

Chapter 3

Student assessment

Student performance in Portugal is assessed by a wide range of instruments, ranging from national standardised tests to ongoing daily formative assessment in the classroom. At the national level, full-cohort educational progress national tests are conducted in Grade 4, the results from which are used as key performance measures towards national goals. These are low stakes for schools, teachers and students. Summative assessment is based on a mix of teacher-based classroom assessments and national examinations. The latter take place at the end of both the second and third cycles of basic education (Grades 6 and 9), in Portuguese language and mathematics, and in secondary education in the last year of each subject (Grade 11 or 12). However, teachers hold most responsibility for summative assessment as the weight of national examinations is never dominant for the final mark. In the first cycle of basic education (Grades 1-4), assessment is generally informal and formative and results are reported in a descriptive and qualitative format. In the second and third cycles (Grades 5-6 and Grades 7-9), the emphasis on formative and internal assessment continues but summative results are reported on an ordinal scale from 1 to 5 and there are external examinations at the end of each of the cycles. The emphasis on formative assessment remains in secondary education, but greater attention is given to both summative and external assessment. A major asset is that student assessment is seen as part of the professional role of teachers in Portugal. Other strengths include the external dimension to assessment; the increased focus on data and results; the provisions to meet the needs of a diverse student population; and the assessment innovations associated with the New Opportunities programme. However, considerable challenges exist in building effective student assessment approaches. These include the difficulty in translating formative assessment as a policy priority into effective practice in classrooms; the limitations of moderation of teacher-based assessment within and between schools; assessment being used simultaneously as a controlling mechanism and a learning process; the little use of data for analysis; and the absence of external assessment in some strands of secondary education.

This chapter focuses on the formative (assessment *for* learning) and summative assessment (assessment *of* learning) used to assess student progress and achievement in Portuguese schools within the context of the overall framework for educational evaluation. 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 also addresses the context for that assessment, the policy framework within which it takes place, the views of stakeholders on assessment processes and the tensions between the articulated aims of the assessment framework for education, and its application in schools and classrooms.

Context and features

A framework for student assessment

Three purposes are identified for the assessment of students in basic and secondary education in Portugal:

- To support the education process in such a way as to sustain the success of all students and permit the readjustment of school and class curricular projects, particularly with regard to the selection of methodologies and resources, in accordance with the students' educational needs.
- To certify the various forms of learning and competencies acquired by students at the end of each cycle and when they leave basic education, by means of internal and external summative assessments.
- To help improve the quality of the education system, and make it possible to take decisions that will improve it and promote greater confidence on its operation on the part of society (Ministry of Education, forthcoming).

These purposes represent a coherent balance between the professional, regulatory and accountability purposes of assessment found within most modern systems of student assessment. These purposes are supported by a programme of internal (conducted by teachers and schools) and external (conducted by agencies outside schools) assessment supported by the Office for Educational Evaluation (GAVE) in the Ministry.

The framework for assessment is drawn from all the subject areas studied by students in the various cycles of education. Educational progress tests and national examinations also draw on the Common European Framework of Reference for foreign languages, the PISA programme, and international practice on indicators and performance-level descriptors as appropriate.

Currently under development, it is envisaged that the Learning Goals project will also form part of the framework, as these goals for student learning will be supported by appropriate assessment strategies. This project aims to provide the school system with a consistent and coherent set of standards so that student progress can be monitored (Ministry of Education, forthcoming).

This emphasis on coherence is both aspirational and pragmatic in origins. The guiding principles underpinning the education system include an assertion that assessment should be integral to all curriculum development and should have a central role in the regulation of teaching and learning. Such an integrated assessment system is a policy goal shared across many countries. From a pragmatic perspective, the Learning Goals project is also a response to pedagogical, structural and administrative differences between the different

cycles of schooling, especially between the first and the other two cycles of basic education (Fernandes, 2009). A further pragmatic function of a structuring framework is as a response to the extension of compulsory schooling to 12 years and a consequential broadening of the range of educational programmes on offer in the system as the cohort becomes more diverse.

Features of the student assessment system in basic education

In the first cycle (Grades 1-4), assessment is generally informal and formative. Summative tests of educational progress (*provas de aferição*) take place at the end of the cycle in mathematics and in Portuguese language. These are provided and marked for all students by the Ministry but are low stakes for schools, teachers and students (see Chapter 6). All other assessment in this cycle of education is internal and the responsibility of the teachers and the schools. Teachers and schools also agree and publish a set of assessment criteria for each subject in the curriculum. These form the basis of formative and summative assessment. While these are based on national guidelines, schools have some flexibility in the weightings they can assign to different components based, for example, on community or local needs. These criteria are used to support internal summative assessment and school-level decisions about student progress to the next grade (repetition of grades is possible for students who cannot progress, with the exception of Grade 1), or about additional support that may be required. Results are reported in a descriptive and qualitative format.

This emphasis on formative assessment is teacher-focused in orientation. The information provided by this continuous assessment is to assist teachers in the design and management of appropriate curricula, and in connecting the learning to the students' characteristics and the forms of learning and competencies they are required to develop. Formative assessment also serves diagnostic purposes in that it generates useful information for teachers about the learning progress of students and allows for programmes to be differentiated and adjusted to match student learning needs.

The external summative tests of educational progress at the end of the first cycle give rise to system-, school-, class- and student-level data which can be used for school self-evaluation and for teacher planning. These results are reported to parents, teachers and schools (see Chapter 6).

