



Chapter 2

PROCEDURES FOR TEACHER APPRAISAL

This chapter considers the features of different approaches to teacher appraisal, i.e. the mix of instruments used in different teacher-appraisal models, including self-appraisal, classroom observation and portfolios, and the elements of performance that are appraised, including planning and preparation, the classroom environment, teaching approaches and contributions to school development. The chapter also examines whether appraisal is voluntary or mandatory, and the frequency with which teachers are appraised.



Teacher appraisals may be mandatory or voluntary, depending on the purpose of the appraisal process. The frequency of appraisal also varies. It can be carried out periodically at regular intervals, at key stages during a career, or on specific occasions, such as at the end of the probationary period or to renew a contract.

Table A.1 in Annex A provides an overview of country practices. Teacher appraisal at the end of a probationary period is usually mandatory. In Ireland, there are two appraisals during the probation period in primary schools, and an appraisal is conducted at least three months before the end of the probationary contract in secondary schools. In Austria, the appraisal is ongoing throughout the one-year probationary period and varies in frequency. In the Slovak Republic, individual schools determine when the end-of-probation appraisal is conducted. In most countries with a specific probationary period, probation lasts one year.

Regular teacher appraisals for performance management are mandatory in all countries that conduct such appraisals, and are usually conducted every few years. In Austria and the Czech Republic, however, appraisals are conducted at the discretion of the school director or the school board. In the Czech Republic, an appraisal may also be conducted in relation to a decision on employment status. In Canada, appraisals may occur at the discretion of the school leader if there are concerns about performance. Regular appraisals for performance management are conducted every year in Korea, Mexico, New Zealand, Poland, the Slovak Republic and Slovenia. They are usually conducted annually in Australia, but there are some variations between states and territories. Education ministers in Australia have agreed to a national framework for teacher performance and development. The framework, which will begin to be implemented in 2013 at a pace and method determined by states and territories, requires a formal review at least once a year. Appraisals for performance management are conducted every other year in Hungary and Portugal, every third year in Israel and the Netherlands (ISCED 2-3), and every fourth year, at least, in the Flemish Community of Belgium, Chile¹ and the Netherlands (ISCED 1). In Canada, experienced teachers are formally appraised every five years. In France, the frequency of appraisal for performance management depends on the status of the teachers and whether they work in the public or private sector. The frequency varies between once a year and once every three years, and may also occur at the initiative of the teacher or in the case of problems. Teachers in France are also appraised by inspectors on an irregular basis (usually every three to four years at ISCED 1 level and every six to seven years at ISCED 2 level).

Appraisal procedures for registration vary among the countries where such processes exist. In New Zealand and Australia, the appraisal is conducted at the end of a “registration period” while in Austria, annual renewal is mandatory. In Sweden, appraisal for registration occurs once at the end of an introduction period and may be conducted again in relation to decisions on employment status, at the discretion of school leaders or at the discretion of the National Agency for Education.

In Chile and Mexico, teachers may voluntarily apply for appraisals that are linked to reward schemes. In Chile, there is also a mandatory, annual appraisal of groups of teachers (i.e. teachers in individual schools), the National Performance Evaluation System (SNED). In Korea, appraisals related to the performance-based incentives system are mandatory and are conducted every year. The period of evaluation for incentives runs from 1 January to 31 December. Incentives are announced at the end of the calendar year and payments are made in the following year.

ASPECTS APPRAISED

Determining the domains to be appraised helps to define the core responsibilities of teachers and the kind of performance that is valued as “good teaching”. Danielson’s *Framework for Teaching* (1996, 2007) groups teachers’ responsibilities into four major areas: planning and preparation; instruction; the classroom environment; and professional responsibilities. Tables A.1, A.2 and A.3 in Annex A show that planning and preparation, the classroom environment and instruction are among the elements countries most frequently mention as key areas for teacher appraisal. Other frequently appraised aspects include the professional development undertaken by teachers, teachers’ contributions to school development, and links to external partners or the community. These elements can be seen as part of teachers’ broader “professional responsibilities”.

Since a teacher’s work involves considerably more than the pedagogical activities associated with student learning, it is appropriate that teacher-appraisal models also consider those professional responsibilities that are less directly related to teaching itself. These include working and planning in teams; working on projects between schools; managing and sharing leadership responsibilities; providing professional advice to parents; building community partnerships for learning; and participating in professional development (OECD, 2005). Considering these responsibilities recognises the fact that the demands on schools and teachers have become more complex and teachers’ responsibilities have broadened as a result.



In a number of countries, the aspects appraised include a range of more specific elements, most of which (but not all) are covered by the four categories above. In Chile, teachers' competence in student assessment is appraised as part of regular performance management. In Estonia, appraisal for promotion considers the teachers work efficiency, in addition to fulfilment of qualification requirements. In Korea, student guidance is an important aspect in all types of teacher appraisal. In Mexico, the teachers' length of service is also considered for the reward scheme. In the Netherlands, teacher appraisal (at the end of probation and for performance management purposes) is intended to focus on a range of competencies closely related to the above aspects, including interpersonal competence, pedagogical competence, subject matter and didactical competence, organisational competence, team co-operation, co-operation with external actors, and reflection and development. In New Zealand, a variety of specific elements to be appraised are outlined in the teaching standards in addition to teaching and learning environments, such as professional relationships and values, bicultural partnership, promotion of inclusive learning environments, responsiveness to diverse linguistic and cultural backgrounds, and analysis and use of assessment information. In France, appraisal for performance management also includes an appraisal of teachers' "way of serving" (*manière de servir*), which includes punctuality, attendance, authority and "radiance" (*rayonnement*) as well as teachers' conformity to national programmes and reforms. In Poland, the intention of performance management is to assess "all aspects of teacher performance". In Australia, in addition to aspects related to the organisation of classroom teaching and professional development, appraisal for registration purposes also covers the teacher's professional engagement and general professional competence in relation to the Australian Professional Standards for Teachers.

