristics to which teachers need to respond in their teaching (Tomlinson, 1999; Tomlinson and Eidson, 2003). These are in short their readiness: what students bring with them to the learning process and how ready and well prepared each student is to engage in what the teacher is planning for him; their interest: what each each student enjoys learning, thinking about and doing; and thirdly students’ learning profile: in what way each student likes to engage in the learning process.

The second condition relates to the nature of the learning tasks and the learning experiences the students are given and the third relates to their motivation and self-confidence. I will come back to these later on.

The teachers’ side the model rests on three underlying forces that shape their teaching practices. These are: teachers conceptions about teaching; their pedagogical and content knowledge and their self-efficacy beliefs.

Teachers conceptions relate to teachers’ own beliefs and ideas about themselves as teachers, the nature of teaching and learning, the purposes of learning and what should be the outcomes of learning (Brown, 2002).

Pedagogical content knowledge is the blend of how well teachers know the content of their subjects and how good they are at teaching it to the intended outcomes (Guðmundsdóttir, 1991). This is obviously a vital precondition for good teaching because not enough to know the subject – it takes a lot more to be a good teacher and it is my guess...
that if you were to choose between the subject specialist and the good pedagogue for your children you would not be in doubt.

The third factor important here is the teachers self-efficacy beliefs. These can be defined as their conviction or belief in their own ability to influence how well students learn or perform which again is an important precondition for good teaching (Bandura, 1994; Goddard, Hoy and Hoy, 2004).

I suggest here that these three factors are the most powerful determinants of teachers planning and actions in three areas: curriculum and content, how the actual teaching is carried out and, last but not least, assessment. It needs to be emphasised here that there must be harmony between these three areas and that they directly affect each other. In particular it should be noted that assessment tends to have direct influence on both the curriculum and the actual teaching. This is to say that teachers have a strong tendency to teach what they think, or know, will be tested. And the converse of that: they tend NOT to teach what they think, or know, will NOT be tested (Linn and Miller, 2005).

It is also clear that teaching is rooted in government mandates and official documents like the NC. It is clear however that teachers stand between these mandates and what happens in schools. Therefore official decisions about education are filtered through teachers’ conceptions, their pedagogical content knowledge and their self-efficacy beliefs before they are implemented in the classroom.

With the model I want to make clear that assessment, in whatever form it is, is one bit of a complex jigsaw of teaching and learning; that it has direct impact on all the other bits and therefore on the nature and outcomes of students’ learning. And this is why we have to think about assessment in context with all the other parts and the way we think about assessment has to fit in to what we think about those parts.

**The purposes of NT**

There is no time here to address the model in details so instead of attempting that I will highlight just a few important issues.

First of all I want to get back to the question I asked earlier: What are the purposes of National testing? Who wants it, who needs is and what for? From the Icelandic point of view it is quite clear that the grade ten examinations serve mainly accountability purposes. They go hand in hand with the detailed NC I described earlier, that is devised to ensure that every child is given the opportunity to receive the same kind and same amount of education at the same time. Their purpose is mainly to provide both national and local educational authorities with information about how well students and schools perform on the standardised learning outcomes set in the NC, and hold schools, teachers, and in a sense students themselves, accountable for their performance.

In addition to that the results of the NT provide the upper secondary schools with information about students’ performance they use to regulate their intake. Then of course the results are a feedback to parents and to students themselves.
The National testing is an example of assessment OF learning where outcomes are tested at the end of the learning process, as opposed to assessment FOR learning where data on student learning is gathered in order to improve teaching and learning and the internal learning conditions in schools (Barber and Fullan, 2005). It is noteworthy that in the regulations about these tests in my country, there is very little reference to the professional accountability of schools and authorities to make use of the test results for improvement purposes.

The grade four and grade seven tests serve the same accountability purposes but additionally they have a professional accountability feature to provide teachers and schools with information about pupils performance that can be used to respond better to their educational needs.

The underlying assumption here seems to be that by centralised decisions about curriculum and assessment and by holding teachers and schools accountable for the results it is possible to improve schools and ensure equity in the school system; that all students have the same opportunities to receive the same kind and amount of education – that no child is left behind as the Americans put it.

This, however, is questionable to say the least. First of all the very idea that it is possible to devise and implement a curriculum where all children receive the same kind and amount of learning at the same time is a simplistic assumption and in fact at odds with everything that modern theories of learning tell us; and, I would argue, at odds with common sense as well. Just reflect on your own children and see if you can agree with that.