The assessment scenario in the second (Grades 5-6) and third cycles (Grades 7-9) is somewhat different for three reasons. First, summative results are reported in a descriptive manner but an ordinal scale from 1 to 5 is also used. Second, because the curriculum at these stages is organised by subject interdisciplinary categories (in the second cycle) and subjects (in the third cycle), the student is taught by more than one teacher. This means that the class board of teachers is involved in arriving at the internal summative marks. The third particular feature of the second and third cycles is that national examinations are taken by all students at the end of both Grade 6 and Grade 9 in Portuguese language and mathematics. The national examinations for Grade 6 students were introduced in the 2011/12 school year and replace equivalent national educational progress tests which were organised for about a decade.

In both the second and third cycles of basic education, the emphasis on formative and internal assessment continues. Internal summative marks are given, but one of the purposes of that summative assessment is to determine whether the student meets the requirements to take the national examinations in Portuguese language and mathematics. These examinations are marked on a five-level scale and they count for 25% and 30% of

the final classification in each subject respectively in the second and third cycles, with the remaining 75% and 70% coming from the schools' internal assessment.

In summary, internal formative assessment is a strong feature of the basic education system, with a focus on providing information to teachers, schools and students on student performance. Summative assessment is used to inform parents and other agents. In both formative and summative processes, teachers are at the centre, and are the main agents of the assessment process, with the exception of the national examinations at Grades 6 and 9. Teachers and schools have considerable autonomy to define assessment criteria and to design internal tests and tasks (Fernandes, 2009).

Features of the student assessment system in secondary education

The emphasis on formative assessment continues into secondary education, although greater attention is given to both summative and external assessment in this phase than in the earlier cycles. In secondary education, the purpose of assessment is to gauge each student's knowledge, competencies and capabilities and to verify the extent to which s/he has achieved the overall objectives for this level of education and the subjects and other components that comprise it. Of note, education for citizenship and Portuguese culture are cross-cutting themes assessed through all the courses (Ministry of Education, forthcoming).

The programmes on offer in secondary education are diversified into scientific and humanities courses to prepare students for progression into higher education, and technological courses which have a more vocational focus. There are also artistic programmes and professional/vocational courses available (see Chapter 1). This diversification is significant for assessment, in that the programme of national examinations is associated only with the scientific and humanities courses. Students in other courses who wish to progress to higher education have to take the national examinations required as entrance tests for the courses they want to take. The opportunities for non-traditional students under the New Opportunities programme (see Chapter 1) for those who may be at risk of dropping out of school or who wish to complete compulsory education are also significant in that many of these are offered in mainstream public basic and secondary schools. Some of these courses are led by teachers who work or had previously worked in secondary education, and all are associated with the forms and processes of assessment generally associated with the vocational sector.

Formative assessment in secondary education is continuous and systematic and it also serves a diagnostic function similar to assessment in basic education. The objectives of internal and external summative assessment in secondary education are classification and certification. Students are marked three times in each year, on a scale of 1-20 with the third mark in each year determining whether the student will be admitted to the corresponding national examination and which is taken in the final year of each subject. A mark of ten is required to take the national examination. In terms of the final classification of students, the internal summative assessments for which teachers have responsibility carry 70% of the weighting, with the national examination making up the remaining 30%. However, for the purposes of entry to higher education, the balance is different, with the school-based component reduced to 50% and the external examination increased to 50%. As for basic education, the national examinations are the responsibility of the Ministry.

The Ministry also makes a series of optional intermediate tests available for use by schools to support the assessment process – these are available for all levels of education even if with greater prominence in the third cycle of basic education and in secondary education. Their purposes are to give teachers insight into the national standards, to give students a sense of their own progress, and to help students to prepare for external assessments. The tests are centrally marked, and the results returned to schools. A particular emphasis in these intermediate tests has been placed on mathematics and the natural sciences. Access to these is provided on line by the Office for Educational Evaluation (GAVE). Use of these is optional, but widespread. The majority of schools use them as summative assessment instruments.

In summary, in secondary education, the teacher’s role in the assessment process is further developed to include the classification and certification process. As in basic education, the teacher appears to be the primary agent in the assessment process, even when national examinations are applied. Of note are the development and spread of intermediate tests and their increasingly summative function, as well as their organisation and marking on an external basis.

Strengths

Assessment is seen as part of the professional role of teachers

From the beginning to the end of schooling, assessment of students is seen as integral to the work of teachers. This is not just the view of the teachers themselves – the review team was struck by how widely this conviction is shared by students, school leaders and parents. While external assessment is available in the form of end of cycle educational progress tests, intermediate tests and national examinations, only in the examinations used for progression to higher education is this external assessment given equal weighting to the assessment of teachers. On all other occasions, the teacher assessment is given greater weighting.

Further, this school-based, teacher-led assessment is based on tasks and tests developed by teachers and schools and informed by assessment criteria that are developed by the pedagogical councils or class boards within each school. Schools have considerable autonomy in the decisions about student progress and certification (Fernandes, 2009).

This apparent autonomy in assessment for teachers and for schools is uncontested and widely supported. While there is an appeals system available for the secondary phase of education, and while it is routinely availed of, the appeals can be made on the basis of a procedural failure only. Substantive assessment judgements cannot be the subject of appeals. This support for the assessment work of teachers is particularly noteworthy given recent controversies around the appraisal of teachers in Portugal discussed elsewhere in this report.

The centrality of teachers in the assessment process, and the support for this teacher agency from inside and outside schools are particular strengths of the Portuguese assessment system. A consequence of that positioning of teachers at the heart of the process is the emphasis on *formative* assessment, an emphasis that is shared in many countries.

The concept of formative assessment is open to a variety of interpretations in assessment policies across the systems of individual countries, including Portugal.