In some countries, there is also a focus on teachers' knowledge in different areas. In Chile, for the teacher rewards scheme, there is an appraisal of the teacher's knowledge of discipline and curricular content as well as pedagogical knowledge. In Mexico, subject knowledge is considered as part of regular appraisal through the universal appraisal system. In Poland, subject area and didactic knowledge is considered as part of appraisal for promotion. In Portugal, both scientific and pedagogical dimensions are considered as part of regular appraisal. In Slovenia, there is an assessment of teachers' knowledge of legislation, the language of instruction and teaching skills. In the Netherlands, as mentioned above, subject-matter competence is one of the key areas to be appraised.

In some countries there are no central regulations regarding the aspects to be appraised. This is the case for regular appraisal and registration processes in Austria, and for regular appraisal in the Czech Republic and Hungary. In these cases, the specific scope of teacher appraisal is determined by the individual evaluators.

INSTRUMENTS AND INFORMATION SOURCES

A range of instruments and information sources is typically used to appraise teachers. As shown in Tables A.1, A.2 and A.3 in Annex A, the most frequently used instruments are classroom observation, interview/dialogue with the teacher, teacher self-appraisal, and portfolio.

Classroom observation

Teaching practices and evidence of student learning are likely to be the most relevant sources of information about professional performance. Most key aspects of teaching are displayed while teachers interact with their students in the classroom. As a result, teacher appraisal is usually firmly rooted in classroom observation. Almost all countries use classroom observations for regular performance management and many countries also use it at the end of the probation period. Classroom observations are also used for registration in New Zealand and as part of appraisal for promotion in Israel and Korea. In Korea, the performance-based incentive system is also based on observation of performance. Classroom observation is usually undertaken by the school leader or a member of the leadership team; but in Chile, a 45-minute class is videotaped and then evaluated by the national institution responsible for teacher evaluation. In Portugal, classroom observations are an optional element of regular appraisal, but they are required for the award of "very good" and "excellent" marks and for advancement to certain career grades.

Objective setting and individual interviews

Most teacher-appraisal models require the individual teacher to set performance objectives for a given period of time in agreement with the school management. The appraisal then assesses the extent to which the objectives were met. The setting of objectives, as well as the appraisal itself, usually involve individual interviews that foster reflective discussions between evaluators and teachers. In addition to classroom observations, interviews and/or dialogues with teachers are also frequently used across countries, usually for regular performance management, but also at the end of the probation period



and for registration processes. In Israel, a dialogue with the teacher is also part of the appraisal for promotion. In most countries, the interview is conducted by a member of the school leadership team; in Chile teachers are interviewed by a peer in addition to the principal.

Teacher self-appraisal

Requesting that the teacher being appraised evaluates his or her own performance is essential, as self-appraisal encourages teachers to reflect on the personal, organisational and institutional factors that have an impact on their teaching. Self-appraisal needs to be distinguished from informal teacher peer review, which has been shown to be an essential element of collaborative professional development. Teacher self-appraisal is used in most countries in regular appraisal for performance-management purposes. It is also a formal part of end-of probation processes in Israel, of registration processes in New Zealand, and of appraisal for promotion in Estonia and Israel.

Teacher's portfolio

An instrument that often complements teacher self-appraisal is a teacher's portfolio. A portfolio can include lesson plans and teaching materials, samples of students' work and sample commentaries on that work, self-reported questionnaires and reflection sheets. Portfolios are used for all types of appraisal. In Scotland, Singapore, and the German state of Thuringia, portfolios are used as one component of teacher appraisals; they are also used in England, but are not required by government regulations.

The content of a teacher's portfolio can be selected and assembled according to the purpose of the appraisal. Several researchers argue that portfolios provide information about the extent to which teachers are meeting educational standards (Klecker, 2000; Campbell et al., 2000; Tucker et al., 2002). Darling-Hammond (2001) argues that teacher development should take precedence in designing portfolios, and that "narrative reflection" is the best way to foster such development. Beck et al. (2005) observe that portfolios that focus on teacher development enhance professional outcomes. Combined with other evaluation instruments, documents prepared by the teacher may be used for summative appraisals.

Teachers sometimes consider the requirement to develop a portfolio as a burden that takes time away from their core work of teaching. Systems that rely on portfolios should thus encourage teachers to design their portfolios in such a way as to reflect a "natural harvest" of the teacher's work. For example, planning documents could describe a unit or lesson that the teacher is actually teaching; and the video, and accompanying commentary, could capture a lesson in class (Santiago et al., forthcoming).

Teacher testing

In some countries, teachers are subject to testing to assess their general and specialised competencies. In some rare instances, the results of these tests can be used for teacher appraisal. Chile and Mexico use teacher tests for their rewards schemes. Mexico also uses teacher tests for regular performance appraisal through the universal appraisal system, and there are several voluntary examinations that teachers can take for entry into the profession and to evaluate their professional competencies. Luxembourg and Slovenia report using a national examination of teachers at the end of the probation period, while Sweden uses a national examination of teachers for registration purposes.

