Chapter 4

Student Assessment

Sweden has a balanced approach to student assessment that captures a wide range of learning dimensions. There is a strong focus on classroom-based assessments through which teachers collect a wide range of evidence on student progress and provide regular feedback to students. National tests at key stages of education are intended to capture a variety of curriculum goals through performance-based tasks including oral assessment and team projects. However, as all other types of assessment in Sweden, the national tests are corrected and graded by the students’ own teachers, and the weight of test results in students’ grades is determined locally. This raises concerns about inequities in grading. Given the key role that national assessments play in the Swedish evaluation and assessment system, it is vital to increase the reliability of these tests. External moderation could help ensure consistency, comparability and equity of the national assessments. Capacity building for effective summative and formative assessment is also key to strengthening teacher and school leader practices.
This chapter focuses on approaches to student assessment within the Swedish evaluation and assessment framework. 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 (EPPI, 2002). This chapter looks at both summative assessment (assessment of learning) and formative assessment (assessment for learning) of students.

4.1 Context and features

**Student assessment and the role of national tests**

Students receive a summary statement of their achievements in school through end-of-semester reports in Years 8 and 9 as well as in the school-leaving reports at the end of upper secondary school. These summary statements of student learning are based on teachers’ continuous assessments in the classroom, which is supported by compulsory national tests in certain subjects and stages of education.

National tests exist for key stages in compulsory school (Years 3, 5 and 9) and in upper secondary school. The results from national tests are one of the bases for teachers to determine students’ overall grades. Teachers grade the national tests for their own students and each school decides how to weigh the national assessments and course grades. According to the National Agency for Education (NAE), the primary purposes of the national assessments are to:

- Ensure that all students have the opportunity to achieve goals for learning, regardless of gender, race, economic background, or place of residence.
- Ensure fair and equitable grading across schools, as the manner and frequency of assessment varies a great deal from municipality to municipality (and even from school to school). The NAE has also established grading criteria for different subjects offered in the curriculum.

National assessments in Years 3 and 5 are intended for diagnostic and formative purposes. These assessments cover Swedish/SSL, mathematics and English (in Year 5 only). They are compulsory and must be administered by schools in a nationally specified period in the spring.

The national tests in Year 9 and those in upper secondary school are summative. Students are required to sit assessments in the core subjects (Swedish/SSL, mathematics and English). In addition, each school administers national tests in one of the science subjects (biology, physics or chemistry), as allocated by the NAE4.

Beginning in 2012, students in Year 6 will take national assessments in Swedish/SSL, mathematics and English. These assessments will be compulsory and will replace those now given to students in Year 5.

Tests are also available “on demand” in different subjects, including foreign languages, social science subjects and selected vocational subjects from a test bank run

4. For this purpose, the NAE created three samples of schools, each representative for the total student cohort in Year 9. The samples consider parents’ level of education, Swedish/foreign background and responsible school organiser.
by the NAE. The tests, which are aligned with the national curriculum, are intended to supplement teachers’ own classroom-based assessments. In addition, the NAE provides diagnostic materials in the core subjects.

**A new grading system**

Currently, students do not receive grades until Year 8. From Year 8 and through upper secondary school, students receive a term grade at the end of the autumn and spring semesters. Currently, the grade levels in Year 8 and 9 are: G (Godkänt – Pass), VG (Väl Godkänt – Pass with distinction), MVG (Mycket Väl Godkänt – Pass with special distinction). If a student does not fulfil the requirements for a passing grade, no grade is awarded in the subject. In upper secondary education, there is also a failing grade – IG (Icke Godkänt). Under the current system, students who meet “goals to attain” as set out in the course syllabi may be awarded a Pass grade. The NAE sets out criteria describing the kinds of performance students must demonstrate for the more difficult to attain VG and MVG grades.

The Government has proposed making grades compulsory from Year 6. The intention is to help students get used to being graded before they start lower secondary education. This option is still under discussion and there are advocates on both sides of the argument. The Government has also proposed the introduction of five grades instead of the current three grades (G, VG, MVG). The introduction of additional grade levels are intended to provide teachers with more options in assigning grades – for example, if they believe a student’s work falls between a VG and MVG performance. Reducing the interval between grade levels is also intended to increase students’ motivation to achieve better results, as the next level will be more easily achievable.