Despite some contestation around meaning, there is a strong commitment at all levels of the Portuguese education system to formative approaches and to maintaining this characteristic feature of the assessment system in any future development. The intermediate tests, though externally set and externally marked, are also designed with teachers' assessment practice in mind – they are designed to give teachers an insight into the national standards, as well as to give feedback to students on their progress.

Student assessment also has an external dimension

One of the challenges faced by any system committed to internal, formative and teacher-led assessment is the need for checks and balances across the system to ensure reliability in the application of standards and criteria and to gather system-wide data for the purposes of evaluating system quality. The provision of end of cycle wholly externally marked educational progress tests at the end of the first cycle, of national examinations at Grades 6 and 9 and of a series of external components for subject-based examinations in the secondary cycle represents a considered attempt to address this challenge. The capacity of the Office for Educational Evaluation (GAVE) in how it approaches and conducts its work is an important feature of the quality processes. The optional but very popular intermediate tests are also significant in this regard; however the growing pressure on GAVE to generate more external tests more often and the shortage of expertise in psychometrics and assessment more generally in the education system should be noted.

Assessment innovations associated with the New Opportunities initiative

There have been a number of recent initiatives, under the umbrella of the New Opportunities programme, to extend the educational provision in schools to students who may have left school, or may be at risk of leaving school, and to adults who might not have completed compulsory education (see Chapter 1). These new programmes have been accompanied by the development of approaches to assessment focused on motivating students, giving high quality feedback, and including the active participation of learners in the assessment process. A key feature of these arrangements and approaches is their location close to the learning process and to the learner. Thus, assessment tends to occur immediately after the completion of a module or portion of a course rather than at the end of a year or cycle. The use of approaches beyond written tests, such as a performance assessment, puts the learner and learning at the centre of the assessment process. The need to use assessment to motivate learners to learn, rather than to simply engage with the assessment or test, is acknowledged as fundamental by those working in this sector. The challenge of this focus is also acknowledged.

Alternative programmes such as the Education and Training Courses feature these approaches to meet the needs of the particular groups of learners accessing this form of provision. In contrast to the assessment system for mainstream schooling which places teachers at the centre of the process, the assessment system for these alternative programmes, which are also internal and conducted by teachers, appear to place the learner at the centre. The progress of the learner is given priority over the gathering of system or school-level data.

A sophisticated information system has been developed to support the work of the New Opportunities Centres. This is more than an administrative tool. It supports the kinds of assessment used, allows for learning to be recorded, monitored and accredited, and provides data for the quality assurances processes used in this form of education. Of note,

in line with the policy in the rest of education, assessment is internal, conducted by teachers and instructors.

These approaches to assessment have been applied in vocational education and training contexts elsewhere in Europe (Green *et al.*, 2000; Tillema, 2003) and seem to be well supported in the settings in which they are now being used in the Portuguese system by the New Opportunities Centres and the National Agency for Qualification (ANQ, and recently recreated as ANQEP). Good assessment design, according to one widely used perspective, draws on three elements: a model of how learners represent knowledge and competence; tasks or situations that allow for performance to be captured or observed; and a means by which inferences can be drawn from evidence gathered from that performance (Pellegrino *et al.*, 2001). The validity and reliability of any assessment depends largely on the quality of the connections between these three elements, and the degree of attention paid to each.

In the system of assessment used in the New Opportunities Centres and in alternative provision more generally, considerable attention has been given to each of these elements in a balance that seems, at least initially, to be having a positive impact on participation and completion. In addition, the attempt to place the learner at the heart of an assessment process that is supported by a robust information system that enables monitoring data to be gathered is noteworthy.

The location of many of the New Opportunities Centres in mainstream secondary schools, or as part of clusters, is a further strength. First, it means that there is some engagement between the students availing of these programmes and secondary school students. Second, it means that, while there can be some tension between the flexibility of the newer programmes and traditional school culture and structures, there is some interaction between teachers working in the two systems with opportunities for professional engagement and exchange. While these opportunities remain to be exploited in full, and the challenges of supporting dialogue within the more individualised culture of schooling are widely acknowledged, they are particular strengths of the New Opportunities programme, and have much to offer the assessment system for all students in schools.

While these developments are relatively new, initial data are beginning to emerge about the value of the initiatives and their positive impact.

Meeting the needs of a diverse student population

In addition to the assessment developments with the New Opportunities programme, particular efforts are made to meet the needs of students whose mother tongue is other than Portuguese. Such students are supported through an initial diagnostic assessment of language proficiency using benchmarks based on the Common European Framework of Reference, and assessment criteria to support progress identified by the school. In national examinations at Grades 6 and 9, taking account of progress, some students are given the option of taking the Portuguese examination as non-mother tongue.

Students with special educational needs are also supported in the system of student assessment by a number of supportive measures. This is in the context of a high degree of inclusion of students with special needs in mainstream schools. Students with special educational needs benefit from Individual Educational Programmes (IEP) and do not participate in the national assessment programme but have personalised targets and assessment criteria associated with their IEP. Test and exam accommodations exist in line

with what students are normally allowed in the classroom but the consistency with which they are administered is unclear. For instance, for students with visual impairment, adapted assessment instruments are available.

In the context of special education, there is a risk that curricula and assessment frameworks may define achievement and progress too narrowly to capture many valuable areas of learning of students with special educational needs. Teachers may not always have the awareness and competencies to ensure adequate and innovative assessment of students with diverse needs and to report accordingly to parents. While in Portugal schools generally have among their staff special education teachers, there are some concerns that regular teachers have not received the preparation (including in their initial teacher education) to adequately respond to the needs of special education students. For instance, they may have insufficient knowledge of formative assessment and/or differentiated teaching to specifically meet the needs of those students. While inclusive assessment practice exists in many schools, the key challenge appears to be to ensure that such good practice is developed consistently in all schools enrolling students with special educational needs.

