Chapter 3

Student assessment

There is a strong focus on summative student assessment in the Slovak Republic. The reliability of final examinations at the upper secondary level (Maturita) has been strengthened with the introduction of an external written component. National assessments are administered at the end of lower secondary schooling and mainly serve the purpose to certify student achievement and help inform selection of upper secondary school. A competency-based curriculum has been introduced since 2008 and implementation has been problematic due to a lack of clarity on the minimum content standards and a lack of guidance, tools and training for educators. In this context, there is room to give more support to teachers, to develop models to assess skills and competencies and to phase in student assessment criteria linked to the standards in the national education programmes.
This chapter focuses on approaches to student assessment within the overall evaluation and assessment framework in the Slovak Republic. Student assessment refers to processes in which evidence of learning is collected in a planned and systematic way in order to make a judgement about student learning (Harlen and Deakin Crick, 2002). This chapter looks at both summative assessment (assessment of learning) and formative assessment (assessment for learning) of students.

**Context and features**

This section sets out the major features and developments in the approach to student assessment in Slovak schools. Overall, while there is a strong commitment among teachers to regular assessment, the methods remain markedly traditional with a predominantly summative rather than formative purpose. There are often significant variations in standards across institutions.

**The impact of educational reform on student assessment**

The School Act (2008) was the most far-reaching attempt to reform the education system since the Slovak Republic became independent in 1993. It has had significant consequences for the curriculum, assessment and organisation of education that many schools and teachers are still getting used to.

One of the most significant changes in the Slovak assessment landscape was the foundation of the National Institute for Certified Educational Measurements (NÚCEM) by the Slovak Ministry of Education. The NÚCEM’s main role is to develop, administer and oversee all national tests and assessments, including tests in basic schools (currently the Testovanie 9) and the examinations at the end of secondary schooling (the Maturita) – specifically the external examination and the written elements of internal assessment. In addition, the NÚCEM has responsibility for liaising with international assessment programmes such as the OECD’s PISA, research and development of methods of measuring and evaluating the quality of education, and ensuring standards and monitoring outcomes.

**The development of national performance standards for student learning**

Student assessment is based on age related educational performance standards for each degree of education that form part of the statutory national education programmes. These standards specify both the minimum content that students should be taught and the minimum standard of performance they are expected to reach in competencies related to knowledge, skills and abilities. Since their introduction in 2008, the National Institute for Education (NIE) has continually revised the standards in response to criticism that the initial standards were vague, unhelpful and did not relate to any established taxonomies of learning (NÚCEM, 2012, p. 85).

**On-going teacher assessment in basic schools**

During the primary phase, on-going teacher assessment combines the formal marking of work, regular oral or written tests, and oral assessment (all marked to a five point scale). The latter is supplemented by verbal feedback to the student and is most common in the early years of basic schools. Formal summative reporting of assessment happens at the end of the first and second semesters during the school year, and results in a certificate for individual subjects. Students are graded for both academic achievement (to
a five point scale) and behaviour (to a four point scale). Additional certificates are also issued on the successful completion of the primary (ages 6-10) and lower secondary (ages 10-15) phases in basic schooling. The school education programme specifies which subjects are marked and how.

A student needs to pass the end of each grade to progress to the next phase of education. If a student fails in a compulsory subject, they have the opportunity to be re-examined during the summer holiday. If the re-sit is failed they then must repeat the grade, though data from PISA suggests this is relatively rare. In 2009, 3.8% of 15-year-old students in the Slovak Republic reported that they had repeated a grade at some stage during primary or secondary schooling, in comparison with the OECD average of 13% (OECD, 2010).

The introduction of national tests in lower secondary education (Testovanie 9)

Since 2003, the Slovak Republic has participated in the PISA international comparative tests. In all three cycles, Slovak students performed significantly below the average for OECD countries in reading literacy, with results for mathematics fluctuating between average and below average (see Chapter 1). Partly in response to these results, national external assessments in mathematics and the language of instruction (Slovak, Hungarian or Ukrainian) – known as Testovanie 9 or T9 – were introduced in 2009 at the end of lower secondary education (ISCED 2). These replaced the previous Monitor 9 tests that were administered from 2004/05, which in turn had replaced a national sample-based test.

The Testovanie 9 tests are summative in purpose as they provide a measure of student performance at the end of one phase of their education and inform the choice of secondary school, notably whether the student would go to an academic “gymnázium” or vocational secondary school. The tests are norm referenced and students receive a percentage mark and a figure showing where their performance sits in the national performance distribution. Testovanie 9 is seen as a way of raising the profile of the types of skills and competencies students need to address the shortcomings revealed in PISA. The implementation of these tests was supported by sample materials and marking schemes and regional seminars for teachers, school leaders, university professors and policy makers.

While individual student results are not published, aggregate Testovanie 9 results are published online in regional tables containing the percentage of marks correct for Slovak language and literature and mathematics over the past three years and where each school sits in the national performance distribution (a national percentile for each school). Information is also provided showing how far a school’s performance relates to the national average.

In addition, the NÚCEM publishes an annual report summarising key findings from the T9 results including: statistical information; comparative data; an analysis of common errors made by students; and recommendations for teaching and learning.