Student results

Student learning outcomes are the essential criterion for the success of an education system. In some countries, the vast majority of teachers receive the highest or very high ratings of their performance in national teacher evaluation scheme and yet, at the aggregate level, student performance is unsatisfactory. Such mismatches between the messages conveyed to teachers and the performance of the education system in terms of learning outcomes underlines the importance of using student learning outcomes systematically as sources of evidence for teacher appraisal. To some extent, such learning outcomes can be measured through student test scores. Of the countries surveyed by the OECD, the Slovak Republic reported that student outcomes are used for teacher appraisal at the completion of probation, and Mexico reports that student outcomes (results of standardised assessments) are used for regular appraisal in the context of performance management. Mexico also uses student outcomes to evaluate teacher performance as part of its rewards scheme. In Chile, students' standardised assessment results are used to evaluate groups of teachers (teachers in individual schools) as part of the National Performance Evaluation System (SNED). England, Scotland and Singapore reported that they use student results at some point in the overall teacher-evaluation process.



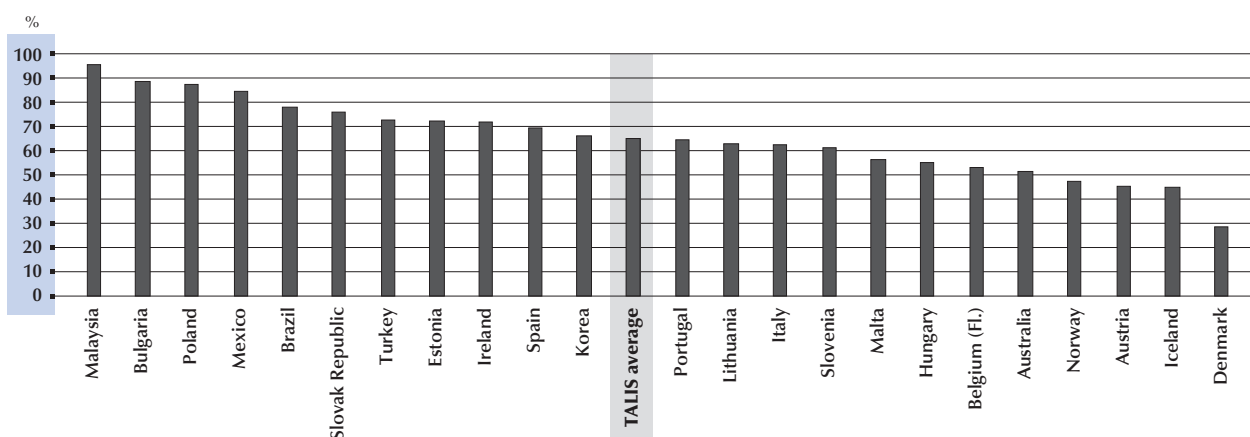
While student results are not formally considered as evidence for teacher evaluations in national appraisal frameworks, in the countries participating in the OECD Teaching and Learning International Survey (TALIS), 65% of teachers reported that they considered student test scores to be an important part of the feedback they received (Figure 2.1).

Teacher-appraisal systems based on student test results are intended to strengthen incentives for teachers to commit themselves to helping all students to meet important, centrally defined standards and fulfil goals within the national curriculum. Student learning outcomes, including student results in standardised assessments, are an appealing measure to assess teaching performance, since the ultimate goal of teaching is to improve student learning. Braun (2005) argues that considering student results is a promising approach for two reasons: first, it moves the discussion about teacher quality towards student learning as the primary goal of teaching; second, it introduces a quantitative and seemingly objective measurement of teacher performance.

Figure 2.1

Student test results as an aspect of teacher appraisal (2007-08)

Percentage of lower-secondary teachers who reported that student test results were considered to be of high or moderate importance in the appraisal and/or feedback they received



Source: OECD, *TALIS Database*.

However, it is difficult to identify the specific contribution a given teacher makes to a student's performance. Learning is influenced by many factors: the student's own skills, expectations, motivation and behaviour; the support students receive from their families and the influence of their peer group; school organisation, resources and climate; and curriculum structure and content. The effect that teachers have on a student's performance is also cumulative: at any given moment, a student is influenced not only by his or her current teachers but also by former teachers. Raw standardised scores thus reflect much more than the impact of a single teacher on student performance (Isoré, 2009).

In this respect, the development of value-added models that control for a student's previous results – and can thus potentially identify an individual teacher's contribution to the student's achievement – represents significant progress (see Box 2.3). However, there is wide consensus in the literature about two points. First, student outcomes should not be the *sole* measure of teacher performance, particularly when career decisions concerning the teacher, including pay, are concerned, because doing so introduces a substantial risk that teachers could be punished or rewarded for results beyond their control (Kane and Staiger, 2002; McCaffrey et al., 2003; CAESL, 2004; Braun, 2005; Ingvarson et al., 2007). Second, using student results as an evaluation instrument is likely to be more relevant for whole-school evaluations than for individual teacher-performance appraisals.

Given that evidence of student learning progress is fundamental, this should not imply that teachers are exempted from providing evidence to demonstrate student progress in their classrooms through, for instance, portfolios. It is also possible to design a system where teachers and school leaders meet and agree specific goals for student learning and for ways to assess student progress towards these goals. Such a system would encourage teachers to work with their colleagues and school leaders to identify measurable learning and performance goals for the entire class as well as for groups of students. For example, a



teacher with many struggling students may have both a class goal and a goal specifically for the struggling students. In this context, it is important that teachers not be penalised for setting high goals that are not always met, because that might result in teachers setting less-challenging goals for their students. Rather, the students' success and progress, even if they fall short of the goals, should be the basis for measuring teachers' contributions to student learning growth (Santiago et al., forthcoming). In New York State in the United States, for example, student learning objectives are used to measure teachers' contribution to students' progress in all subjects. Teachers receive guidance in setting appropriate learning objectives for their students, and districts exercise considerable discretion in approving appropriate assessments and measures to determine student progress. The measure of that progress constitutes 40% of teacher-evaluation scores while other state-approved measures, such as classroom observations, surveys and portfolios, constitute the remaining 60% (Santiago et al., forthcoming).