A focus on data and results

A further strength of the assessment system is the attention paid to data and results at school, regional and national levels, by teachers, parents and students, and by the general public. While this attention can also have negative consequences discussed elsewhere in this chapter, the focus on outcomes and results, and the commitment to collect and monitor data is noteworthy. The work of both the Information System Co-ordinating Office (MISI) and the Office for Education Statistics and Planning (GEPE) (recently integrated in the Directorate General for Education and Science Statistics, DGEEC) in gathering and disseminating system- and school-level data is an important resource for decision makers at every level of the system (see also Chapter 6).

At school level, data on student achievement are gathered and discussed by class boards and pedagogical councils. The review team was struck by the degree to which engagement with data and results was cited by so many teachers and others associated with the work of schools. Such engagement was seen as key to school development and improvement.

While the generation and collation of data are associated with an administrative task for schools, nonetheless, the commitment to data and evidence is admirable and widely regarded as important for school and system improvement (Earl and Katz, 2002; Eurydice, 2009; Levin, 2010).

Challenges

Formative assessment is stronger in policies than in classroom practices with some significant consequences for pedagogy and learning

Finding a balance between formative and summative assessment is a challenge shared by many education systems. It is made more complex by the wide range of understandings of the meaning of the term *formative assessment*, and the difficulty of managing the tensions between a stated commitment at policy and school level to an assessment process focused on students and their learning, and public, parental and political pressure for accountability in the form of scores and rankings (Harlen and James, 1997; Wiliam and Black, 1996; Newton, 2007). This pressure for summative scores, and

a conflation of formative and summative purposes in education policy documents in many countries has also resulted in confusion that in some cases may have hindered sound assessment practice, especially in the development of formative assessment (Harlen and James, 1997; Newton, 2007).

In some cases, and certainly in the case of Portugal, the distinction that has been drawn between formative and summative assessment has been effective in foregrounding the assessment work of teachers and the role of assessment in support of student learning. The emphasis placed on the formative role of assessment in basic and secondary education in Portugal and the centrality of teachers in that process is highlighted in all of the relevant policy documents, and in teacher, student and parent discourse (Ministry of Education, forthcoming).

However, the review team did not find that this formative emphasis was reflected to the same degree in classroom and school practice. On the contrary, in classroom and schools, the formative seems to be increasingly displaced by the summative and a focus on the generation of summative scores. While the attention to results and data is a positive feature of the system in Portugal, an over-emphasis on these may be having a negative impact and undermining the formative role of teachers and assessment so highly valued in policy goals.

Such over-emphasis generated particularly strong views in the course of discussions with the review team. The team heard about an obsessive attention to results, the drive for results skewing the education system, media hype around examination results, classroom practice dominated by examination and test preparation, non-compulsory tests gaining the status of compulsory tests, and the quality of teaching being equated to the quality of results. It is noteworthy that such phenomena are more usually associated with the backwash from high- or medium-stakes *external* assessment, but in the case of Portugal, they seem to be associated with apparently low-stakes assessment where teachers play a central role.

Two further challenges were identified in the course of these discussions, and supported by site visits. The first is that while the majority of marks for any grade are awarded by the teacher, these marks appear to be awarded for the same kinds of activity assessed through the external component. In the secondary cycle in particular, classroom practice over-emphasises preparation for traditional paper-based tests. The second challenge is that while the focus on the teacher as the assessment leader is a strength of the Portuguese system, it seems to have led to an under-emphasis on student and learner agency and an over-reliance on marks and grades in the feedback process to the detriment of rich feedback that can help learners to identify the next steps they need to take.

OECD (2005) defines formative assessment as the frequent assessment of student progress to identify learning needs and adapt teaching. However, there is debate over whether assessment is truly formative unless it involves the student, and unless feedback to the student on his/her progress features in the process (Shepard, 2000). Assessment is not truly formative, it is also suggested, unless the evidence gathered is used to inform subsequent instruction. Feedback in Portugal tends to be focused on test performance and results rather than on learning. This applies to feedback from both internal and external assessment. Such feedback on test scores is an important component of learner feedback, but only in the context of broader evidence of learning and feedback on that learning. The absence of this broader focus reflects a tendency to adopt a more behaviourist approach to pedagogy, and little engagement with constructivist principles.

In the course of the Review, the issue of low levels of student motivation was raised, along with concerns about the ability of some students to learn independently. However, in general, these were presented as attributes of the students, rather than as consequences of practice. Research indicates, however, that these behaviours are often associated with particular approaches to teaching and assessment. If the focus of classroom activity is on the test, rather than on the learning process, then teachers and students may emphasise a performance orientation over a mastery orientation (Dweck, 1986; Middleton and Midgely, 1997). When getting a good mark on a test becomes the goal of classroom activity, ironically, it can be at the price of learning. Pursuing a test performance can work against students' engagement and persistence in learning, self-regulation, learner autonomy, and motivation. It also leads to a transmission style of teaching supplemented by the rehearsal of test performance.

Such an orientation is not surprising given the admission by those working in policy, research and initial and continuing teacher education that assessment and motivation theories are given little attention in the various dimensions of the Portuguese system.