On-going teacher assessment in secondary schools

On-going teacher assessment in secondary schools follows a similar pattern to in basic schooling, consisting of formal marking and regular oral and written tests graded on a five point scale. Based on these grades, students are issued with certificates to confirm that they have successfully completed the course at the end of first and second semester
(mid-year and end of year). Again, students can re-sit examinations, but have to repeat the
grade if they fail a second time. Assessment is often heavily influenced by text books,
many of which contain tests. This has recently become problematic in that new text books
focused on competencies have not been well received, leading some teachers to continue
to use text books – and consequently tests – up to ten years old that do not reflect the new
curriculum.

The Maturita and curriculum reform

The compulsory Maturita examinations mark the end of upper secondary schooling
(ISCED 3) and act as a means of determining university entrance. There are compulsory
examinations in Slovak language and literature and a foreign language, and these consist
of an internal and an external component. Students are also examined in two elective
subjects via an oral examination within the school. Since 2011, the Maturita has only one
level, with the exception of foreign languages, which has two levels linked to the
Common European Framework for languages.

Examinations in the language of instruction (Slovak, Hungarian or Ukrainian) and
foreign languages include an internal component that is split into an oral examination and
a written examination. The oral examination is set by the school and can take a number of
forms such as presenting a complex task, a performance, the defence of a project or a
combination of these. This part is assessed by an examination panel including a member
of the subject commission from another school with the chair appointed by the regional
school authority. The written part of the internal component is based on a topic set by the
NÚCEM. This is then internally marked by two teachers using a common marking
scheme. This scheme aims to create transparency by providing teachers with clear criteria
for marking student work. The NÚCEM has received feedback from many teachers that
the marking scheme does not reflect all the necessary aspects to be assessed. As such, the
NÚCEM has made attempts to review the common marking schemes for all foreign
languages, as well as for the languages of instruction. In the case that the two teachers
marking the internal component cannot agree on the student marks, an arbitrator must
help them decide, this person is typically the Head of the subject at the school. Teachers
may also contact the NÚCEM directly to request advice on aligning their marking to the
common scheme. There are no further processes for ensuring that marks or grades are
awarded appropriately and consistently.

The external component takes the form of a written test that is devised and
administered by the NÚCEM. Over time the external component of the Maturita has
gradually increased in significance, reflecting the fact that in the Slovak Republic and
other countries in Eastern Europe there has not been a tradition of external assessment at
the end of secondary schooling.

From 2011 the Maturita examination was reformed to better reflect the revised
curriculum by focusing more on developing skills and competencies. However, hurried
implementation led to many students complaining of a mismatch between what their
teachers had prepared them for and what appeared in the examination. This however
appears to have been an issue specific to that year and the OECD review team formed the
impression that curriculum and examinations were now once more synchronised.

Internal school structures for student assessment

Other than the compulsory national tests and examinations at the end of basic
schooling (ISCED 2) and upper secondary schooling (ISCED 3), schools determine their
own approaches to student assessment based on Ministry guidelines. The School Act stipulates that school leaders are responsible for determining the method of assessment in individual subjects after negotiation with the school pedagogical board, one of the internal advisory structures available to school leaders. The pedagogical board evaluates the overall summative student assessment procedures in place in the school. Subject commissions, groups of teachers collaborating on specific subjects, e.g. mathematics, can also be used for this purpose. Individual teachers are responsible for the continual assessment of students. While these varying methods can lead to a wide range of different approaches to assessment across schools, there is some evidence of greater consistency of practice being developed through the co-operation of subject commissions across schools (see below).

Developing future national assessments

Building on the current national testing model, the NÚCEM is preparing new tests for the end of primary schooling (ISCED 1) to measure pupil attainment at the end of this phase of education and provide additional information to schools (see below). These new tests are referred to as Testovanie 5.

Strengths

Establishing the NÚCEM has led to reforms and innovation in student assessment

Since its founding in 2008, the NÚCEM has played an important role in developing and reforming assessment in the Slovak Republic. For example, at the time of the OECD review, the structure of the Maturita qualification was imbalanced in relation to the weighting of both the written and oral components of the internal assessment and the internal and external assessments as a whole. This had led to questions about the validity and reliability of the qualification. Specifically, in the Maturita internal examination, a student needed only to obtain a grade 4 (one grade above a fail) to pass the oral examination. However, if the student exceeded this and scored a grade 3 or higher in the oral examination, they would pass the internal component regardless of the grade they achieved in the written examination. Subsequent to the OECD review visit, the NÚCEM addressed this anomaly by increasing the weighting of the written component within the internal examination so that a student would have to succeed in both elements in order to pass the examination overall. Furthermore, as of September 2012, students who fail to get 33% or higher on the external component will not be able to pass the Maturita.

The NÚCEM has also brought an international perspective to student assessment and has introduced changes to the Maturita such as including reading comprehension questions based on the PISA model, for example, to bring it more into line with other European assessments and better address aspects of the revised curriculum.

In addition to reforming existing assessments, the NÚCEM has had an impact in developing new qualifications that have the potential to develop assessment practice in the Slovak Republic; for example, Maturita examinations in new subjects are currently being piloted. During the OECD review, schools appeared to view this development positively, seeing it as a way of increasing the status of these subjects and bringing greater objectivity to grading. The new Testovanie 5 assessments currently being prepared by the NÚCEM will also enable value added measures to be introduced by comparing student performance at the end of ISCED 1 with the T9 results at the end of
ISCED 2. Again, there were strong indications that teachers welcomed this and felt it would give a fairer and more realistic picture of student progression.