Box 2.1 Delaware, United States: Incorporating measures of student learning into teacher evaluations

In addition to measuring the usual components of teaching performance, such as planning, instruction, and classroom environment, education authorities in the United States are starting to incorporate measures of student learning. For example, the state of Delaware's evaluation system, the Delaware Performance Appraisal System (DPAS), was initially based on Danielson's *Framework for Teaching* and the four aspects of practice recommended for appraisal: planning and preparation, classroom environment, instruction and professional responsibilities. When the state revised the DPAS (now called DPAS II), it added a fifth aspect, known as Component V, on student improvement.

In addition, Delaware has identified methods for gathering information about student progress as an aspect of teaching performance. Component V calls for teachers to use three measures of student progress: one based on the state standards test, an alternative measure based on an assessment other than the state standards test, and goals for student progress developed by the educator. In the 2011-12 school year, Delaware engaged hundreds of the state's teachers to develop a wide-ranging library of resources that support implementation of DPAS Component V.

Delaware was also an early leader in efforts to refine observation instruments. For each of the aspects appraised, the rubric describes specific teaching competencies, which it calls criteria. The purpose of the rubric is to allow "the teacher and evaluator to develop a common understanding of the teacher's strengths and areas for improvement" and to help "ensure evaluator consistency".

Source: United States Department of Education response to OECD survey.

Surveys of students and parents

Surveys can provide information about how students, parents and others who have continuous contact with a teacher perceive that teacher's performance (Peterson, 2000; Peterson et al., 2000; 2003; Jacob and Lefgren, 2005). Such surveys are rarely used systematically in formal teacher appraisals, however. Across the countries surveyed by the OECD, Mexico reported using student surveys for performance management, while New Zealand and the Slovak Republic reported sometimes using student and parent surveys for regular appraisals for performance management. Parent surveys are also used for regular appraisal in Canada. Poland reported that the opinion of the parents' council is considered for the rewards/promotion scheme. In Korea, a multi-dimensional method, including questionnaires completed by students and parents, is used for regular teacher appraisal for professional development. Of the ten additional countries surveyed for this report, only the United States reported using surveys of students or surveys of parents at some point in the appraisal process.

Other indicators of teacher performance

A number of countries report using appraisal instruments and information sources that do not readily fit the above categories. In France, for example, the instruments used at the end of probation are a report prepared by the teacher's tutor, the opinion of the school leader and, in rare cases, the inspection report. In Italy, the information used at the end of probation concerns formal aspects of a teacher's work, such as attendance at school and participation in training. In the Netherlands, an extensive description of a teacher's competencies is available for both completion of probation and regular performance management. In Austria, the evaluators choose which instruments to use for performance management and registration appraisals. In Portugal, the instruments used for regular appraisal are outlined by each school in the school-development plan. In some countries, including Sweden, Hungary and Estonia, school leaders choose the instruments of certain appraisal processes.



Box 2.2 Italy: The “Valorizza” (valorisation) experiment to identify and reward teachers

Italy currently has no system to appraise teachers' performance. School heads are not entitled to do so, nor are the local authorities, nor the inspectors, who only act as a last resort in cases of serious misconduct. Equally important, Italy has no data systems in place that would allow for systematic measurement of individual teacher or student performance on which appraisal could be based.

It was in this context that the Ministry of Education launched an experiment in 2010 to base teacher appraisal on the collective experience and views of key stakeholders – principals, teachers parents and students - in the school. Some 33 volunteering schools were invited to join this experiment and, within each school, teachers were invited to volunteer for being evaluated. Each teacher completed a self-evaluation questionnaire, focused on their professional behaviour, attitudes and practices, and provided information on their professional background and career. All parents of the school and the students of the last two grades (12 and 13) were also given an evaluation form to fill in, where they were asked to name up to three teachers whom they considered as the most highly and widely respected for their professional behaviour. The teachers' assembly elected two of its members who, along with the school head, formed the evaluating committee.

Each of the three committee evaluators carefully reviewed the questionnaires and professional background of the candidates and, later, the evaluation forms completed by parents and students. No single element of this evaluation process was given priority over others. The analysis was carried out individually, without prior agreement and without exchanging views with other participants during the evaluation process. At the end, each of the evaluators identified his or her personal list of those teachers they considered most highly and widely respected for their professional behaviour (up to 30% of the candidates). Afterwards, the three evaluators met for the first time and compared their lists; the teachers who had been listed by all three evaluators were selected. Among those candidates who were included in two lists were compared and those who were “comparatively better” were added to the overall list, with the aim that the total would not exceed 30% of the candidates in the school. More than 900 teachers were evaluated; 276 of them were selected as being highly and widely respected for their professional behavior within their school.

It is interesting to note that two-thirds of the selected candidates were independently identified by the three evaluators, and that nearly all in the remaining third were identified by two of the evaluators. These outcomes suggest that key stakeholders in a school share very similar views as to who is highly and widely respected for their professional behaviour, even if those stakeholders may not be able to define or agree upon the exact criteria that characterise that professional behaviour.