There is further debate among assessment researchers as to whether feedback should also include information for the learner not just on how they have done, but on the steps they need to take to improve or to progress further as agents of their own learning (Butler and Winnie, 1995). The pedagogical implications of interactive feedback have only recently begun to be considered in educational research and debate but are generally recognised as significant but under-emphasised (Black and Wiliam, 2009). In Portugal, these pedagogical implications have yet to be explored at policy or school level; neither are they considered in initial teacher education.

Moderation is focused on procedures rather than on outcomes of learning or shared understandings of standards

One of the challenges faced by all systems of student assessment that rely heavily on the judgement of teachers and schools is maintaining the quality of moderation and improving the processes by which standards across schools are assured to be assessed consistently. Some attention is given to this in the Portuguese system with a system of checks and balances in place once the teacher has scored the test and produced a sheet of results as the outcome of an assessment event (Ministry of Education, forthcoming). There is checking that the procedures have been followed and the criteria applied correctly. However, this process does not include discussion or analysis of student work, across classes in schools, across schools, nor at national level. Examples of what is expected (except in numeric terms, with targets set for acceptable percentages at each classification level) are not available. In first cycle schools, the review team encountered some evidence of such sharing of examples of student work, but little was found in the second or third cycle, and none in the secondary system, nor any concern expressed that it was not available.

Because the moderation is focused on results and on the application of procedure, there is little focus on teacher judgement, and none on the work of students, to arrive at a particular score or mark. Thus, the “product” of learning is seen as a grade or a test score rather than any “real” work (Torrance and Pryor, 1998). An emphasis on results is not without its merits. But such numbers are a *representation* of achievement, and if too much attention is paid to the representation rather than to the real evidence of achievement, then the entire system can become oriented on the “score” rather than on the “game”. Earlier in this chapter the model advocated by Pellegrino *et al.* and others for

quality assessment was discussed. The third element of that quality model – an interpretation method for drawing inferences from the evidence gathered – is of note for this discussion. The process by which inferences are drawn is not discussed by teachers, nor is there any consideration of the interpretation involved. This has serious implications for the validity of the internal assessment process – a process which, given its location in schools and the role of teachers in its design, should be highly valid.

This phenomenon goes well beyond the widely documented backwash effect of high-stakes external testing on teacher behaviour. Rather, a combination of internal assessment, published criteria and public interest in results can lead to what Torrance (2007) has described as assessment as learning – a system where the achievement of the criteria becomes the focus of learning, and an over-reliance develops on the accumulation of marks and the documentation of competence and a narrowing of the learning experience for students and teachers.

In Portugal, the focus on clear criteria is an obvious system strength. These are locally developed and shared with students and guardians. But this unwavering focus on clarity and on the procedures by which the criteria are used in schools has led to a lack of attention to the moderation of standards between classes in the same school and between schools. In Portugal, moderation processes focus on results only to ensure comparability of grade distribution, but discussion of what is represented by the grade awarded – examples of student work at each grade – are rarely considered in schools, and never between schools nor with the public at large.

This over-emphasis on what teachers in one school called the “daily proliferation of indicators” may create a sense of security and comfort, but it can also act as a barrier to looking at “real” learning in “real” classrooms. Engagement between teachers may also be reduced to a consideration of procedures and the application of the criteria, rather than any collegial discourse on successful teaching and learning. What emerges from this engagement is an atomised account of teaching and learning, carefully mapped onto criteria and standards. Some of those who discussed this issue with the review team warned that the benchmarks for the new Learning Goals project might suffer the same fate with a focus on the goals themselves being lost in the pressure to measure and report on progress.

Assessment is used as a controlling mechanism as well as a learning process

The autonomy given to schools to specify and publish their own assessment criteria within national guidelines is used by many schools to include criteria for student behaviour and discipline within this framework. Student behaviour was mentioned to the review team as a challenge in a number of schools.

The monitoring and reporting of student behaviour should be part of the day-to-day activity of schools; these tasks form part of an overall student support strategy in any school to ensure, in the first instance, that good behaviour and engagement are rewarded and that poor discipline can be identified early and a speedy intervention made (Task Force on Student Behaviour in Second Level Schools, 2006). However, the inclusion of a rating or mark for discipline or behaviour as part of the teacher-led internal assessment that contributes to the final academic mark is not found in international practice and seems to breach one of the fundamental principles of sound assessment practice to avoid combining achievement in different domains in the same assessment event or process. There is also a risk that it can contribute to some confusion in students about the purpose of assessment in schools and to what has been termed self-level feedback, rather than

feedback focused on the task, or on progress, or on achievement, with consequent implications for student self-efficacy (Stobart, 2001).

It should be noted however, that in none of the meetings during the visit was the inclusion of a non-academic dimension in summative assessment raised as an issue, a challenge or a concern. However, the inclusion of this aspect of student-school engagement in the academic system adds further to the already uncertain messages about assessment and its function discussed above.

Data are gathered, but not subject to interrogation or secondary analysis

The review team was struck, on the one hand, by the commitment to gathering evidence of achievement through marks and scores, but on the other, by the lack of robust interrogation of that data beyond quality assurance checks to ensure that procedures had been followed. Teachers beyond the first cycle seemed reluctant to engage in discussions about the quality of student learning while giving considerable time and commitment to the scores and marks awarded. Such reluctance may be understandable given the lack of professional development opportunities for teachers on assessment and pedagogical issues, and the low priority given to these areas in initial teacher education.

The data on student achievement in tests – whether generated externally or internally – were generally not subject to any secondary analysis. No further investigation was pursued at school or cluster level to monitor the progress of particular groups of students within the cohort, or to check if different rates of progress were apparent over time for different groups of students, in different subject areas for example. This seemed at odds with the policy commitment to an inclusive system and some initiatives to meet the needs of a widening range of learners.