As well as developing the content and structure of qualifications, the NÚCEM, working closely with other organisations such as the Institute for Information and Prognoses of Education (ÚIPŠ), is using technology to introduce innovative new ways of delivering qualifications. The Maturita online project is currently trialling the use of electronic testing as a way of increasing the objectivity of external assessments and improving feedback to teachers. Early responses have been extremely favourable with 86% of students involved in the trial saying they would recommend it to their classmates and 85% of co-ordinators and administrators expressing a preference for online delivery over a paper-based equivalent.1

The NÚCEM has also utilised online delivery to provide feedback to schools. Analyses of tests administered by the NÚCEM, including score distribution and differences according to the school type, region and gender, are published on the NÚCEM website. The ranking of schools is also provided in the form of an interactive map. More detailed reports providing an analysis of test items or curriculum areas that students found difficult are sent directly to school leaders in order to inform decisions about curriculum development and areas requiring greater coverage (NÚCEM, 2012, p. 90).

**Improvement in supervision for external testing**

From 2011 the NÚCEM has introduced more rigorous administration procedures to improve the reliability of internal assessment and address continuing concerns about the potential for malpractice. A teacher from a different school is now present as an observer during T9 tests and a parallel procedure has been introduced in secondary schools where the chair of the subject commission from a different school is part of the panel for the internal component of the Maturita to ensure objectivity. They have the authority to stop an assessment if they have any doubts about the way it is being conducted. There is evidence that in 2011, the number of schools investigated for malpractice decreased significantly as a result (Polgáryová, Kurajová Stopková, and Kubiš, 2011, p. 4).

The NÚCEM also works closely with the Slovak State Schools Inspectorate (ŠŠI) to identify and discourage malpractice in high stakes assessment, i.e. assessment with important consequences for students, in both basic and secondary schools. Inspectors observe the conduct of the internal component of the Maturita in approximately forty schools nationally. While this is a relatively small percentage, the visits are targeted at schools where there are suspicions of irregularities. A similar process happens for T9 testing, although in this case about 150 schools are chosen at random and visited unannounced.

**Examples of emerging good practice in internal teacher assessment**

While contextual issues, such as the lack of national criteria to inform on-going teacher judgements, have to some extent impeded the development of reliable models for on-going assessment, the OECD review team found emerging evidence of good assessment practice in many of the schools visited.

In several cases schools provided opportunities for subject teachers to meet with other colleagues in the region to discuss best practice in assessment, and although these meetings often focused mainly on sharing tests they had devised, in the longer term it
provides an excellent model for schools to collaborate on developing shared criteria and standards. An important starting point in developing teachers’ confidence and expertise in assessment is to allow time for them to meet, discuss judgements they have made on student’s work, and develop a shared understanding of what evidence of attainment at a particular level looks like.

There was good evidence of schools exploring ways to develop consistency in their assessment practice. Many schools visited reported using their subject commissions to set or validate tests, which provided a degree of internal comparability and coherence in their assessment activities. Evidence of using subject commissions to improve the consistency of assessment judgements through internal standardisation processes was present, but less common. Positively however, all relevant teachers have received training in this area, and this is seen as a long term process that will require continuing development.

A consistent feature of the OECD review was that both students and parents were extremely positive about the way in which assessment outcomes are shared with them. Parents especially felt that teachers were accessible and willing to discuss progress if they had any concerns, with many schools operating an “open door” policy. Parents also formally meet with class teachers twice a semester where they can discuss attendance, behaviour, grades progress and receive advice on how their children could improve.

**Emerging recognition of the potential that technology can bring to student assessment**

There appears to be an emerging interest among schools in the use of technology to support student assessment. The OECD review team saw evidence of a commercial, web-based assessment tool, the basic version of which has been distributed free to schools, being used to share assessment data online with parents. The tool allows teachers to input marks and grades from on-going assessments and use these to record and track student progress. The tool also has the capacity to be used diagnostically, producing graphics that provide analysis and comparisons. Though this functionality seemed to be less widely used, it has the potential to give valuable feedback on students’ strengths and weaknesses that could be used formatively, as well as to inform planning.

During the OECD review, the Institute for Information and Prognoses of Education (ÚIPŠ) presented its plans to develop a new information system to record student information electronically. The aim is to reduce bureaucracy by allowing all relevant bodies and organisations to access the same database that includes assessment information for each semester.

**Challenges**

**Assessment is not sufficiently formative in purpose**

Although formative assessment “does not have a tightly defined and widely accepted meaning” (Black and Wiliam, 1998a, introduction), it can broadly be defined as on-going assessment whose main purpose is to recognise and respond to student learning while it is happening to give feedback that will help the student learn and improve (Cowie and Bell, 1999, p. 101–116.) In contrast, summative assessment seeks to grade or define the level of attainment a student has achieved at a particular point more formally (most often at the end of a unit of work, school year or phase of education). In short, summative can be characterised as assessment of learning, whereas formative represents assessment for learning.
The OECD review team formed the impression that the idea of formative assessment was not well understood by teachers, students and parents. While schools clearly recognise the importance of regularly assessing their students, with several using “input and output tests” at the start and end of a unit of work or topic to measure student progress, it did not seem clear that the results of these tests were being used in a formative way. The predominant culture appeared rather to be the use of regular and frequent summative assessment (such as weekly quizzes followed by more formal tests at the end of a topic) with the main purpose to provide evidence towards students’ grades at the end of each semester. This frequent testing of small discrete areas of knowledge can lead to a fragmented picture of a student’s overall progress and give neither the student nor teacher a clear sense of their overall progress. Interviews with students indicated that feedback was mainly limited to marks, brief comments, or a discussion of which questions they got wrong, focusing more on what the student needed to improve rather than how they could achieve this improvement. However, such feedback “is not formative unless the interaction is designed to help students to learn…marks, levels, judgemental comments or the setting of targets cannot, on their own, be formative” (Mansell and James, 2009, p.10). Further, there were some concerns raised by older students that teacher feedback could be based on the teacher’s perception of the student’s preparation and effort as any meaningful analysis of what their next steps should be to progress. Internal evaluations of national assessment results in mathematics and grades assigned to students in mathematics have certainly highlighted a marked degree of subjectivity in teacher assessment (NÚCEM, 2012, p. 89).

