



2

What Behaviours Do Teachers Reward?

This chapter examines the uses and significance of school marks. It discusses the kinds of behaviours, habits and attitudes that teachers reward with better marks, some of which may be unrelated to student learning. It also explores differences in how countries and economies distribute marks and suggests ways of improving the effectiveness of school marks.



Within the classroom, one of the most important levers teachers have to guide the behaviour of students is marks/grades. Marks serve three main instructional purposes: they provide formative feedback to students, they provide teachers with feedback for instructional planning, and they certify that students have mastered skills considered relevant by the teacher or school (Carey and Carifio, 2012). Such is the recognised importance of in-class assessments that, in PISA, practically all students attend schools that assess students' work or progress through teacher-developed tests and/or assignments that the students return. In almost all countries and economies over 95% of students are assessed this way (see Table IV.3.10 in OECD, 2010). Through marks, teachers signal what kinds of attitudes and behaviours they reward, and signal what they believe is important for student success. In general, teachers reward achievement – students' mastery of skills and knowledge, as measured by standardised tests. Yet the relationship between marks and test scores is not perfect, and this difference is related to the behaviours and attitudes that teachers reward (Brookhart, 1993; McMillan et al., 2002).

While teachers use marks as a diagnostic tool, they also use them to communicate expectations and foster motivation in their students (Stiggins and Conklin, 1992), because students react to the incentives marks represent by modifying their behaviour (Bonesrønning, 1999). Marks as a mode of communication and a source of incentives influence student interest in school and the subject matter, self-efficacy, motivation, and future performance (Brookhart, 2009; Docan, 2006; Guskey, 2004). Used effectively and correctly, marks can motivate students to put forth more effort and change their behaviours and attitudes in a direction that is beneficial to students and learning. Marks can, however, discourage and disengage students when they are inconsistent with student models of motivation (Covington, 1984; Kohn, 1993).

In the context of the importance and pervasiveness of marks, this chapter answers the following questions: what kinds of behaviours, habits and attitudes do teachers reward with better marks? Do certain teachers, schools and countries/economies offer rewards for attitudes that are unrelated to student learning? Do countries and economies differ in the way they distribute marks and reward student behaviour? And, can more efficient grading patterns be identified among school systems?

Generally, teachers can be categorised according to the way in which they use marks to promote students' learning. While some teachers subscribe to the "keeping things moving" approach, others follow a "developmental" approach to marking (Kelly, 2008). Teachers who adopt the "keeping things moving" approach reward in-class co-operation and penalise disruptiveness and disciplinary problems. These teachers reward behaviours that are not necessarily directly related to future achievement, but that help to keep instruction moving smoothly (Stiggins and Conklin, 1992). Teachers who adopt the "developmental" approach, on the other hand, use marks to foster student engagement. These teachers focus on cultivating interest, concentration and effort, because they believe that student engagement is a necessary precondition for learning (Fredricks et al., 2004). These teachers seek to motivate what they believe are key student competencies that will help students become better learners inside and outside the classroom. They emphasise fairness and are sensitive to how students will react to the marks they receive so as to promote self-esteem and future engagement (Brookhart, 1993). Another incentive that motivates the way teachers determine marks is the pressure to align their teaching and their marks to the contents and results of external standardised assessments, where these are implemented and have high stakes for students, teachers and/or schools.

Marks vary according to the incentives teachers want to put in place for their students. The meaning and purpose of marks thus vary with respect to the objectives and nature of assessments. Assessments, however, should not be a synonym of marks. Assessments are the gathering of information for the purpose of making instructional or career-path decisions, and the information from assessments does not always need to be transmitted to students; not all assessments need to produce marks. Whenever assessments are graded, marks should provide clear and useful information for the purpose of enhancing learning. The criteria used in marking and the behaviours or content that is being evaluated in them should be clear to students so that the information that is contained in marks is interpretable and useful for students and teachers to improve learning (Guskey and Bailey, 2001; Marzano, 2000; O'Connor, 2002; O'Connor, 2009; Tomlinson, 2005).