Some strands of secondary education are not supported by external assessment

Since 2006, external assessment has been applied only to the scientific and humanistic courses, and not to technological, professional/vocational, artistic or other specialised courses. Successful completion of these latter courses is not subject to any national assessment, but to local, internal assessment only. Of itself, this should not present any particular challenge to the quality of a system of student assessment. However, in the Portuguese context it gives rise to two concerns. The first is that given the reservations expressed about the quality of moderation in the discussion above, some questions arise about the reliability of assessment that is completely internal. The second is that in light of the fact that external assessment is expanding across the school system (intermediate tests for example), not including some element of external assessment in certain elements of provision in secondary education seems to signal that reliability matters less for these courses and these students – traditionally, students who do not progress to higher education.

Policy recommendations

Develop the quality of feedback to learners to strengthen formative assessment practice

The challenge of ensuring educational initiatives and reforms have an impact on the experiences of students in classrooms is an enduring one in education policy. Education systems tend to focus on change as an end in itself, with a proliferation of initiatives on a conveyor belt of reforms that just keeps moving (Harris, 2011). The challenge to focus on

where it really matters – classrooms – is perennial and global. The temptation to focus on structures, systems, procedures, programmes and resources is a strong one but, while these components of policy are important, changing any or all may not impact on the quality of student learning. Classrooms are hard to reach. A commitment to formative assessment on paper needs to be matched with engagement with formative assessment practices in classrooms and schools.

Building on the culture of evaluation, and the centrality of the teacher in the assessment system in Portugal, greater focus on a culture of feedback on student learning would deliver a number of wins for the system. First, it would serve to highlight the role of the learner in the learning process, and encourage a greater sense of agency and responsibility in learners of all ages in the school system. The lack of such agency was a concern reported by many teachers, school leaders and educational administrators during the review visit. Assessment that is truly formative, which includes formative feedback, allows for the role of learners and their peers in the learning process to be clarified. Traditionally, teachers have been regarded as responsible for establishing where learners are in their learning, where they are going, and what needs to be done to get them there. It is important to flag the role of the learner in all of these processes; learning is a shared responsibility (Black and Wiliam, 2009). This message is as important for learners as it is for teachers.

Second, a greater focus on rich feedback would support classroom and school discussions on the quality and process of student *learning*, to accompany the current widely supported emphasis on marks and *results*. Third, it might go some way towards engaging the public and the media on the outcomes of education beyond a numeric mark. Finally, extending such a focus to external assessment arrangements might usefully counterbalance the drive to add further to external assessment, by developing richer feedback on current external assessment already in place. Those sectors of secondary education without any external assessment at present might serve as a useful developmental space for an approach to rich feedback. This point is discussed further below.

The New Opportunities Centres and the work of the National Agency for Qualification (which recently became ANQEP) offer a starting point for this work as they already provide support for student feedback and encourage learner participation in discussions about progress and achievement in the particular programmes provided under their auspices. Sharing this practice widely across the system would be a useful first step. Improving the theory and practice of student feedback should also be a focus of work in initial teacher education, and an emphasis in continuing professional development.

Use the Learning Goals project to focus on examples of student work to support better moderation and greater teacher professional collaboration around student learning

The Learning Goals project and the intermediate tests offer some potential in developing a focus on feedback for learning and on exemplification of student work in support of that focus. The benchmarks and indicators that are being developed to support the Learning Goals project could be extended to include real examples of student work to illustrate expectations at the different levels, with student and teacher commentary. Reporting on progress towards the Learning Goals might incorporate similar examples, as well as, or even as a replacement for, numeric scores.

The concerns expressed by Portuguese policy makers and those working in system evaluation that the articulation of targets (as proposed in the Education Programme 2015) and reporting on progress had become a matter of routine without inspiring those working in the system to use the targets to support practice are not unique to Portugal. Data are good, but more data is not always better even if it is tempting for all working in education to believe this to be the case (Sahlberg, 2011). A useful distinction can be made between evidence and data and on what is required to work with both. Hargreaves and Shirley express a concern that policies which have “stapled teachers to their spreadsheets, kept them calculating and concentrating on tested achievement gains instead of inspiring animated professional discussion about students and their learning” (2009, p. 92). What is needed, they suggest, is that teachers engage with *evidence* as well as with data.

Work on evidence is already underway in the project promoted by the Office for Educational Evaluation (GAVE) to share good practice in student assessment. However, the Learning Goals project, with its particular emphasis on student learning and the EU key competences, offers a new context in which to develop a focus on showing what students know and can do at each level of education, not just to teachers and education professionals, but to the general public. Suspicions that this new project will simply graft another layer of indicators onto a system already bearing a heavy load would be allayed if a fresh approach were taken to benchmarking using real examples of student work.

The project already offers much to the system in terms of coherence and consistency in connecting the different cycles of the school system. The current aims of the Learning Goals project are focused on providing the school system with a consistent and coherent learning standards supported by indicators (Ministry for Education, forthcoming). Extending those aims, and the project deliverables, to include exemplification and evidence would connect those indicators more closely to practice and to classrooms. It would also give them a public interest.

Generating and sharing such evidence at school level might also be the basis for shaping the many meetings of teachers about results and grades towards genuine professional learning communities which, when they work effectively, can improve the quality of moderation and reduce in-school variation in student outcomes (Harris, 2011). Engaging with evidence might also ensure a better balance across the three strands of quality assessment as described by Pellegrino *et al.* (2001) and cited earlier in this chapter.

Given the particular attention paid to ICT in the education system in recent years, the use of online environments for sharing examples of student work to support engagement with questions of standards and processes of moderation should be explored both from a public access and a cost-effectiveness perspective.