**Tension between the shift to skills and competencies in curriculum and an assessment model that is better suited to testing knowledge**

When there are significant shifts in the curriculum, assessment and pedagogy also need to be refocused so that the reform is “joined up”. In this context, the dominance of testing in the Slovak system is particularly problematic as the revised curriculum emphasises competencies such as tolerance and independence, which are difficult, if not impossible, to assess using traditional testing methods. Implementation in schools has lagged behind the vision of the new curriculum, with many schools and teacher organisations mentioning the rapid pace of reform and a lack of guidance on how to assess competencies as a considerable barrier to being able to implement the changes effectively.

**An overemphasis on testing as an assessment strategy**

While the OECD review revealed examples of schools using different forms and methods of internal assessment, including oral assessment and marking of on-going work, testing as an assessment tool, whether oral or written, predominates. In many cases, assessment was seen as being synonymous with testing.

The Assessment Reform Group (2002a) found strong evidence of the negative impact of testing on students’ motivation, partly due to the reduction in self-esteem of those who did not do well. In particular, they found that, unsurprisingly, students who tended to already be successful in the tests enjoyed them and were able to use appropriate test taking strategies to improve their results, whereas low achievers “become overwhelmed by assessments and demotivated by constant evidence of their low achievement” (ARG, 2002a, p. 4). In turn, this increases the gap between high and low achieving students. Further, “when tests pervade the ethos of the classroom, test performance is more valued than what is being learned.” This leads to a focus on performance outcomes (such as
scores or grades) rather than on learning processes (ARG, 2002a). This further undermines any formative purpose that might help students improve. The focus becomes on improving their grades or results rather than developing deeper understanding or identifying next steps and how to take them successfully (Butler, 1988, p. 1-14).

Equally, an assessment model predominantly focused on testing is better suited to traditional approaches to teaching and learning based on transmission of a fixed body of knowledge rather than encouraging students to develop their own ideas and interpretations independently. During the OECD review, several groups of students interviewed said the tests were mainly limited to closed or short response questions. This does not give students the opportunity to develop explanation, analysis or their own interpretations and therefore does not support or encourage higher order thinking (Paris et al., 1991, p. 12-20). This is particularly significant in relation to the Ministry for Education’s aim to address poor performance in relation to higher order reading comprehension skills in the PISA tests.

While testing clearly has an important role to play in teachers’ on-going assessment of students, it should be used mainly for what it is best suited to measuring; in most cases the transmission of knowledge. Students should have the opportunity to demonstrate their ability in different assessment contexts. For example, a student may be able to show conceptual understanding in a discussion with a teacher or peer that they would not necessarily be able to reproduce in a test setting. Typically, the current assessment system in many Slovak schools does not provide a variety of ways for students to demonstrate their competencies. A classic test assessment system disadvantages students who prefer more active and creative learning experiences and favours students who are able to master information sequentially (ARG, 2002a, p. 6).

The OECD review team found examples in some schools visited of other types of on-going assessment evidence being gathered, including practical assignments, presentations, oral assessments and reading diaries, although this was more common in basic schools. However, even where oral assessment was used, the potential for providing a different type of deeper evidence of students’ understanding was rarely realised. For example, oral assessment tasks would focus on recall (learning and reciting a poem for example) rather than dialogic discussion to demonstrate understanding and interpretation of the poem itself.

While some teachers were aware of a distinction between using tests primarily to assess knowledge and assessing application of skills through, for example, projects; they also recognised that the lack of any national assessment criteria beyond the minimum standards included in national education programmes meant that it was often difficult to have a shared sense of what was being assessed and what progression looked like. This was echoed by some students who, despite feeling that being assessed through their on-going work gave them greater freedom to show what they can do, were often unclear whether or how this was being taken account of in their grades.

**The predominant normative assessment model in Slovak schools limits feedback to students**

The normative focus of both internal and external assessment in the Slovak Republic, in which students are ranked in relation to others in the class, school year or nationally in percentiles, is problematic. While normative assessment is effective in establishing which students are better than others, it provides little evidence of why, and so has limited use in providing feedback to students on how to improve. In relation to national testing, there is
also a lack of clarity as to the relationship between the relative results and outcomes based on a normative model and those related to the National Standards for education, which imply measurement against objective levels of achievement. The relative nature of this assessment model also means that it is not appropriate for determining the effectiveness of schools, as it cannot effectively track the progress over time of the school population as a whole (see Chapters 5 and 6).

Norm referencing relies on the questionable assumption that the ability of a cohort remains unchanged from year to year. This can produce unreliable comparisons as attaining a high grade is as dependent on the performance of other students being assessed as a student’s own ability. Put simply, a student working at a particular standard could receive significantly different grades if they were assessed as part of an able or relatively weak cohort.