**Formative assessment**

Formative assessment, which the OECD (2005a) defines as the frequent assessment of student progress to identify learning needs and adapt teaching, is supported in Swedish schools through:

- Regular development talks with students and their guardians.
- Individual Development Plans (IDPs).
- Student involvement in goal setting and self-assessment.

Individual school leaders set out the general template for the IDP that will be used in their school. The IDP is to include an assessment of the student’s current performance levels in relation to learning goals set in the curriculum and syllabi, and steps the student should take to reach those goals. Whether to include additional information, such as the student’s more general development (e.g. the student’s ability to take on responsibility, their social skills, and so on) is up to the school leader. The written IDP is to include the student’s and guardian’s input from the regular development talks, which usually take place once a semester. For students who are experiencing difficulty, schools are required to document plans as to how they will help students achieve goals.
4.2 Strengths and challenges

Strengths

Strengths associated with performance-based assessments

Sweden’s national assessments measure student progress toward standards embedded in the national curriculum. The assessments are performance-based – that is, students are scored on open-ended performances, such as written essays, oral communication skills, demonstrating reasoning processes, collaborative problem solving, and so on. Compared with close-ended testing formats, performance-based assessments are often seen as being more effectively aligned with curricula that emphasise development of higher-order thinking skills and capacity to perform complex tasks. Such tests assess a range of integrated knowledge and skills by asking students to perform a task rather than to select a correct answer (Wren, 2009). The national assessments, each of which may take several hours, and may be spread out over several weeks, cover a wide swath of the curriculum. Teachers thus have more information on student performance across a range of tasks, and a better idea of each student’s development and progress.

Assessments for students in Years 3 and 5 are used solely for diagnostic and formative purposes. Several of the younger students we interviewed commented that they had enjoyed taking the tests. The Year 9 students we spoke to, as well as those taking examinations in upper secondary education, were naturally more anxious about the assessments, as the results are one of the bases on which teachers determine student grades. At the same time, several of the older students felt that the assessments had been valuable in getting them to reflect upon what they had studied during the year. The fact that students who do not pass one or more of the national subject tests may re-take them helps to lower the personal stakes for students somewhat.

Both primary and secondary students with whom we spoke said that they believed the tests were set at a reasonable level of difficulty. Some students even said that they felt the tests were too easy. This may reflect the current two-tiered system of standards, which sets out goals to attain and goals to strive for. However, this two-tiered system will be removed when the new curriculum comes into force in July 2011.

Advantages of teachers scoring their own students’ performance on national assessments

Teachers score their own students’ performance on the national assessments. Often teachers work in teams, with colleagues from their own school, or with teachers from other local schools who teach the same subjects, to score the tests (Swedish Ministry of Education and Research, 2010). While there are challenges associated with this approach, as will be noted below, there are also advantages. For example, the scoring experience may serve as an important form of professional development for teachers. Teachers are able to discuss views on student performance with their peers.

5. By contrast, standardised assessments with close-ended answers, such as multiple-choice, true-false or fill-in-the blank tasks, tend to focus on content rather than thinking skills. In such standardised tests – which are often machine-scored – tasks are treated as discrete items and may not capture reasoning processes behind student responses.
They can also use the national tests for formative purposes, returning the work to students along with feedback on strengths and weaknesses. Such feedback can be given to individual students, but also collectively to the whole class (Crooks, 2004). The reviewing of national tests provides opportunities for teachers to analyse the impact of past teaching and learning approaches and adapt instruction.

Another clear advantage of teacher involvement in scoring is that schools have test results much more rapidly than they would if they were relying upon a separate organisation to deliver the results. This is important in terms of teachers’ ability to use the results in a formative fashion (more will be said below about the importance of timing for giving formative feedback). Teachers are also much more likely to refer to the results of the national assessments and to adapt instruction to meet student needs when they have spent time directly reviewing their own students’ performance.