Ensure that assessment is relevant and responsive to students with special educational needs

Special education challenges many common assumptions about student assessment. The presence of students with special educational needs in mainstream classrooms provides opportunities for teachers to be innovative and creative in the design of teaching and assessment approaches. The key challenge is to ensure that rich assessment opportunities are systematically offered to all students with special educational needs regardless of where they go to school. This calls for a better preparation of all teachers to respond to the needs of students with special educational needs, with inclusive assessment further included and developed in both initial education and professional development for

teachers. In the same way, test accommodations for special education students need to become systematic across the school system. The Individual Educational Programmes (IEPs) offer a framework to identify strengths and interests of students with special educational needs and set specific and manageable goals for development. The IEP should bring together the school, parents, student and possibly other agencies around the basic processes of assessing, objective setting, teaching, monitoring, evaluating, reassessing and further planning to support the learning of the student.

New Zealand provides an example of dedicated resources for the assessment of students with special needs. The New Zealand Ministry of Education is supporting innovative approaches to assessment and reporting for diverse students and has launched a project on Assessment for Learners with Special Education Needs, which includes the development of “narrative assessment” exemplars, guidance, and resources. Two key resource documents, *Narrative Assessment: A Guide for Teachers* and *The New Zealand Curriculum Exemplars for Learners with Special Education Needs*, are available to support teachers in maximising learning opportunities and pathways for children with special educational needs within the New Zealand Curriculum (see <http://assessment.tki.org.nz/Assessment-tools-resources/Assessment-tools-for-Learners-with-Special-Education-Needs>) (Nusche *et al.*, 2012).

Consider means of assessing student discipline other than through the inclusion of a mark for student behaviour in the academic mark

The inclusion of student discipline and behaviour in the academic mark as a matter of routine sends confusing signals as to the purpose and focus of assessment in schools. As part of any effort to place a greater emphasis on student learning, this practice might be replaced by other means of recording and reporting on student behaviour at school level, and across the system as a whole as a means of monitoring the well-being and engagement of students at school. Consultation with students on this matter might be useful.

This is not to say that the affective dimensions of learning such as persistence, application to challenging problems, and resilience in the face of failure for example, should be beyond assessment or excluded from feedback. These are learning behaviours and should be part of any system of student assessment. Given that so much of student assessment is internal, including these aspects of learning should be a matter of routine, and should be part of the rich feedback provided to students. Moving away from the inclusion of discipline and behaviour in the award of marks will need careful management and should be seen as part of a set of measures to place a greater focus on learning system-wide while finding more appropriate means to support positive student behaviour in school.

Build on the attention paid to results to generate richer data

The attention paid to results, the strong concern to promote an inclusive education system, and the already robust data collection systems in place offer a unique opportunity to move towards richer data for school and system use. This would align Portugal with some of those countries widely recognised as being at the forefront of education innovation (McKinsey & Company, 2007, 2010). The recent development of mega-clusters has created developmental sites where some approaches to richer data could be developed without incurring the costs associated with whole-system developments. As discussed above, more data does not necessarily deliver better systems. Aiming for a

balance between rich data and evidence would be a useful policy goal. It would further allow for Portugal to develop the system that would best suit its needs and its schools. It would also allow for approaches to be trialled, and for existing expertise within the clusters to be harnessed for local school and community benefit and for national insight.

At several meetings with schools and clusters, teachers and school directors spoke enthusiastically about the data collection systems they had in place but none had moved to anything beyond rudimentary analysis (distribution of scores, comparisons with previous cohorts for example). Yet schools also held data on students' family circumstances, migrant status and socio-economic situations. These data were not connected with academic achievement data so no information was available on differential performance of different groups within the school cohort. For some schools serving poorer communities, this was a source of frustration as their school results were compared with the results of schools drawing from more affluent communities, and the generation of comparison tables by the media resulted in some legal guardians, especially in urban areas, moving their children to what were perceived as "better" schools (see also Chapter 6).

Consultation between GAVE, GEPE (now integrated in DGEEC), teachers, school directors and parents on such an initiative, particularly in clusters where such information platforms might be developed, would be important and would support schools in using the data, along with evidence (see above) to inform the education programme on offer and plan for improvement. The balance between data and evidence is important to restate if Portugal is to avoid the pitfalls experienced by other education systems of burying the education system in continuously collected data (Sahlberg, 2011).

Fernandes (2009) in his evaluation of the Portuguese system concluded that the elements of a high quality student assessment system were in place; what was needed, he suggested, was a better balance between them so that the system might focus more on processes of teaching and learning which he identified as the greater challenge in the system.

Make greater use of external assessment as support for moderation and teacher professional development

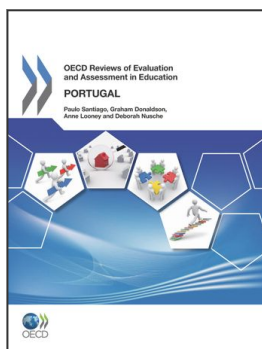
The intermediate tests, and the development of some external assessment for the components of secondary education not currently subject to any national examinations offer some opportunities for developmental work by GAVE on modelling richer assessment, reporting and feedback practice for teachers and sharing the assessment expertise developed by that agency over time. Discussions of standards, examples of student work (as discussed above), examples of performance assessment (even if only used on a sample basis) might be shared through the online environment already used to good effect for the dissemination of some tests. The intermediate tests seem to be feeding a drive for data to the detriment of evidence. This drive is also contributing to a number of trends identifiable in the system, more generally found in systems dominated by high- or medium-stakes external assessment. Sharing examples of classroom success in tests with teachers and students, developing and sharing commentaries on examples of work collected in the tests and from teachers in schools might feature in the next phase of development of these tests.