**Limitations and potentially negative impact on student motivation of age related standards in national education programmes**

Although the introduction of revised standards in national education programmes linked to Bloom’s taxonomy at the end of each phase of schooling is a positive step in unifying assessment practice, significant inconsistencies remain across subjects. For example, the standards for Slovak language and literature in the second stage of basic schools comprise two performance levels (minimal and optimal), whereas other subjects, including mathematics, only outline a minimum or do not specify a performance standard at all. Furthermore, using these standards to judge whether a student has met the minimum requirement for further study can lead to problems of student motivation as more able students will be able to achieve the minimum standard with little effort, leaving them little incentive to exceed this (William, 2001).

The use of age related national standards also makes it very difficult to measure progress meaningfully. For example, a student may achieve the “optimal” level of performance at the end of the first phase of their education, the minimal requirement at the end of the second phase and below minimal for the next phase; giving the impression that they are regressing, despite making steady progress in absolute terms (National Curriculum Task Group on Assessment and Testing, 1988, p. 31). Similarly, while the school’s interpretation of a student meeting the minimal requirement at different reporting points would be that they are making satisfactory progress, the student may feel demotivated by the fact that, despite making significant improvements over time, their grade has not changed (National Curriculum Task Group on Assessment and Testing, 1988).

Other research has found that in order to improve, learners need to believe that ability is “incremental”, i.e. that they can get better at something through practice and engagement, regardless of whether they feel they are “good” at it (Dweck, 1986, p. 1040-1048). This stands in contrast to the perception that their ability is “fixed”, which is when they believe that they are naturally good or bad at something and this is how they will stay. As in the example above, if students achieve the same grade over time, this potentially re-enforces the belief that their ability in that particular subject is fixed, while if the assessment system encourages students to see ability as incremental: “Rather than thinking of themselves as a “level x” person, they would think of themselves as someone who has so far reached level x, but, with more work, would be able to progress to “level x+1” (William, 2001, p. 7).
3. STUDENT ASSESSMENT – 55

Concerns about how the outcomes of high stakes assessment are used

The outcomes from high stakes student assessments at the end of ISCED 2 (Testovanie 9) and ISCED 3 (the external part of the Maturita) are used in a range of ways including ranking students in relation to their performance, providing feedback to schools to inform teacher planning, and as “entrance examinations” for the next phase of education (NÚCEM, 2012, p. 90). In addition, there appears to be an intention to use the results to compare school performance as a way of driving improvement (see Chapter 6 for further information). However, the OECD review team questions whether some of these uses are fit for purpose. As noted above, the normative testing model is not best suited to providing a reliable summative picture of a student’s attainment because it is linked to relative, rather than absolute, performance.

Currently, there is also inconsistency in how the results of external testing are used by gymnasia and universities as a proxy for entrance examinations in selecting the most able students for entry to their institutions. Notably there appears to be great variation in how universities use the Maturita results: some allocate places based solely on Maturita results; others where there is a high demand for places (such as medicine) take account of Maturita results, but also set their own entrance examinations; other universities are inclined to take students regardless of their Maturita results. During the OECD review, interviews with schools and regional governments suggested that the latter is a consequence of the introduction of per capita funding in 2002, which has driven universities to maximise their funding by increasing the number of students they admit. This has had the unintended effect of lowering entry standards and consequently undermining one of the main purposes of the Maturita. This situation is echoed with the Testovanie 9 tests, which interviews suggested was no longer being used consistently as a performance discriminator to inform entry as schools, particularly gymnasia, also seek to expand student numbers to increase income.

Lack of consistency in internal student assessment judgements within and across schools

Currently, students are graded for their on-going work by teachers according to a five-point scale (where grade 1 is very good and grade 5 is unsatisfactory or fail). However, there appeared to be wide variation in how these are applied, leading to significant inconsistencies within and across schools. Evaluation by the NÚCEM found large variations between grading across schools: “students who are regularly marked with grade 1 would be marked grade 3 if they studied at another school” (NÚCEM, 2012). During the OECD review, these findings were backed up by interviews with parents who shared a lack of confidence in the reliability of assessment, feeling that different schools graded more harshly or leniently. Given this context, it is unsurprising that the subjectivity of on-going teacher assessment is acknowledged to be a significant weakness in the Slovak system (NÚCEM, 2012, p. 89).

The OECD review team formed the impression that while some schools may develop internal criteria to judge the grading of student work, there was no guarantee that these judgements would be linked in a meaningful way to the National Standards for Education. In general, if there is no clear sense of what students are being graded in relation to, teacher judgements can be viewed as subjective and unreliable. This subjectivity also raises issues of student transfer, where the assessment information provided by one school may not match the standards of another school. Where assessment is not linked to external criteria, teacher judgements of students become more
open to the “halo effect”, that is, when awarding a grade to students, teachers are influenced by how they perceive students’ level of effort or behaviour. For example, a student who the teacher felt had tried hard, might be awarded a grade higher than the performance merited to encourage them. Further, the OECD review team formed the impression that teachers rarely shared criteria or learning objectives explicitly with students. This means that it is not possible to provide clear formative outcomes for the student as there is nothing specific to define what they have to do to reach the next stage of performance.

**Lack of effective use of assessment data at transition points**

As well as providing a summary of achievement at the end of a particular phase of education, one of the key uses of national summative assessment is to provide information that influences and informs future provision. When the information passed on is timely and sufficiently detailed, it supports transition by allowing schools and teachers in the next phase to adapt their approaches and curriculum planning to better suit the needs of the learners, as well as eliminating the need for re-testing on entry (Bew, 2001, p. 68-69).