**A strong focus on classroom-based assessments**

Schools in Sweden have full autonomy in determining the criteria for the internal assessment of students (Annex 5). Each school decides how to weigh the results of the national examination and the teachers’ assessment.

Although there are concerns about the unevenness of teacher grading both within and between schools – and these concerns will need to be addressed – the practice of basing final grades on a broad range of evidence on student achievement is important and should be continued. Teachers have many more opportunities to observe students over time and performing a variety of tasks, including extended projects, and in this sense their observations have higher validity. Teachers are also less likely to “teach to the test” when they are able to take into consideration a range of experiences and observations of student performance. The fact that classroom-based assessment takes place on multiple occasions reduces the risk of student assessment-anxiety.

Classroom-based assessments are further supported by the fact that national test banks are available, so that teachers may choose assessments they would like to use for their own purposes. Teachers with whom we spoke noted that they used tests from the NAE’s central test bank from time to time. Given that students learn at different rates, the fact that teachers are able to download tests when they believe students are ready is also very positive. Control over timing of the tests also means that teachers may provide students with feedback on their test performance when it is relevant to what they are learning. Scotland uses a similar approach, and these tests are very popular. The NAE is currently expanding the number of tests available through the central test bank.

**A firm foundation for formative assessment**

Classroom-based formative assessment involves the minute-to-minute, day-to-day interactions between and among teachers and students that help to uncover how well students understand new concepts, and where teachers may need to adjust teaching to better meet learning needs. In classrooms featuring formative assessment, students are frequently engaged in assessing their own and their peers’ work as they build their skills for learning to learn. The results of summative tests (classroom-based tests or national assessments) may also be used formatively. However, assessment is considered as formative if and only if it shapes subsequent learning (Black and Wiliam, 1998; Wiliam, 2006).
Sweden’s focus on engaging students in setting goals for learning through the IDP, and developing skills for self- and peer-assessment are important for the effectiveness of formative assessment. Teachers are generally more likely to focus on formative assessment when they have tools and guidelines to support the process (OECD, 2005a). The IDP, as a core feature of Swedish education, ensures that both teachers and students are focused on identifying individual learning goals, and developing strategies to address any shortcomings. It can be a powerful tool for developing students’ own assessment skills, as well.

While the review team did not have the opportunity to observe any classes, both teachers and students indicated that Swedish classrooms do support many of the elements of effective formative assessment. For example, students interviewed for the review said that their teachers give them regular feedback on the quality of their work, and that they usually know how well they are performing and what they need to do to improve their work and reach learning goals. Several students commented that they felt well supported in the learning process. They also said that they frequently assess the quality of their own or their classmates’ work, and they found this process useful.

While some stakeholders expressed concerns that students should get used to receiving grades earlier in their education, in terms of formative assessment, Sweden’s relatively low-key focus on student grades is a positive point. In their review of the literature on formative assessment, Black and Wiliam (1998) found that the grading function in schools tends to be overemphasised while learning is underemphasised. In many ways, formative assessment is fundamentally about the quality of interactions between and among students and teachers. In this regard, Sweden’s strong focus on student-centred learning and on the importance of helping all students to achieve are major strengths. As will be discussed below, additional support for building teachers’ skills in different approaches and techniques will further strengthen formative assessment in day-to-day teaching, learning and assessment.

Challenges

Some cautions regarding reliability and generalisability of national test results

While performance-based assessments, the format used for the Swedish national tests, have many advantages over standardised assessments, some cautions must also be noted. The first is that, while performance-based assessments are, in principle, more effectively aligned with curricula that emphasise higher-order thinking skills, this is not necessarily always the case. Researchers in the United States found that performance-based assessments frequently do not measure the skills and processes intended (Baxter and Glaser, 1998; Hamilton et al., 1997; Pellegrino et al., 1999). In other words, they may be of limited validity. While the stakeholders with whom we spoke were very positive about the national assessments, any judgment on the validity of the current assessments (in other words, whether they measure what they are intended to measure) would require a more in-depth evaluation.