Empirical research examining teachers' marking tendencies and how students react to the marks they receive generally focuses on a specific country. PISA is unique in that it allows for a comparison of how marks are distributed across countries and economies. PISA can also provide information on the proportion of schools that motivate behaviours that are conducive to learning. A careful study of marks is of utmost importance because student marks – insofar as they convey information not only about academic abilities but also about students' non-cognitive skills (e.g. engagement, approaches to learning, attitudes towards challenges, etc.) and students' rank within the class – are one of students' most important sources of information about their potential to succeed in further educational endeavours and in the labour market later on. In fact, while marks reflect, in part, students' skills and mastery of content, their power to predict students' future educational and social outcomes is stronger than the effect of academic achievement alone. For example, students' marks, net of achievement, predict educational plans, course-taking patterns, eventual achievement and earnings nine years after graduation (Kelly, 2008; Rosenbaum, 2001).

Students who take the PISA test – and similar tests – are undergoing an assessment of the skills and competencies that they have accumulated over the years. Most students are not familiar with such objective performance measures, however, and in most



countries such standardised tests only occur at the end of specific academic cycles. This means that while educators and policy makers are prompted to improve the efficiency and equity of their own school or of entire education systems by the school's/system's performance in standardised tests, student motivation is stimulated by a different set of incentives altogether. Marks represent the most concrete form of information students have about their abilities and potential. A careful assessment of grading practices in different education systems reveals the type of incentives that are put in place to facilitate the learning process where most learning takes place.

This chapter uses data from the PISA Educational Career questionnaire (ECQ). The questionnaire was optional for countries participating in the PISA 2009 assessment. Students in 17 countries and economies were asked, among other things, to report what mark (grade) they received in their language-of-assessment class. This chapter identifies the attitudes and behaviours that are most strongly related to marks, with a special focus on those habits, attitudes and behaviours that are not necessarily conducive to improving cognitive skills, such as those associated with students' educational context and background.¹

Because countries differ in the way in which they instruct their teachers to mark students and in the scales used to distribute marks, this chapter also offers a thorough look at the different schemes countries have developed to determine and distribute marks to students. Countries vary remarkably in the scales teachers use to mark students; but in all countries considered, teachers reward similar attitudes and behaviours. For example, students who perform well in standardised tests like PISA are more likely to receive better marks, indicating that teachers reward, first and foremost, skills and mastery of the types of knowledge of in-class content that standardised tests measure. In many school systems, teachers reward behaviours that promote learning, such as engagement with reading and control strategies. Yet in many school systems, attitudes, values and behaviours unrelated to learning are also rewarded with higher marks. In these systems teachers often mark on a curve and/or use a marking system that favours students with particular background characteristics.

HOW COUNTRIES GRADE THEIR STUDENTS: MARKS IN PISA

Very few countries share the same grading schemes. Notwithstanding the fact that PISA did not compile information on the content that is marked and the standards that are applied in grading students, Figure 2.1 highlights the different grading schemes adopted by teachers in the countries and economies participating in this study. The education systems in Austria, Hungary and Serbia use a scale from 1 to 5 while those in Croatia, Poland and the Slovak Republic use a scale from 1 to 6. The education systems in Iceland, Italy and Latvia mark from 1 to 10, and the Flemish Community of Belgium, Ireland, Singapore and Trinidad and Tobago use a scale from 1 to 100. In Portugal, students in ISCED level 2 programmes receive marks on a scale from 1 to 5 while students in ISCED level 3 programmes receive marks on a scale from 1 to 20. In most countries and economies, higher values on the scale are indicative of better evaluations. However, in Austria and the Slovak Republic, a value of 1 indicates the better evaluation. In all these countries, the scales are discrete, that is, students do not report receiving marks with decimal places.

■ Figure 2.1 ■

Marking schemes across countries and economies

Failing mark	Marking range					Schools or regions have different marking systems
	1 to 5	1 to 6	1 to 10	1 to 20	1 to 100	
Many possible values	Portugal ISCED 2		Iceland, Italy, Latvia	Portugal ISCED 3	Belgium (Flemish Community), Ireland, Singapore, Trinidad and Tobago	Macao-China, Mexico, New Zealand
One value	Austria, Hungary, Serbia	Croatia, Poland, the Slovak Republic				