The OECD review revealed that while student assessment information is passed on at transition points between basic and secondary schools, staff in both types of school feel that this could be done and used more effectively. Typically, secondary schools receive a simple certificate confirming that the student has satisfactorily completed basic education. Other examples of information passed on from basic schools may include the student’s grade average from Years 7 to 9 and/or the student’s percentile ranking in Testovanie 9, which provide limited information for secondary schools on how to plan differentiated provision for students, e.g. for gifted students or those with special needs. At the same time there appeared to be an established perception among teachers in basic schools that secondary schools were only interested in student performance information as a basis for selecting or rejecting a student for entrance into their schools.

**Use of tests provided by private companies may not support the revised curriculum**

A number of Slovak schools use commercially produced test-based assessments to inform their judgement of student attainment (NÚCEM, 2012; also confirmed during the OECD review). Further, certain commercial tests enjoy direct support from some municipal authorities and consequently allow the feedback to schools of comparative data, as students in schools throughout the municipality use these tests. These tests aim to provide summative data of the students’ level of knowledge in different subjects. Schools receive feedback of results that provide information on areas for individual students to improve and allow the comparison of student results to students in other similar schools. The use of additional externally developed tests can provide schools with useful information as part of their wider student assessment systems. However, during the OECD review a number of concerns came to light. The commercial tests most widely used by schools appeared to contain mainly closed response style questions, which are best suited to assessment of knowledge-based elements of the curriculum. It was not clear to the OECD review how well these tests reflect the revised curriculum with its focus on skills and competencies. Related to this point, the OECD review team understood that a major motivation of schools to buy and use such commercial tests is to better prepare their students for the Testovanie 9 tests. This reinforces both the impression that student assessment is heavily associated with testing (see above), and the importance of
establishing how useful commercial tests are in providing feedback on student progress against the national education programmes.

**Over-reliance on testing may narrow the curriculum**

In general, there is a risk that an over-reliance on testing as a form of assessment may narrow the curriculum. The potential risk may be exacerbated when the results of tests are used for accountability or performance measurement for either individual teachers or a school as a whole (see above and Chapter 5). This can create pressure to improve results at the expense of curriculum coverage, development of overall knowledge, skills and understanding, and student engagement. The curriculum can become narrowed as teachers focus only on what is assessed and a disproportionate amount of curriculum time is spent on practice tests or examinations. Consequently, teaching and learning becomes focused on “learning to pass the test” rather than on helping students to demonstrate what they can understand, do and apply independently. The OECD review team found some evidence of this “teaching to the test” in relation to both the external part of the **Maturita** and **Testovanie**. In the worst case, this means that, as well as an impoverished curriculum experience for students, improvement in school results over time may be as much due to students becoming more familiar with the test model as improvements in their attainment.

**Policy recommendations**

Although the OECD review team has identified a number of challenges, there are also several good initiatives that the following policy recommendations aim to build on:

- Phase in assessment criteria linked to standards in national education programmes.
- Integrate formative assessment into teaching and learning.
- Develop models to assess skills and competencies.
- Provide greater support for teachers in implementing curriculum and assessment change.
- Clarify the purpose and use of different types of student assessments.

**Phase in assessment criteria linked to standards in national education programmes**

For any assessment system to be reliable and comparable, students should be measured against the same, clearly defined, standards or criteria. The OECD review team recommends the introduction of assessment criteria that would show progression towards and be integrated with the National Educational Standards, and ultimately replace the 1 to 5 grade scale currently used. These criteria would describe the stages of progression leading up to the current standards in national education programmes and would show how students are working towards them. For example, a student assessed at level X will have met the criteria for level X and will be working towards level Y. A more detailed explanation of how this model would work in practice can be found in the report by the English National Curriculum Task Group on Assessment and Testing (National Curriculum Task Group on Assessment and Testing, 1988, p. 91-173).
Adopting this approach would provide a foundation to begin to address a number of the challenges identified earlier in this chapter and would support:

- Tracking of progression for individuals and cohorts.
- Greater consistency and reliability within and across schools in grading.
- Providing parents and students with a clear, fair and transparent understanding of how well they are achieving in relation to age related expectations.
- Formative assessment outcomes as the criteria will provide clear learning goals for students.
- Evaluation of impact for educational or pedagogical projects. For example, if a school introduces a programme to address the needs of particular student groups (e.g. students excelling academically, students of Roma background, students needing support in a particular academic area), assessment criteria would provide a tool to monitor how far the targeted students had improved.

In the first instance, we recommend that the introduction of assessment criteria should be focused on supporting on-going internal assessment in Slovak language and literature and mathematics in order to make the implementation more manageable for teachers. In the longer term, it could support a shift to criterion based national testing.

This would involve co-operation between the NÚCEM, the National Institute for Education and the Ministry of Education to produce assessment criteria that match the new curriculum, are compatible with the existing standards in national education programmes, and provide sufficient detail to be clear and unambiguous without being unwieldy to implement and use. The OECD review team recommends that the criteria are fully trialled with pilot schools before they are implemented and that teachers and educational researchers are involved in their development to ensure they are fit for purpose and appropriately set. The introduction of these assessment criteria should be supplemented by guidance and training to show best practice in how the assessment criteria could be used, and annotated examples of student work at different levels to establish a shared standard and support teachers’ professional judgements.