Subsequently, the Ministry asked two respected Foundations (Associazione Treille and Fondazione per la Scuola della Compagnia di San Paolo) to investigate the extent to which the outcomes from the evaluating committee were considered appropriate by the other stakeholders in each school. A questionnaire was given to all teachers (candidates and not) and to the parents and students of the schools in the sample. Once again, an average of two-thirds of the respondents fully agreed with the results.

It is noteworthy that all schools accepted the experiment favourably. In particular, the exercise was perceived to provide recognition of teachers for their professional behaviour, and was seen as evidence of trust in the capacity of school communities to appraise their own members, without having to rely on external experts. In addition, the high level of agreement on the selected candidates helped to avoid conflict and disputes about the evaluation process.

The fact that each stakeholder offered his or her view separately from others, yet there was a high degree of convergence among the perspectives, also suggests a high level of face validity of the process and that a good reputation is a shared opinion, even though it may be based on different criteria or motivations.

As a next step, the Valorizza experiment suggests giving the selected teachers a yearly allowance equal to two to three months of their ordinary salary for three consecutive years and to repeat the Valorizza appraisal process every three years, always on a voluntary basis.

Source: Fondazione Treille.

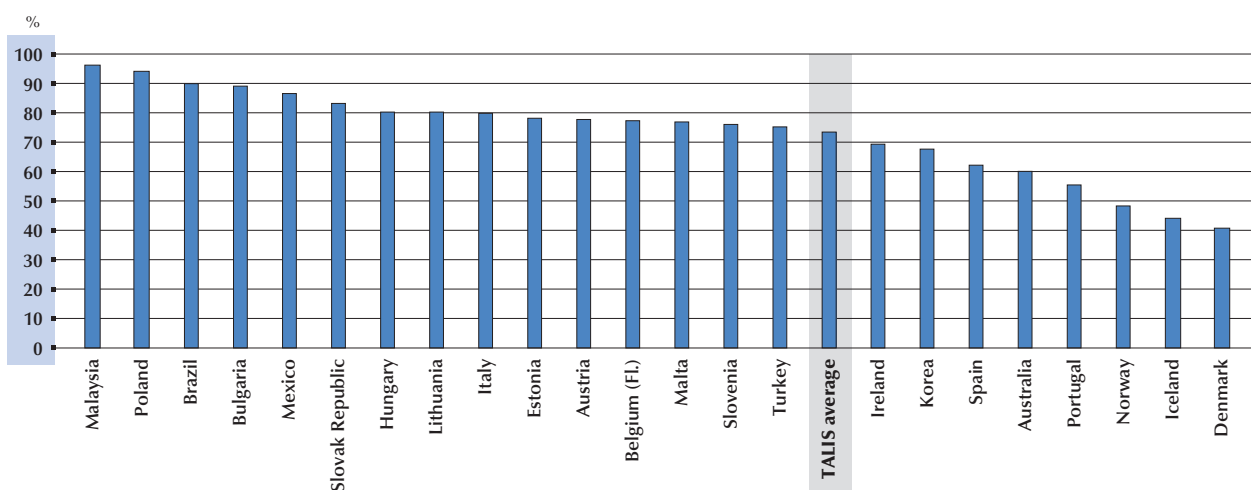


Improving teaching and learning through teacher appraisal

Most countries that have formal policy frameworks for teacher appraisal now include classroom observation as a key instrument for appraisal. However, evidence from TALIS indicates that there are differences across countries in the degree to which teachers regard classroom observations as an important component in their appraisal. On average across the countries that participated in the TALIS survey, 73.5% of teachers reported that they considered the direct appraisal of their teaching in the classroom as moderately or highly important; however, only 40.7% of teachers in Denmark, 44.1% of teachers in Iceland, 58.4% of teachers in Norway and 55.3% of teachers in Portugal shared that view (Figure 2.2).

Figure 2.2

Direct evaluation of classroom teaching as an aspect of teacher appraisal (2007-08)
 Percentage of lower-secondary teachers who reported that direct appraisal of classroom teaching was considered to be of high or moderate importance in the appraisal and/or feedback they received



Source: OECD, *TALIS Database*.

In countries where teacher appraisal is more informal and not regulated by central frameworks, school leaders may not be expected to enter classrooms and observe teaching practice with an evaluative focus. Rooting teacher appraisal firmly in classroom observations can be challenging in countries where there are strong traditions of teacher autonomy and little experience in classroom observations with an evaluative focus by school leaders or other teachers. In those countries, the main source of professional feedback for teachers is often a dialogue with the school leader on issues such as working conditions, responsibilities and salaries without, however, systematically including observation of and feedback on actual teaching practice (Nusche et al., 2011a; 2011b; Shewbridge et al., 2011). The advantage of classroom observations, however, is that they are more likely to identify weaknesses in practice that can be addressed through appropriate professional-development activities. Other proxies for teaching quality, such as lessons plans or evidence of communication with parents, are important too, but they cannot substitute for what actually occurs in the classroom (Santiago et al., forthcoming).

In most countries, classroom observations are conducted by personnel within the school, generally school leaders. Several researchers have criticised the practice of conducting annual announced evaluations as they do not provide an authentic picture of day-to-day teaching and often do not involve constructive feedback or coaching for improvement (Klinger et al., 2008; Daley and Kim, 2010; Danielson, 2011; Marshall, 2012; Papay, 2012). In several countries surveyed by the OECD, teachers were observed only once every few years.

For classroom observations to be useful, each school must have the capacity to conduct them effectively. This requires that leadership teams be trained in conducting observations and in engaging in constructive discussions with teachers. Training should include teachers as well, since it is critical for them to understand how their performance will be assessed.