Performance-based assessments also tend to have lower reliability and generalisability than do standardised assessments. Research in other countries has shown that it is very difficult to generalise from hands-on performance-based tasks to make judgements about student competencies. Shavelson et al. (1990), for example, found that performance-based assessments in science were highly task dependent. This is in line
with research showing that higher order thinking skills are context and situation specific (Linn et al., 1991).

There are important challenges for teachers in Sweden and elsewhere in scoring open-ended performance assessments. The Swedish Schools Inspectorate has recently undertaken the first of three national corrections of teacher scoring of student performance on national tests and concluded that the current grading and assessment system is not reliable. Overall, the Inspectorate re-corrected 35 000 tests taken by students in Year 9. While there were no big differences between the marking of teachers and cross-checkers in mathematics, there were indeed large discrepancies in Swedish (open-ended questions). Overall, the reviewers have found that teacher scoring of national tests continues to be very uneven.

**Concerns about inequities in teacher grading**

According to the stakeholder groups with whom we spoke, the current standards and learning goals have long been considered as being too vague to guide instruction and assessment. Teachers may interpret learning goals and the grading criteria in many different ways which leads to inequities in teacher grading. While the new curriculum has yet to be released, several of the stakeholders we interviewed had had the opportunity to review at least some portion of the new standards and learning goals. They reported that the new goals within the curriculum are more concrete and believed they will help to address at least some of the unevenness in teacher grading within and between schools and municipalities.

Another area of concern is in regard to equivalence of student grades (reliability) across schools. There is no national guideline as to how much weight should be given to the national test result within the overall grade assigned to students. A 2009 study by the NAE stated that there are great differences between how teachers designate students’ overall grades in relation to their national test results. There were large differences both between schools and between teachers within a school. According to the study, some teachers set grades that are significantly higher than test results, others parallel to test results and some assign grades lower than the test results (NAE, 2009b).

**Limits to the use of the national tests to diagnose student needs**

Year 3 and Year 5 teachers consider the national assessments as diagnostic and formative. However, schools have to administer the tests in a nationally specified period in the spring, which means that results are only available very late in the school year. This poses less of a problem for schools where teachers remain with the same students for more than one year. Teachers and students would nevertheless derive more benefit from having results of the tests early in the school year.

In addition, such standards-based assessments typically do not provide the level of detail needed to develop profiles of individual student needs. As is typical for standards-based assessments, scoring of Sweden’s national assessments is criterion-referenced. That is, scores describe student performance relative to performance targets. Criterion-referenced scores are usually reported as broad proficiency categories, such as

---

6. Note carefully that “standards-based assessments” refer to assessment of progress toward learning standards, while “standardised” assessments refer to a testing technology.
basic, proficient and advanced (Cizek et al., 2004). The “cut score” is the level at which students pass the test.

While criterion-referencing is appropriate for standards-based assessments, if the proficiency categories are too broad they will mask significant heterogeneity in student performance. This is particularly important for students who perform “below expectations”, and for whom it is important to diagnose the source of learning difficulties. For example, Rupp and Lesaux (2006) found that it was virtually impossible to disentangle the cause of reading difficulties based on the single global score or classification of standards-based assessments. Thus before teachers are able to develop remedial plans, lower-performing students may require additional diagnostic testing. A number of empirically validated diagnostic tools are designed to identify the source of learning difficulties, and better shape remediation programmes. Teachers should be aware of the limits of standards-based assessments and should have training to use validated diagnostic tools.

Another challenge with the national assessments in Year 5 is that these tests are used over two successive years. It seems inevitable that teachers, who will have spent considerable time administering and scoring the tests in the first year of its use will tend to “teach to the test” in the second year. There is therefore a risk of score inflation – that is, scores will overstate improvements in student learning. The information on student performance might therefore be less useful in diagnosing student needs.