It is important that the work on the new curriculum standards by the National Institute for Education is accompanied by the development of not only curriculum examples for schools to be used in their development of the School Education Program, but also sample student test items. These test items will provide teachers examples of how to assess student progress against the national standards. For example, Germany (Bildungsstandards) and the Netherlands (with the national project on student reference standards in mathematics and language teaching “referentieniveaus”) provide examples of the development of national standards with accompanying curriculum and testing materials for all levels and school types. Members of the Consortium of Institutes for Development and Research in Education in Europe are well placed to refer to recent international experience in these areas (see www.cidree.org).

**Integrate formative assessment into teaching and learning**

Valid and reliable on-going assessment is vital for students to know how their learning is progressing and what they need to do to improve; and for teachers to know whether their students have understood what they have been taught, what level of attainment they have achieved and how planning can be improved. To meet both aspects a balance of summative and formative assessment is needed.
The meta-analysis of a wide range of research has shown that formative feedback, especially when from both teacher to student and student to teacher, is one of the most powerful and effective tools to improve student progress and the quality of teaching and learning (Hattie, 2008). In addition, a review of 20 studies of classroom assessment found that “innovations which include strengthening the practice of formative assessment produce significant, and often substantial, learning gains” (Black and Wiliam, 1998b). As such, supporting teachers in understanding and implementing formative assessment could be a key way of developing assessment expertise in schools and improving student achievement and independence.

Box 3.1 Principles for effective formative assessment

The Assessment Reform Group identified ten principles that could form the basis for introducing formative assessment practice into Slovakian schools, arguing that it should:

1. be part of effective planning
2. focus on how students learn
3. be central to classroom practice
4. be a key professional skill
5. be sensitive and constructive
6. foster motivation
7. promote understanding of goals and criteria
8. help learners know how to improve
9. develop the capacity for self-assessment
10. recognise all educational achievement.


Box 3.1 presents a set of principles that could be used for the basis of introducing formative assessment into Slovak schools. Implementing these principles effectively will involve a substantial shift in the teaching culture as a whole that will require a strong, long-term commitment from the Slovak government. The shift in emphasis from summative to formative assessment “may challenge [teachers] to change what they do, how they think about learning and teaching, and the way they relate to their students” (Mansell and James, 2009, p. 9). The professional development of teachers therefore, needs to go beyond simply delivering training on formative assessment, and address creating cohesion and coherence between curriculum planning, pedagogy and assessment practice.

In formative assessment for example, the tasks and questions planned need to be “fashioned in the light of their potential to engage students in making contributions that can reveal key strengths and weaknesses in their understanding” (Mansell and James, 2009). In other words, to have an effective formative outcome, teachers need to consider what they are teaching and how they teach it, in order to find out what they want. This approach also needs to give students the opportunity to take more responsibility for their own learning and develop their skills in “learning to learn”.

OECD REVIEWS OF EVALUATION AND ASSESSMENT IN EDUCATION: SLOVAK REPUBLIC © OECD 2014
A large scale study of Portuguese primary schools shows the impact of introducing formative “learning to learn” practices in the classroom. Teachers in the study who trained their students to use self-assessment strategies found that their attainment showed a 50% improvement in comparison with the control group (Fontana and Fernandes, 1994, p. 407-417). A further study showed that students able to use self-assessment techniques became “less inclined to attribute outcomes to luck, and are better able to identify the real causes of the academic events that happen around them” (Fernandes and Fontana, 1996, p. 309).

It is possible to use testing formatively if teachers discuss students’ responses with them in a way that develops their knowledge and understanding of how to improve. However, teachers should be encouraged to use a wider range of assessment tools and methods, such as dialogic questioning (Alexander, 2008), observation of student interaction, and peer and self-assessment. This will involve learners in the assessment process and provide a richer source of evidence that gives a more rounded picture of the student’s knowledge, skills and understanding when providing feedback.

**Develop models to assess skills and competencies**

During the OECD review, regional directors and employer’s associations acknowledged that the Slovak economy needs a highly skilled workforce that meets the needs of the modern labour market. The revised Slovak curriculum aims to address this by developing competencies that can help young people to cope with social, economic and technological change, and succeed in school and the world beyond. However, if these competencies are to be recognised, valued and developed, they need to be assessed using appropriate tools.

While many countries and jurisdictions, most notably various Australian States and New Zealand, have moved towards a greater focus on wider skills in their curriculum (Lucas and Claxton, 2009), there are different opinions on how best to assess them. It is agreed, however, that there is a need to develop assessment tools and methods that are suited to recognising skills and competencies, and to capture evidence from a wider range of sources (see Box 3.2). Without this there is the risk of a significant mismatch between curriculum content and assessment, which means the latter is not fit for purpose.

**Box 3.2 Assessing broader skills and competencies**

The assessment of personal skills and competencies should (Futurelab, 2007):

- be integrated into the learning process
- be sensitive to context and complexity
- promote self-worth and development
- be meaningful to and owned by learners
- act as a bridge and currency between learners and diverse communities
- enable multiple comparisons and lenses
- recognise collaboration.
Box 3.2 Assessing broader skills and competencies (continued)

Suggested methods to support the assessment of broader skills and competencies include (Lucas and Claxton, 2009):

- self-report questionnaires of various kinds
- evaluation of students’ learning portfolios or diaries, or other written reflections
- structured teacher observation in terms of various quasi-objective “ladders of progression” for each of the wider skills
- “learning stories”: short vignettes and digital photos or videos that capture a series of increasingly accomplished “leading edge moments” in individual students’ learning careers
- periodic 360-degree assessments of student progress drawing on testimony from parents, friends, teachers and coaches, as well as documentary evidence of various kinds.