While high-quality classroom observations appear to be related to better student outcomes, the quality of classroom observations depends on how well trained the observers are (Kane and Staiger, 2012; Kane et al., 2010; Milanowski, 2004 in Santiago et al., forthcoming).

Using multiple sources of evidence for teacher appraisal

Using several appraisal instruments and multiple sources of evidence for teacher-appraisal systems allows evaluators to measure different knowledge and skills to obtain a comprehensive picture of teachers' abilities (Goe et al., 2008; Peterson, 1987; Rockoff and Speroni, 2011). This is particularly important when appraisal results are used for high-stakes decisions, such as promotion and tenure (Sykes and Winchell, 2010).

While teacher-appraisal models that use multiple instruments and evaluators are more likely to provide a solid basis on which to appraise teachers, limited resources make trade-offs inevitable. As explained in Isoré (2009), comprehensive teacher-appraisal procedures imply greater direct and indirect costs at every stage of the process: agreeing on the design of the system requires time for consultations with all stakeholders; training evaluators is expensive and time-consuming; conducting appraisals implies additional work for both teachers and evaluators; and aligning broader school reforms, such as professional-development opportunities, with appraisals requires more resources.

Box 2.3 United States: Measures of Effective Teaching Project

The Bill and Melinda Gates Foundation in the United States funded a three-year study, the Measures of Effective Teaching (MET) Project, to help "build and test measures of effective teaching to find out how evaluation methods could be best used to tell teachers more about the skills that make them most effective". The Foundation released its third and final report from the MET study in early 2013. The study developed and tested a variety of measures of teaching in order to determine how those measures might be used to predict student performance. Over 3 000 teachers in six US states volunteered to have their lessons videotaped. The teachers' effectiveness was assessed using three different types of measures: classroom observation, student-opinion surveys, and progress in student achievement. Teachers' lessons were videotaped and rated by several people who were trained in the use of the observation instrument. The teachers' students were surveyed using the *Tripod Student Perception Survey*. Teachers who participated in the second year of the study were randomly assigned students to control for prior academic performance. At its conclusion, the study had student-improvement data on 1 600 teachers over two years.

The final report draws three conclusions: that effective teaching can be measured; that using multiple measures with balanced weights helps to understand the different aspects of effective teaching; and that a second observer of classroom practice increases the reliability of classroom-observation measures. In short, the report demonstrates that teachers whose students show improvement were likely to get good results in the future, and that teachers who scored well on other measures, such as classroom observations and student surveys, were likely to have students who showed improvements.

The report also considers the implications of using multiple measures to evaluate teacher performance. It constructs different balances among multiple measures and concludes that equal weights create a more accurate assessment of teacher effectiveness than other models in which one measure is given a greater weight over others. Relying too heavily on a single measure makes it more likely that other valued practices are obscured by a measure that does not consider them.

Finally, the report recognises the value of classroom observation in measuring teacher effectiveness. Feedback from observations can be a powerful tool because it offers teachers actionable advice on how to improve classroom practice. But the report makes the point that observations become much more reliable tools when there is more than one observer involved in the process.

Source: http://metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf.



Box 2.4 China: Using multiple measures to evaluate performance

In China, teachers are appraised against four aspects of their performance: professional integrity or values; capability, including skills and competencies (virtues); diligence; and their own and their students’ achievement. The emphasis in these evaluations tends to fall on teachers’ professional integrity and students’ learning results.

Teachers are evaluated by an elected task force that includes union representatives, party leaders, teachers and other elected officials. This task force looks at teacher self-appraisal and peer and parent reviews to assess a teacher’s integrity and capability. For diligence, a teacher’s attendance record is reviewed. To assess a teacher’s achievement, student outcomes are reviewed, and any awards or certificates the teacher has received, any papers he or she has published, and the teacher’s participation in research projects are considered.

The appraisal results inform a variety of decisions about the teacher’s future career, including salary and professional titles, and are added to a teacher’s portfolio as proof of teaching experience. The Director for Instruction can also use this information to reduce a teacher’s workload if he or she is identified as low-performing or even to transfer the teacher to another post.

Source: Chinese government response to OECD survey.