**Teacher training for assessment competencies is still limited**

Teachers may receive training to build their assessment competencies (for both formative and summative assessment) during initial or in-service training. However, based on feedback from the stakeholders we spoke to, up to now teachers have not been required to take courses in assessment and evaluation during their initial training or in professional development courses. This is a significant gap, as skills for both formative and summative assessment are key to the success of Sweden’s approach to “management by objectives” in education. Teachers and school leaders we interviewed also noted that they took a somewhat ad hoc approach to choosing courses for professional development. Very often, they choose courses according to their own interests, which are not necessarily aligned with the school’s overall development needs.

In a 2007 research project, Peterson and Vestman (2007) described the range of courses addressing student assessment and organisational evaluation available at Swedish universities offering teacher education. The different programmes described in the project tended to offer one or two courses in these areas, but they were often electives or embedded in courses on other subjects. Some of the educators with whom the OECD review team spoke noted that their own teacher training had placed very little emphasis on assessment and evaluation skills. The new initial teacher training to be implemented from July 2011 may help provide teachers with better basic assessment literacy. It contains specific goals and a mandatory course related to assessment and grading. Assessment topics are also expected to be integrated into the didactics of every subject of the new initial teacher training. The new School Leadership Training Programme also includes a module on working in a results-oriented system.
The challenge of integrating formative assessment in day-to-day practice

The results of national assessments, inspection reports, student IDPs and classroom interactions may all be used formatively. The distinguishing feature of formative assessment for any of these approaches is that the information be used to make improvements (Bloom, 1968; Scriven 1967). But the way in which information is used and the timescale for decisions may be very different. Wiliam (2006) distinguishes between long-, medium, and short-cycle formative assessment. According to Wiliam, long-cycle formative assessment occurs across marking periods, semesters or even years (four weeks to a year or more); medium-cycle formative assessment occurs within and between teaching units (three days to four weeks); and a short-cycle formative assessment occurs within and between lessons (five seconds to two days).

Sweden’s student IDP and the emphasis on the use of the national assessments as a diagnostic and formative tool (at least for younger students), might be considered as long- and medium-cycle formative assessments. These assessments are important for identifying areas of need, developing broad teaching strategies to address needs identified within the student cohort, planning, allocation of resources, and so on. But short-cycle formative assessment – the daily interactions between and among students and teachers – has the most direct and measurable impact on student achievement (Looney, 2011). In short-cycle interactions, formative assessment is part of the classroom culture, and is seen as an integrated part of the teaching and learning process. Teachers systematically incorporate formative assessment methods in their course planning – for example, in how they intend to develop classroom discussions and design activities to reveal student knowledge and understanding. These interactions encompass effective questioning to uncover student misconceptions and identify patterns in student responses, feedback on student performance and guidance on how to close learning gaps, and student engagement in self- and peer-assessment.

The way in which teachers approach these different tasks is also important. For example, studies show that feedback which does not provide students with specific guidance on how to improve, or that is “ego-involving”, even in the form of praise, may have a negative impact on learning. Feedback that is focused on the process of learning and that tracks student progress over time is more effective (Köller 2001; Mischo and Rheinberg, 1995). Questions focused on causal effects or that aim at uncovering misconceptions are much more effective than “yes or no” questions or questions that stress recall rather than reasoning processes, which are much more typical in classrooms (Black, 1993; Black and Wiliam, 1998; Stiggins et al., 1989). If formative assessment is to be effective, teachers need to have strong skills to adapt teaching, as well. Teachers thus need to be able to call upon a broad repertoire of teaching methods to better meet individual student needs.

Sweden already has a firm foundation for effective formative assessment - particularly in the value it places on student-centred learning and in use of the IDP as a tool for individualised student assessment – but formative assessment can still be further improved. Based on our conversations with school professionals, our impression was that Sweden may further strengthen classroom-based formative assessment by placing a stronger focus on short-cycle classroom interactions, and in building teachers’ repertoire of research-based formative assessment techniques as well as ways to respond to identified learning needs and capacity to adapt to individual student needs.
4.3 Pointers for future policy development

Sweden is taking a number of positive steps to strengthen formative and summative assessment of students – including the introduction of a new curriculum with clearer and more concrete goals for learning. Policy makers may want to consider additional strategies. The preceding discussion of strengths and ongoing challenges within the Swedish assessment system suggest a number of potential directions for policy. These include:

- Strengthen reliability and generalisability of the national assessments.
- Invest in initial training and professional development to strengthen teachers’ assessment skills.
- Develop tools to support teacher assessment.
- Strengthen short-cycle, classroom-based formative assessment.