Sources:

Lucas, B. and C. Claxton (2009), Wider Skills for Learning: What are they, how can they be Cultivated, how could they be Measured and why are they Important for Innovation?, NESTA, London.

Any assessment model developed needs to recognise what is distinctive about assessment in the context of skills and competencies, Box 3.2 sets out some principles for this. Significantly, these principles have implications not only for assessment practice, but also for pedagogy and student behaviour. This requires a broader change in school culture that will take time and require an investment in training, guidance and monitoring (see also below).

In recognising and gathering the often more ephemeral evidence of wider skills, teachers in basic schools and beyond may have much to learn from their kindergarten colleagues. Here, as the curriculum is predominantly play-based, assessment and evaluation is primarily observational, with the dual purpose of establishing the child’s level of development in key competencies, and diagnostically influencing teacher planning. Rather than relying on testing or written evidence, the teacher uses their own notes from observation and dialogue with the student to inform their assessment judgements (Eurydice, 2009, p. 15).

Provide greater support for teachers in implementing curriculum and assessment change

Changes as radical as those introduced in the School Act (2008) will inevitably take time to implement and will require a significant shift in culture within schools. For example, the change in focus of the T9 test to better reflect the revised curriculum has led to greater emphasis on higher level reading comprehension skills and more contextualised tasks in mathematics. In both cases the questions are intended to move away from the more traditional style of testing that mainly requires students to remember and reproduce information.
However, there is evidence in the overview of the 2011 T9 results that many students, and by implication teachers, found this shift problematic. For example the NÚCEM’s overview of Testovanie 9 results for 2011 found that while students were able to apply learned formula to calculate the volume and surface areas of solids in mathematics, they found questions that involved selecting and applying a relevant strategy or linking knowledge from different areas much more challenging (NÚCEM, 2012). Similarly, in the Slovak language and literature test, students were confident on questions that required them to identify, select and retrieve relevant information from texts, but found higher level comprehension skills including interpretation, contextual awareness and questions requiring responses in their own words considerably more difficult (NÚCEM, 2012, p. 88). This is further evidence that not all schools have been able to successfully meet the challenge of the new curriculum. Indeed, the OECD review team formed the impression that in general, teachers were not prepared or sufficiently supported in the move towards greater autonomy, as school leaders, teachers and students have been used to a more traditional “top down” organisation.

Here, the Methodology and Pedagogy Centre (MPC) has an important role to play and at the time of the OECD review was in the process of revising the professional standards for teachers. Recognising that student evaluation was a weakness, as teacher training had traditionally focused more on content and knowledge than pedagogy and assessment practice, diagnosis of students will be one of the three dimensions in the revised standards. This is an excellent opportunity to ensure that the principles and practice of formative and skills assessment are included in this dimension (see also Boxes 3.1 and 3.2).

In parallel with this, the MPC should also provide training and guidance for existing teachers, both nationally and through its regional offices, to develop teachers’ practice in line with the revised professional standards. There is also the potential to make greater use of NGOs to facilitate networking, peer support and sharing of best practice in assessment.

**Clarify the purpose and use of different types of student assessments**

In common with other countries such as the United Kingdom, the OECD review team notes that the national assessments in the Slovak Republic are being used for numerous purposes. These include providing an objective and comparable measure of students’ performance at the end of a particular phase of education; giving schools information that can be used to inform teaching and learning and curriculum planning; assessing student readiness for the next phase of education; and as a means of measuring school performance.

However, the more purposes a particular assessment has, the less effectively it is able to achieve them and the greater the risk of unintended negative consequences. For example, if a national test aims to produce an objective and comparable measure of performance, its scope will tend to be restricted to what can be marked reliably. This in turn limits the range of what can be assessed and therefore the diagnostic potential of the outcomes. Further, as mentioned previously, a strong emphasis on the results as a means of ranking or evaluating the effectiveness of teachers and schools can have a distorting effect on the curriculum, where drilling students to pass the test takes precedence over providing a rounded curriculum that enables students to apply their knowledge and skills independently.
In this context, the OECD review team recommends that the NÚCEM review national assessments to ensure that they offer the most effective means of measuring a particular outcome or providing a particular type of information. For example, on-going formative teacher assessment may deliver a more responsive and effective means of informing and adjusting curriculum planning than an assessment that happens at the end of a phase of education which, by definition, can only provide information retrospectively when it is too late to have an impact on that cohort of students. Similarly, while performance in national assessments may provide one indicator of the effectiveness of a teacher or school, in order to avoid an over emphasis on a narrow range of assessment outcomes at the expense of a broad and balanced curriculum, other measures and information should be taken into account. This latter point is also considered in more detail in the policy recommendations for Chapter 5 on school evaluation.

Finally, better use should be made of both test results and teacher assessment information to ensure improved transition between phases of education. Rather than merely supplying grade averages, other forms of more qualitative information could be provided that outline, for example, student background, interests and aptitudes and include examples of work. Representatives from secondary schools reported that such information would be very helpful and may help them see transition information as a means of ensuring challenge and curriculum progression for students rather than primarily as a means of selection.
Note


References


Lucas, B. and C. Claxton (2009), *Wider Skills for Learning: What are they, how can they be Cultivated, how could they be Measured and why are they Important for Innovation?*, NESTA, London.