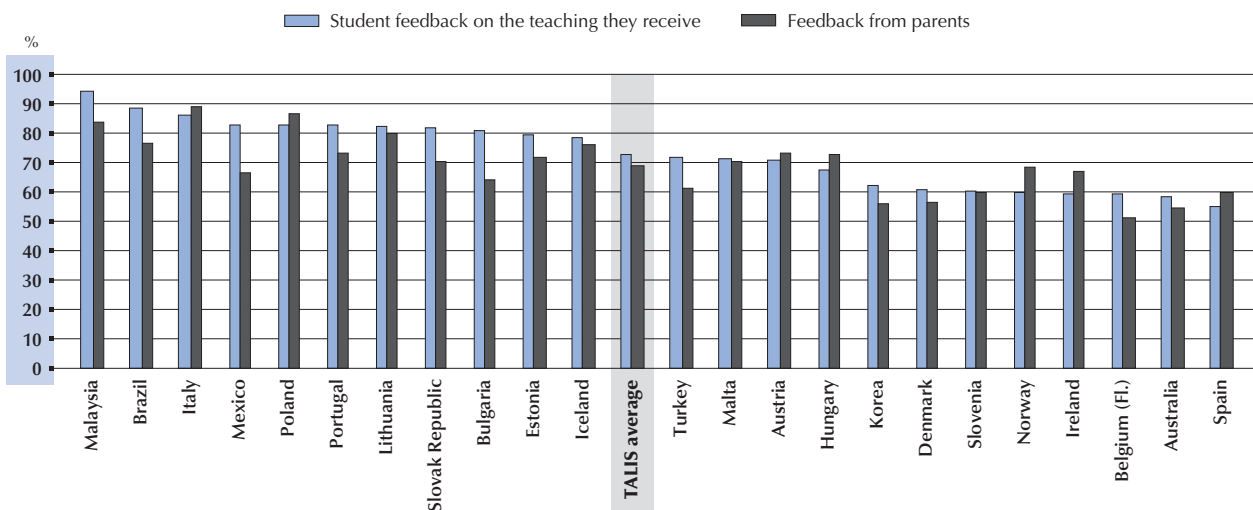
Using stakeholder surveys formatively

Although stakeholder surveys only rarely form part of countries’ formal teacher-appraisal frameworks, evidence collected in TALIS indicates that many teachers consider feedback from parents and students an important aspect of their appraisal (Figure 2.3). In some countries, teachers design their own student surveys to obtain feedback on their teaching practices and their students’ progress.

Figure 2.3

Student and parent feedback as an aspect of teacher appraisal (2007-08)

Percentage of lower-secondary teachers who reported that student and parent feedback was considered to be of high or moderate importance in the appraisal and/or feedback they received



Source: OECD, TALIS Database.



Student and parent surveys can provide important formative feedback to teachers (see Box 2.5). In most contexts, student surveys are not reported to higher levels of the school administration and are generally used only for improvement purposes by the concerned teacher. Parent surveys are more relevant for whole-school evaluations than for individual teacher-performance appraisals. As explained by Isoré (2009), the sparse evidence on this subject shows that parents value qualities in teachers that may have little to do with student achievement, including “the teacher’s ability to promote student satisfaction” (Jacob and Lefgren, 2005), “humane treatment of students” and “effective communication and collaboration with parents” (Peterson et al., 2003).

Box 2.5 Using student feedback to help teachers improve their teaching

The Union of Education **Norway** (the largest union of teachers and school leaders in the country) and the Norwegian Student Organisation have been working together in a national initiative to develop principles and guidelines for teacher appraisal by students. The groups recommend that student surveys should: focus on teaching practice rather than on the teacher as an individual; include students’ self-assessment and assessment of peers so as to allow for an analysis of how student effort and motivation influence the learning environment; feature questions on teaching approaches that are relevant for student learning, such as adapted education and feedback to students, as well as questions on the general context of teaching, such as materials and physical conditions; be carried out anonymously so as to ensure students give honest answers; and be analysed by the teacher and students together with the aim of improving the classroom environment and learning outcomes. This exercise should be followed up with a joint report by the teacher and a group of students on their analysis of results and agreed changes to be made. This report, together with relevant data, should be submitted to the teachers’ closest supervisor. While not all stakeholders agree with the recommendations that have emerged from this project, most have accepted the general idea that student views are an important source of feedback that teachers can use to improve their practice.

Reflecting the student-centred approach to education in **Sweden**, teachers often conduct surveys among their students with the aim of obtaining feedback on their teaching practices. These surveys are organised at the teachers’ own initiative and results are used exclusively by the teacher concerned, often in interaction with the students. Peterson et al. (2000) argue that students respond reliably about teacher quality if questions are formulated in a simple and relevant way. Teachers interviewed by the review team said that students provide useful insights into their strategies for teaching and learning, and that they find this opportunity for feedback important. Student surveys are kept within the classroom and used only to help the teacher improve his or her practice.

Sources: Nusche et al. (2011a), Nusche et al. (2011b).



Note

1. In Chile, the appraisal is conducted every fourth year if the results are satisfactory; otherwise it is conducted every year or every other year.

References

Beck, R., N. Livne and S. Bear (2005), "Teachers' Self-Assessment of the Effects of Formative and Summative Electronic Portfolios on Professional Development", *European Journal of Teacher Education*, Vol. 28, No. 3, pp. 221-244.

Braun, H. (2005), *Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models*, Educational Testing Service, Princeton, New Jersey.

CAESL (Center for Assessment and Evaluation of Student Learning) (2004), "Using Student Tests to Measure Teacher Quality", *CAESL Assessment Brief*, No. 9.

Campbell, D., B. Melenzyer, D. Nettles and R. Wyman (2000), *Portfolio and Performance Assessment in Teacher Education*, Allyn and Bacon, Needham Heights, Massachusetts.

Daley, G. and L. Kim (2010), "A Teacher Evaluation System that Works", Working Paper, National Institute for Excellence in Education, August 2010.

Danielson, C. (2011), "Evaluations that Help Teachers Learn", *Educational Leadership*, December 2010/January 2011, pp. 35-39.

Danielson, C. (1996, 2007), *Enhancing Professional Practice: A Framework for Teaching*, 1st and 2nd Editions, Association for Supervision and Curriculum Development, Alexandria, Virginia.

Darling-Hammond, L. (2001), "Portfolio as Practice: The Narratives of Emerging Teachers", *Teaching and Teacher Education*, Vol. 17, pp. 107-121.

Goe, L., C. Bell and O. Little (2008), *Approaches to Evaluating Teacher Effectiveness: A Research Synthesis*, National Comprehensive Center for Teacher Quality, Washington, DC.

Ingvanson, L., E. Kleinhenz and J. Wilkinson (2007), *Research on Performance Pay for Teachers*, Australian Council for Educational Research, Camberwell.

Isoré, M. (2009), "Teacher Evaluation: Current Practices in OECD Countries and a Literature Review", *OECD Education Working Papers*, No. 23, OECD Publishing, Paris. www.oecd.org/edu/workingpapers; <http://dx.doi.org/10.1787/223283631428>.