**Strengthen reliability and generalisability of the national assessments**

Given the key role that national assessments play in the Swedish evaluation and assessment system, it is vital to increase the reliability of these tests. There are several options of doing so.

Some of the concerns about the reliability and generalisability of teachers’ marking of the national assessments may be addressed through training for raters. Caldwell *et al.* (2003) have found that such training can increase the reliability of scores. This would require further investment in improving teacher capacity to assess students specifically on the national tests by providing more detailed guidelines on scoring and participation in scoring workshops for different disciplines. Results from the series of three reviews by the Schools Inspectorate should provide useful information on the extent of variation in reliability of teacher grading.

Consideration should also be given to establishing a systematic external validation of national test results, for example via a random checking procedure conducted by a competent authority or even establishing an examination authority to score student results in national tests. External moderation of teacher-based assessments can help increase consistency, comparability and equity of teacher-based assessment (for an example from Queensland, Australia, see Box 2). An external checking procedure would require additional resources at the central level with competent psychometricians, but also could benefit from collaboration with educational professionals in scoring different disciplines (see Chapter 7).
In Queensland, there is no whole-cohort external testing or examining in secondary schools. In 1972, Queensland abolished external examinations and replaced them with a system of moderated internal assessments. School-based assessments for the Senior Certificate (Year 12) are currently moderated for those subjects that count towards university entrance. The moderation processes for the Senior Certificate involve subject-based panels of expert teachers providing advice to schools on the quality of their assessment programme and their judgments of quality of student performance based on sample portfolios. The system involves follow-up where panels identify difficulties. There is negotiation of the final results to be recorded on the Senior Certificate. Results are expressed in terms of five relative grades or ‘levels of achievement’ expressed in terms of standards descriptors (referred to as “exit standards”).

Source: Sebba and Maxwell (2005).

In addition, the use of so-called “complex assessments” that combine both performance-based assessment and standardised close-ended questions may help. Such an approach builds on the strengths of both types of assessment: higher validity of performance-based assessments, and the reliability and generalisability of standardised assessments (Linn et al., 1991; Pellegrino et al., 1999).

Another option would be to invest in the development and use of high quality ICT-based programmes to assess complex performances and track students’ problem solving skills. As of yet, the use of computer-based assessments is very limited in Sweden7, although there has been some discussion regarding this possibility. The student councils have also suggested that this would be an important step in improving the national assessments. International test developers are now devoting significant attention to developing effective computer-based assessments that can measure students’ reasoning processes and other higher-order cognitive skills and score “open-ended” performances, such as student essays8. Technology-based assessments may also include simulations, student collaboration and constructed response formats. Students may receive feedback on their performances as they are taking the test – blending formative and summative functions of the assessment (Bennett, 2001; Lewis, 1998; Mislevy et al., 2001). However, there is still quite a bit of development work on computer-based assessments and these approaches are not yet widespread.

It is important to note that each of these proposals would add to the cost of the current system. But the level of confidence in results will be much higher. Training to support improved scoring should be in addition to, not as a replacement for, other professional development to support instruction and assessment competencies. Complex assessments would add a standardised portion to the current assessments – not replace it. While test developers have piloted some pioneering ICT-based assessments, more work needs to be

---

7. The NAE has developed ICT-based assessments for Swedish for Immigrants (SFI) but due to technical difficulties these are for currently only used in some municipalities.

8. ICT programmes that score “open-ended performances” are still in the relatively early stages of development and while they may facilitate scoring of large-scale assessments, cannot replace human raters. Further studies are also needed to determine the validity and generalisability of different automated essay scoring tools (see Wang and Brown, 2007).
done before these kinds of tools are available across different subjects for students at different levels of development.