Secondly there is very little evidence from the school improvement research that government mandates about curriculum and external accountability are effective ways of improving schools. On the contrary the success of school improvement efforts is dependent of internal conditions in schools, such as the initiative, empowerment and commitment of teachers; factors that lie beyond the reach of educational authorities (Hopkins and Lagerweij, 1996).

Turning back to the model, let us also come back to the questions I asked earlier: How does National testing fit into what we know best of effective teaching and learning – and what are their effects on the teaching and learning we want to take place in schools. To address these questions I will now highlight four components of the model: These are: Assessment, curriculum and content, student traits and at last motivation and self-confidence.

**Assessment and the curriculum**

I have already argued that National testing is an example of assessment OF learning. It measures the end product of the learning process, rather than assessing the learning process as it unfolds in order to assist students and make their learning experiences more meaningful. This should indeed be the main purpose of assessment, and this is emphasised time and again in the Icelandic National curriculum (Menntamálaráðuneytið, 1999). Secondly, the tests have what is often called high-stakes nature. This means that there is something at stake for everyone involved and the students face the risk of failure or something unfavourable to happen if they do not get the desired results (Kohn, 1999).
high-stakes nature of the Icelandic National testing is further enhanced by making the results public. Third, the purpose of the tests is comparison. They are not intended to reveal to what extent John or Jane met their own learning criteria. They are intend to reveal where John and Jane, or their class, stand in comparison with the year cohort at the national level. In that sense the tests are competitive in nature. Last but not least Standardised tests in general are almost entirely multiple-choice tests. And they are fine when we want to assess how well students have memorised factual information where there is only one possible answer. And sometimes we may want to do just that. On the other hand when there is more than one possible answer, when we want to test initiative, creativity, ability to synthesise, evaluate and solve problems and when we want to test written and oral expression multiple-choice tests are useless (Popham, 2001). Bearing in mind what I said earlier, that we get what we test and DO NOT get what we DO NOT test, the effects of multiple-choice testing on teaching and learning is obvious. It will potentially direct teaching and learning towards testable information but away from higher order thinking, problem solving and deep learning where students can construct their own knowledge and deep understanding. And I think everybody has to ask themselves: authorities, teachers and not least you as parents: It that what we want? Is that what produces the learning outcomes we desire?

I will just give you one example from my own research to highlight this. The Icelandic Science curriculum has been criticised by teachers for being too extensive to implement within the time allocation for the subject. As a result teachers have to pick and choose from the content areas of the curriculum. For the test to make sense in that situation the choice of curriculum content has to be centralised. This has, at least up to now, been done by so called content tables issued by the National testing institute where the content of the curriculum is classified and ranked in an order of priority. To take an example learning about the infrastructure of cells has high priority; the human body, sex education an family planning has middle priority and the ocean around Iceland and local ecosystems have low priority. And you will of course guess what happens in the teaching – I don’t think I need to tell you that. The textbooks are also a comprehensive packet of six successive books to be covered in grades 8–10, some 900–1000 pages, and tested by the end of grade 10. This leads to an overwhelming emphasis on textbook work and textbook-related exercises at the expense of hands on work, experiments and field work that takes longer time and is not regarded by teachers as effective test preparation.

**Student traits**

If we look at standardised testing, in the light of student traits a peculiar paradox becomes evident. On the surface such tests are objective and fair in the bureaucratic sense of treating everybody equally, but this very characteristic of standardised tests makes it difficult or even impossible to respond to individual differences. In practice this means that treating people humanly comes second after treating them equally (Proppé, 1983). At present there is a growing concern, not least in the USA and the UK, that standardised testing linked with standardised curricula may not be the right way to ensure fairness and equity in the school system. It is argued that this may in fact be detrimental to the nature
and outcomes of education, particularly for students who are at risk in the mainstream school system for reasons like learning difficulties, low social and economic status and ethnic origin (Ryan and Cooper, 2004). If that is the case, schools could be adding to and maintaining discrimination within themselves that in the long run may well feed into the society at large.

Two well known American educationalists Lorrie Shepard and Linda Darling-Hammond argue that the under the pressure of high-stakes testing schools are tempted to reject more students with learning difficulties or send them to discriminating special provisions inside or outside the school (Darling-Hammond, 1997; Shepard, 2000). They also argue that high-stakes tests encourage schools and teachers to react to learning difficulties in an ineffective way, by making students repeat the so called basics they have not mastered, do more of the same, instead of adjusting the teaching and learning process to their needs. By that, they argue, learning difficulties are made the problem of the student, not the responsibility of the school and they question the validity of making students repeat a learning experiences that did not suit them in the first place. Shepard (ibid.) goes as far as to argue that such teaching strategies are not only ineffective, but may even CAUSE learning difficulties by requiring students to learn facts and gather information without understanding and out of context with the broader picture of reality and meaningful tasks.