Given that teachers' professional competence in assessment and feedback was identified as a system challenge, providing reliable high quality material as the basis for professional dialogue and collaboration at school level will be important.

References

- Black, P. and D. Wiliam (2009), “Developing the Theory of Formative Assessment”, *Educational Assessment Evaluation and Accountability*, Vol. 21, No. 1, pp. 5-21.
- Butler, D. and P.H. Winnie (1995), “Feedback as Self-Regulated Learning: A Theoretical Synthesis”, *Review of Educational Research*, Vol. 65, pp. 245-281.
- Dweck, C.S. (1986), “Motivational Processes Affecting Learning”, *American Psychologist*, Vol. 41, pp. 1040-1048.
- Earl, L.M. and S. Katz (2002), *Leading Schools in a Data-Rich World*, in K. Leithwood, P. Hallinger, G.C. Furman, P. Gronn, J. MacBeath, B. Mulford and K. Riley (eds.), *The Second International Handbook of Educational Leadership and Administration*, Kluwer, Dordrecht, Netherlands, pp. 1003–1022.
- EPPI (Evidence for Policy and Practice Information and Co-ordinating Centre) (2002), *A Systematic Review of the Impact of Summative Assessment and Tests on Students’ Motivation for Learning*, Review conducted by the Assessment and Learning Synthesis Group, EPPI, London.
- Eurydice (2009), *National Testing of Pupils in Europe: Objectives, Organisations and Use of Results*, European Commission, Brussels.
- Fernandes, D. (2009), “Educational Assessment in Portugal”, *Assessment in Education: Principles, Policy and Practice*, Vol. 16, No. 2, pp. 227-247.
- Green, A., A. Wolf and T. Leney (2000), *Convergence and Divergence in European Education and Training Systems*, Institute of Education, London.
- Hargreaves, A. and D. Shirley (2009), *The Fourth Way: The Inspiring Future of Educational Change*, Thousand Oaks, California.
- Harlen, W. and M. James (1997), “Assessment and Learning: Differences and Relationships between Formative and Summative Assessment”, *Assessment in Education: Principles, Policy and Practice*, Vol. 4, No. 3, pp. 365-379.
- Harris, A. (2011), “Reforming Systems: Realizing the Fourth Way”, *Journal of Educational Change*, Vol. 12, No. 2, pp. 159-171.
- Levin, B. (2010), “The Challenge of Large-Scale Literacy Improvement”, *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, Vol. 21, No. 4, pp. 359-376.
- McKinsey & Company (2007), *How the World’s Best Performing School Systems Come out on Top*, McKinsey & Company.
- McKinsey & Company (2010), *How the World’s Most Improved School Systems Keep Getting Better*, McKinsey & Company.

- Middleton, M. and C. Midgely (1997), “Avoiding the Demonstration of Lack of Ability: An Under-Explored Aspect of Goal Theory”, *Journal of Educational Psychology*, Vol. 89, pp. 710-718.
- Ministry of Education (forthcoming), *Country Background Report for Portugal*, prepared for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, www.oecd.org/edu/evaluationpolicy.
- Newton, P. (2007), “Clarifying the Purposes of Educational Assessment”, *Assessment in Education: Principles, Policy and Practice*, Vol. 14, No. 2, pp. 149-170.
- Nusche, D., D. Laveault, J. MacBeath and P. Santiago (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD, Paris.
- OECD (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD, Paris.
- Pellegrino, J.W., N. Chudowsky and R. Glaser (eds.) (2001), *Knowing What Students Know*, National Academy Press, Washington D.C.
- Sahlberg, P. (2011), “The Fourth Way of Finland”, *Journal of Education Change*, Vol. 12, No. 2, pp. 173-185.
- Shepard, L.A. (2000), “The Role of Assessment in a Learning Culture”, *Educational Researcher*, Vol. 29, No. 7, pp. 4014.
- Stobart, G. (2001), “How Can Educational Assessment Improve Learning”, paper presented to the 27th Annual Conference of the International Association for Educational Assessment, Rio de Janeiro, May.
- Task Force on Student Behaviour in Second Level Schools (2006), *Schools Matter: Report of the Task Force on Student Behaviour in Second Level Schools*, Department of Education and Science, Dublin.
- Tillema, H.H. (2003), “Integrating Developmental Assessment with Student-Directed Instruction; A Case from Vocational Education in the Netherlands”, *Journal of Vocational Education and Training*, Vol. 55, No. 1, pp. 113-128.
- Torrance, H. (2007), “Assessment as Learning? How the Use of Explicit Learning Objectives, Assessment Criteria and Feedback in Post-Secondary Education and Training Can Come to Dominate Learning”, *Assessment in Education: Principles, Policy and Practice*, 14:3, 281-294.
- Torrance, H. and J. Pryor (1998), *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom*, Open University Press, Philadelphia, PA.
- William, D. and P.J. Black (1996), “Meanings and Consequences: A Basis for Distinguishing between Formative and Summative Functions of Assessment?”, *British Educational Research Journal*, Vol. 22, No. 5, pp. 537-548.



From:
**OECD Reviews of Evaluation and Assessment in
Education: Portugal 2012**

Access the complete publication at:
<https://doi.org/10.1787/9789264117020-en>

Please cite this chapter as:

Santiago, Paulo, *et al.* (2012), "Student assessment", in *OECD Reviews of Evaluation and Assessment in Education: Portugal 2012*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264117020-6-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.