Jacob, B. and L. Lefgren (2005), "What do Parents Value in Education: An Empirical Investigation of Parents' Revealed Preferences for Teachers", *NBER Working Paper*, No. 11494.

Kane, T.J. and D.O. Staiger (2012), "Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains", MET Project Policy and Practice Brief, Bill and Melinda Gates Foundation, Seattle, Washington.

Kane, T. and D. Staiger (2002), "Volatility in School Test Scores: Implications for Test-Based Accountability Systems", *Brookings Papers on Education Policy*, Washington, DC.

Kane, T.J., E.S. Taylor, J.H. Tyler and A.L. Wooten (2010), *Identifying Effective Classroom Practices Using Student Achievement Data*, National Bureau of Economic Research, Cambridge, Massachusetts.

Klecker, B. (2000), "Content Validity of Preservice Teacher Portfolios in a Standards-Based Program", *Journal of Instructional Psychology*, Vol. 27, No. 1, pp. 35-38.

Klinger, D.A., L.M. Shulha and C. DeLuca (2008), "Teacher Evaluation, Accountability, and Professional Learning: The Canadian Perspective", *Rev. Pensamiento Educativo*, Vol. 43, pp. 209-222.

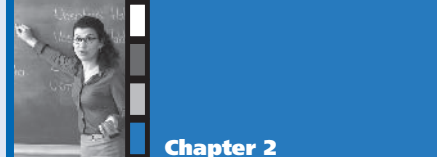
Marshall, K. (2012), "Let's Cancel the Dog-and-Pony Show", *Phi Delta Kappan*, Vol. 94, No. 3, pp. 19-23.

McCaffrey D. F., J.R. Lockwood, D. Koretz and L. Hamilton (2003), *Evaluating Value-Added Models for Teacher Accountability*, Rand Corporation, Santa Monica, California.

Milanowski, A. (2004) "The Relationship Between Teacher Performance Evaluation Scores and Student Achievement: Evidence from Cincinnati", *Peabody Journal of Education*, Vol. 79, No. 4, pp. 33-53.

Nusche, D., L. Earl, W. Maxwell and C. Shewbridge (2011a), *OECD Reviews of Evaluation and Assessment in Education: Norway*, OECD Publishing, Paris. www.oecd.org/edu/evaluationpolicy; <http://dx.doi.org/10.1787/9789264117006-en>.

Nusche, D., G. Halász, J. Looney, P. Santiago and C. Shewbridge (2011b), *OECD Reviews of Evaluation and Assessment in Education: Sweden*, OECD Publishing, Paris. www.oecd.org/edu/evaluationpolicy; <http://dx.doi.org/10.1787/9789264116610-en>.



OECD (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264018044-en>.

Papay, J.P. (2012), "Refocusing the Debate: Assessing the Purposes and Tools of Teacher Evaluation", *Harvard Educational Review*, Vol. 82, No. 1.

Peterson, K. (2000), *Teacher Evaluation: A Comprehensive Guide to New Directions and Practices*, 2nd Edition, Corwin Press, Thousand Oaks, California.

Peterson, K., C. Wahlquist and K. Bone (2000), "Student Surveys for Teacher Evaluation", *Journal of Personnel Evaluation in Education*, Vol. 14, No. 2, pp. 135-153.

Peterson, K., C. Wahlquist, J. Esparza Brown and S. Mukhopadhyay (2003), "Parents Surveys for Teacher Evaluation", *Journal of Personnel Evaluation in Education*, Vol. 17, No. 4, pp. 317-330.

Peterson, K.D. (1987), "Teacher Evaluation with Multiple and Variable Lines of Evidence," *American Educational Research Journal*, Vol. 24, No. 2, pp. 311-317.

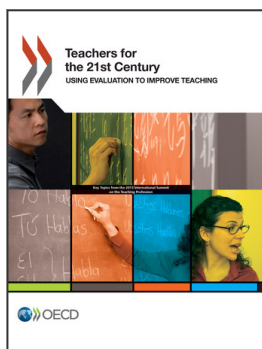
Rockoff, J. and C. Speroni (2011), "Subjective and Objective Evaluations of Teacher Effectiveness: Evidence from New York City", *Labour Economics*, Vol. 18, pp. 687-696.

Santiago, P., F. Benavides, C. Danielson, L. Goe and D. Nusche (forthcoming), *OECD Reviews of Evaluation and Assessment in Education: Teacher Evaluation in Chile*, OECD Publishing, Paris. www.oecd.org/edu/evaluationpolicy; <http://dx.doi.org/10.1787/9789264172616-en>.

Shewbridge, C., E. Jang, P. Matthews and P. Santiago (2011), *OECD Reviews of Evaluation and Assessment in Education: Denmark*, OECD Publishing, Paris. www.oecd.org/edu/evaluationpolicy; <http://dx.doi.org/10.1787/9789264116597-en>.

Sykes, G. and S. Winchell (2010), "Assessing for Teacher Tenure", in M. Kennedy (ed.), *Teacher Assessment and the Quest for Teacher Quality*, Jossey-Bass, San Francisco, pp. 201-224.

Tucker, P., J. Stronge and C. Gareis (2002), *Handbook on Teacher Portfolios for Evaluation and Professional Development*, Eye on Education, Larchmont, New York.



From:
Teachers for the 21st Century
Using Evaluation to Improve Teaching

Access the complete publication at:
<https://doi.org/10.1787/9789264193864-en>

Please cite this chapter as:

OECD (2013), "Procedures for teacher appraisal", in *Teachers for the 21st Century: Using Evaluation to Improve Teaching*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264193864-4-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 and any 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.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.