StatLink  <http://dx.doi.org/10.1787/888932733526>

Some countries and economies have important internal variation in the schemes teachers and schools use to mark their students. Certain schools within countries or certain teachers within schools may use one scheme while other schools and teachers use other schemes. This is the case for Macao-China, Mexico and New Zealand. Many 15-year-old students in grade 11 in New Zealand, for example, receive marks of "Not Achieved," "Achieved," "Merit," or "Excellence" that are consistent with work towards the national certificate of achievement. However, students attending lower grades, or following specific programmes (e.g. International Baccalaureate) or in schools that choose other marking systems may receive their marks as a percentage or on another scale. In Macao-China's schools, teachers are completely autonomous in the way they choose to assess and mark their students. In Mexico, education levels have different marking schemes. As a result, Macao-China, Mexico and New Zealand asked PISA 2009 students not about the specific mark they received in the past year, but to report whether the student received a mark that placed him/her a) below the passing mark or b) at or above the passing mark.

Different education systems establish different criteria for communicating to students that they have failed the class or the assessment. In some countries, the marking scheme allows for only one possible failing value, meaning that students who fail do not know how

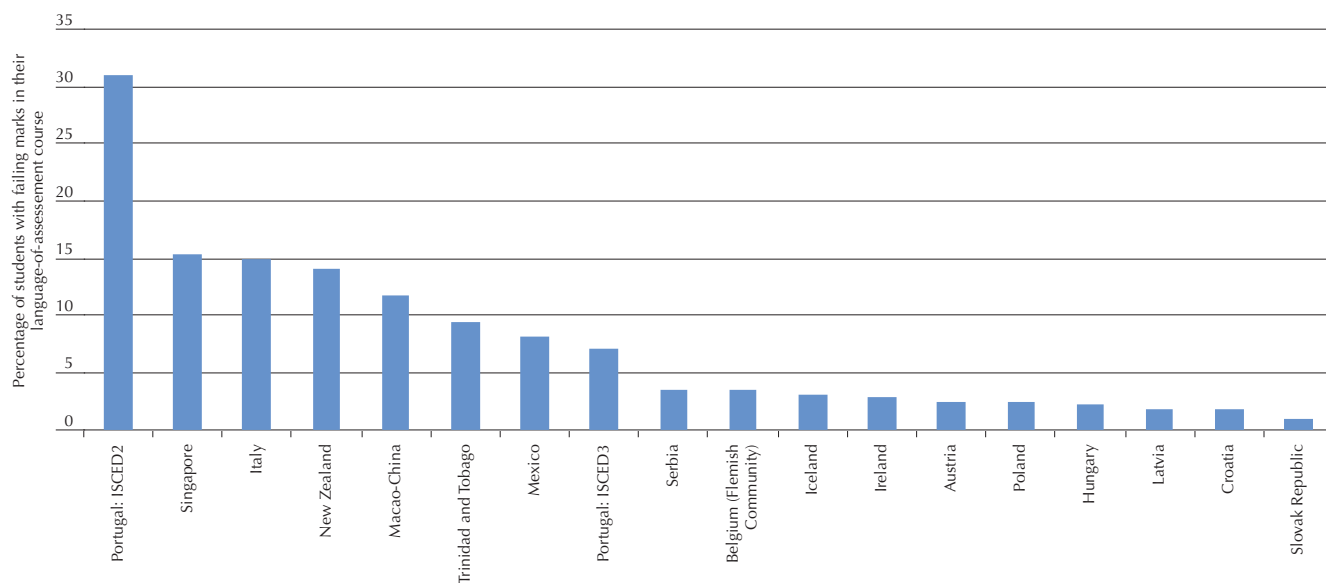
far they are from meeting the passing criteria. This is the case in Austria, Croatia, Hungary, Poland, Serbia and the Slovak Republic, for example. The remaining values on the marking scale reflect the quality of the passing assignment as, for example, “sufficient”, “good”, “very good” and “excellent”.

Other countries establish the passing mark somewhere in the middle of the scale, allowing for marks to communicate how far students are from the minimum passing criteria. The Flemish Community of Belgium, Italy and Singapore, for example, establish the cut-off point in the distribution at the 50% range of the marking scale. In Ireland, the grading scale ranges from 0 to 100, but only scores below 40 are considered unsatisfactory or failing.² Research that evaluates the efficacy of marking schemes to motivate students to put forth more effort suggest that using extremely low values fails to motivate students to improve, especially when a student receives a minimum mark at the beginning of the school year. Moreover, students receiving an unusually low mark may have a hard time recovering to pass the course, because a low mark skews the mark average and undermines motivation. This is particularly the case for scales that allow failing marks drop far below the passing mark (Carey and Carifio, 2012; Guskey, 2004; Reeves, 2004).