Finally in order to ensure that Sweden’s performance based assessments measure the skills and processes intended, it would be important to evaluate the validity of the assessments. The review team heard very positive feedback regarding the national assessments, from teachers as well as students. However, it is important to ensure that the assessments are valid – *i.e.* that they measure what they were intended to measure. External valuation of the validity of assessments would add to confidence, as well as usefulness of tests for policy decisions, or development of instructional strategies.

**Invest in initial training and professional development**

In Sweden’s goal-oriented education system, strong teacher skills for both formative and summative assessment are essential to monitor progress towards learning goals. As discussed above, training is particularly important to ensure the reliability of teachers’ scoring of national tests. Training for teachers scoring tests for students in Year 9 and in upper secondary school is particularly important, as the results of these assessments have important consequences for students. Moreover, in a system where teacher-based assessments have an important place, teachers also need opportunities to develop and improve their own skills for test development. Teacher-based assessments also need to meet criteria for validity, and to be aligned with central learning goals, particularly since teachers’ assessments largely determine students’ final grades.

For teachers in Years 3 and 5, where the tests are used primarily for diagnostic purposes, teachers need skills to interpret results, to understand whether further diagnostic testing of some students may be warranted, and to identify areas where curricular strategies may need adjustment, or where they may invest resources in new programmes to meet student needs. Ongoing attention to teacher training in formative assessment is also vital. Effective formative assessment requires that teachers develop sophisticated skills for uncovering students’ level of understanding, for providing feedback and adjusting teaching strategies to meet identified needs, and for helping students to develop their own skills for learning to learn. Sweden’s emphasis on student-centred learning also means that teachers need skills to help students develop their own skills for self- and peer-assessment.

Training to develop assessment competencies should start with basic assessment literacy, for example, the ability to understand different aspects of validity – what different assessments can and cannot reveal about student learning. Assessment training should also overlap with knowledge of how students learn in different domains so that teachers are able to interpret patterns of student responses to identify misconceptions, and to respond with appropriate instructional strategies.

**Develop tools to support teacher assessment**

The steps already taken to improve the standards and learning goals embedded in curriculum will go a long way toward improving assessment. The NAE may want to consider providing additional tools to support teacher assessment, such as exemplars illustrating student performance at different levels of achievement, and scoring rubrics listing criteria for rating different aspects of performance. This can help guide teacher assessment. For example, the Ministry of Education in Newfoundland and Labrador, Canada, disseminates rubrics with specific guidelines and criteria for evaluating student
work (OECD, 2005a). The rubrics describe levels of quality for each of the criteria, usually on a point scale. Teachers may also use rubrics for classroom-based assessments, sharing the different criteria with students so that they understand different levels of quality work.

Ensuring that the national tests in Years 3 and 5 take place earlier in the year would also help support teachers’ diagnostic assessments. France, the French-speaking community of Belgium and Spain all take this approach. In these countries, national assessments are administered to students who have just made key transitions in their schooling (e.g. from primary to lower secondary schools). At the policy level, trends identified within the aggregate data help to shape policy and identify areas where the majority of students are performing below expectations. At the school level, teachers may identify areas where several new students are having particular difficulty, and adjust curricula to meet these needs (Looney, 2011).

Tests made available through the central test bank can also be used by teachers to design their own classroom-based summative and formative assessments – providing ideas on questions or process that will help identify student misconceptions in different learning domains.

**Strengthen short-cycle, classroom-based formative assessment.**

Sweden has a strong foundation for effective formative assessment. The IDP, the focus on using data from national assessments for younger students to improve teaching and learning, and the strong student-centred culture are all very positive. Policy makers may consider strengthening teachers’ skills day-to-day formative assessments – including skills for setting up learning situations, developing sophisticated questions, providing timely feedback, and so on. These short-cycle formative assessments are likely to have the most direct impact on student achievement. They also are important for ensuring that assessment is not an “add on”, but is integrated with teaching and learning. Formative assessment becomes a part of the culture of the classroom. In this way, Sweden might develop an even stronger framework incorporating long-, medium- and short-cycle formative assessment to improve learning and outcomes for all students.