In my own research in Iceland I have found a clear indication of that where optional subjects in grades 9 and 10, that are meant to expand the possibilities of students to pursue their own interests, are taken into the service of the National tests to make struggling students repeat yet again the basics of Icelandic and Maths.

**Motivation and self-confidence**

I have already made the distinction between assessment OF learning and assessment FOR learning. The former is by nature separated from the learning process and comes as a grand finale when all tasks have been completed. For some students this may no doubt be rewarding, but only if it confirms their belief that they are doing well and their effort will lead to success.

But not all students are winners. In every class there is a number of students who are struggling with what the school demands of them. For those students some struggles are lost while others are won and every won and every lost struggle will impact their motivation and self-image in a different way. But we have to face that there is also a number of students with a continuous history of failure. For those students assessment OF learning is generally yet another occasion where they fail to keep up with their peers, and the devil takes the hindmost. For those students the more the pressure is, the higher the bar of official standards is raised, and the higher the stakes are, the more is their risk of failure; the less are their chances to succeed, and the more is the likelihood that they simply give up (Stiggins, 2004). For those students assessment OF learning is neither motivating nor stimulating; it is discouraging and defeating.

I can still hear the voice of the thirteen year old girl who was speaking to me with a group of her classmates and telling me that she dreaded the upcoming national tests in Maths and Icelandic. And when I asked her why she was so anxious she whispered so quietly that I
had to ask her to repeat: ‘Because I am so pathetic’. That was what six years of compulsory schooling had taught her about herself.

There is a mistaken assumption around, that standardised testing will motivate students to learn and by intensifying the pressure on students to succeed we can cause them to try even harder and do even more. I say a mistaken assumption because I believe that for all students, and in particular low achieving students, the motivation lies in the teaching – not in the testing, – and in the possibility to experience success. It was therefore a quite worrying observation of a group of grade 10 students (although I admired them for their analysis!) that their teachers did not bother to make their teaching interesting and motivating because they relied on the pressure of the National tests.

So if we decide to have National tests we should not disguise them as a motivator for students. Let us instead admit that they are not; that they are simply a means of collecting data about the performance of the school system. And if we do have a National testing system there is an absolute demand that it goes parallel with a system of continuous assessment FOR learning at the classroom level, that supports student learning and helps them experience more success than failure.

Few have captured this better than the British assessment specialists Black og William when they say:

> What is needed is a culture of success, backed by a belief that all can achieve. Formative assessment (Assessment for learning) can be a powerful weapon here if it is communicated in the right way. Whilst it can help all pupils, it gives particularly good results with low achievers where it concentrates on specific problems with their work, and gives them both a clear understanding of what is wrong and achievable targets for putting it right. Pupils can accept and work with such messages, provided they are not clouded with overtones about ability, competition and comparison with others ... Feedback to any pupil should be about the particular qualities of his or her work with advise on what he or she can do to improve, and should avoid comparisons with other pupils. (Black and William, 2001)

### Some final words

To conclude here, I have argued that National or Standardised testing, as a means of assessment, have to be seen as a part of a broader picture of teaching and learning and that they will have certain impact on the learning process, as will all other forms of assessment. I have used my time to highlight some issues derived from a model of teaching an learning to underpin that argument. I want to make it quite clear that I am NOT suggesting that official data about the performance of the school system should not be collected. NOR am I suggesting that teachers and schools should not be held accountable for what they have been trusted to do. On the contrary I think that should be done. National tests may well play a part in data gathering about the school system but we can not rely on them alone for that purpose. Our data must be far more comprehensive than that, and we must learn to use the data we gather constructively to inform improvement. If we, however, decide to use National tests we must be aware of their shortcomings and consequences. The high-stakes pressure must be lifted off students and the tests should not be disguised as a means of...
improvement or motivation when they are not. Most of all I am arguing, though, that assessment of learning both at the national and school level needs to go hand in hand with continuous assessment for learning at the classroom level that is intended to help students to enrich their learning experiences and improve their learning outcomes; experience victories instead of defeats; success instead of failure.

References