Considering the different marks school systems use to identify failing students, more than 30% of students in ISCED level 2 programmes in Portugal receive failing marks, a percentage that is consistent with the large number of students who report repeating a grade (OECD, 2010). Percentages of students receiving failing marks are also high in Italy, Macao-China, New Zealand and Singapore, where at least 10% of students reported having received a failing mark in their language-of-assessment course. Receiving failing marks is comparatively uncommon in Austria, Belgium (Flemish Community), Croatia, Hungary, Iceland, Ireland, Latvia, Poland, Serbia and the Slovak Republic, where fewer than 5% of students receive failing marks (Figure 2.2). As described above, most of the school systems that have a small percentage of students receiving failing marks in their language-of-assessment courses have only one value to identify failing, but a more detailed scale to identify passing results.


■ Figure 2.2 ■

Percentage of students with failing marks in their language-of-assessment course



Countries are sorted in descending order of the percentage of students with failing marks in their language-of-assessment courses.

Source: Table B2.2.

StatLink  <http://dx.doi.org/10.1787/888932733545>

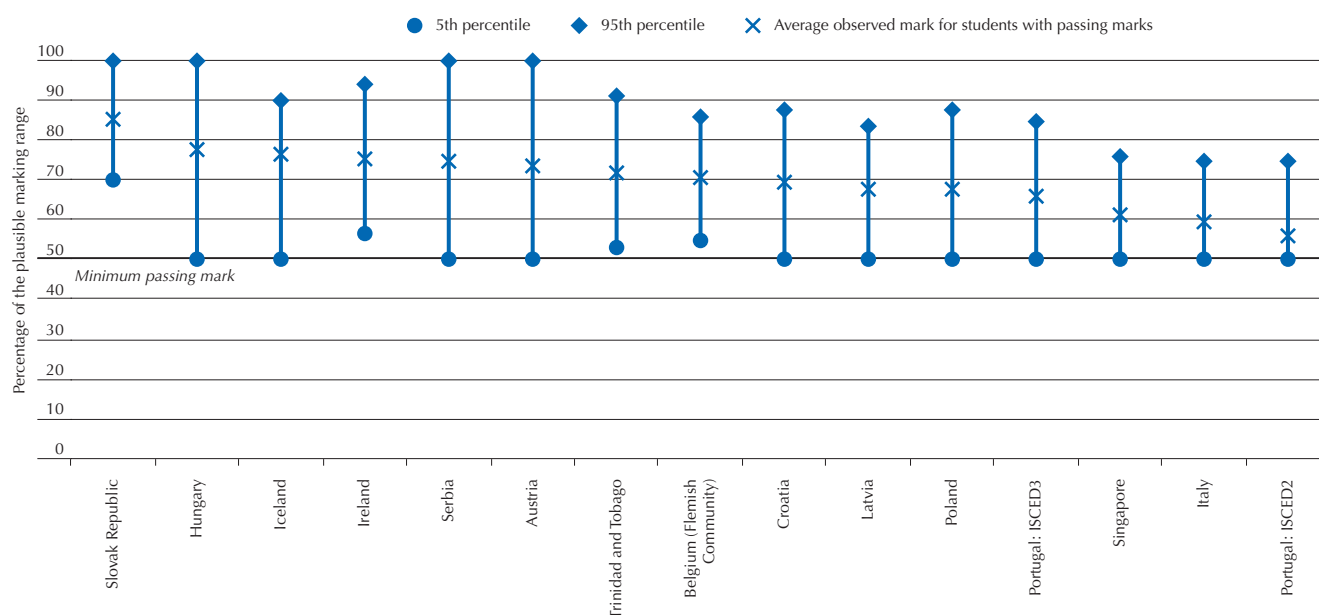
Although school systems vary in the detail with which they describe failing marks, all school systems distinguish passing students in at least four categories (signalling achievement that is “sufficient”, “good”, “very good” or “excellent”) or on a continuous scale. Figure 2.3 presents the distribution of marks for only those students who received a passing mark and presents that distribution on a scale that ranges from 50 (equivalent to the minimum passing mark) to 100 (equivalent to the maximum mark).³ At least 5% of students receive the maximum mark in Austria, Hungary, Serbia and the Slovak Republic, but highest marks are comparatively uncommon and reserved for a smaller percentage of students in the remaining school systems. Highest marks are especially uncommon in Italy, Latvia, Portugal (ISCED 2) and Singapore. In these school systems the highest scoring five percentage of students receive marks that are equivalent to 75% of the maximum passing mark or equivalent to the midpoint of the passing scale.



The average mark received by passing students is close to the midway point of the possible passing values (75% in Figure 2.3) in Austria, Hungary, Iceland, Ireland and Serbia. It is more common, however, that the average passing mark is below this midpoint, signalling that in many countries and economies, the distribution of marks among students who pass their language-of-assessment course is skewed away from the top marks. The average mark in these participating countries is 70% of the plausible range for passing marks. Average marks are in the bottom part of the marking range in Italy, Latvia, Portugal (ISCED 2) and Singapore. Only in the Slovak Republic is the average student more likely to receive a mark that is closest to the top of the marking scale (Figure 2.3).

Figure 2.3

Distribution of marks received by students who passed their language-of-assessment course



Note: Macao-China, Mexico and New Zealand are omitted from this figure because they measured student marks as pass/fail as their marking systems differ by region and/or school.

For comparison purposes, only passing marks are considered; the lowest passing mark is set at 50 and the highest possible mark at 100.

Countries are sorted in descending order of the average passing mark.

Source: Table B2.1 and B2.2.

StatLink <http://dx.doi.org/10.1787/888932733564>

Across participating countries and economies with sufficient data to evaluate distributions, the standard deviation in passing marks is 12% of the plausible range. The distribution of marks is more compressed in Singapore, where the standard deviation in marks represents around 8% of the plausible range in passing marks. By contrast, the spread in marks is greatest in Austria, Hungary and Serbia, all of which have a marking scheme that identifies only four, but clearly distinct, categories of passing marks. Marking schemes that allow for a more subtle distinction between students by offering a wider range of values in which to place students (e.g. Belgium [Flemish Community], Ireland and Trinidad and Tobago), do not necessarily capitalise on the opportunity to better discriminate between students. In these countries and economies, the marks students receive are more concentrated than those received by students who are in school systems where only few values are used; but those few values are used effectively to capture important differences in the performance of passing students (e.g. by distinguishing passing achievement as “sufficient”, “good”, “very good” or “excellent”) (Table B2.2).

As will be discussed later, overall marks within a country, and especially the variation of these marks, need to be interpreted carefully. Marks reflect not only students’ mastery of the courses’ content, but also students’ behaviours and attitudes toward learning and school, and their relative position with respect to their peers in the class or school. In this regard, as marks sometimes have a normative component (students’ marks also convey information about their relative position in the school or class), and the skills relevant to define mastery of content varies across different programmes, variation in marks among education levels and schools and the meaning of marks have to be interpreted in the context of how students are distributed across grades, schools and programmes in a particular country. The section “Towards a model of marks” develops these ideas in evaluating how different student and school attributes influence the marks students receive in their language courses.



Box 2.1 **Standard-Based Assessment, reforming marking schemes and practices in Korea**

In December 2011, the Korean Ministry of Education, Science and Technology (MEST) announced the “Plans to Improve the Secondary School Academic Affairs Management” to meet the demands for creativity and character required in a global, knowledge-based society. A key feature of the plan is a change in the assessment and in-class grading system within Korean schools, known as the Standard-Based Assessment. Prior to this new academic performance assessment method, Korea’s grading practices remained largely normative: assessments and marks were aimed to rank students to see who has achieved more. As a result of the new plan, grading practices have shifted to become criterion-based to evaluate whether, and to what extent, individual students achieved the national curriculum standards developed for each subject. Students in local schools throughout the country are now marked in a uniform five-category scale (with values A, B, C, D and E) that is easily recognisable and understood in Korea and an international setting.

The objectives of the new grading practices involved in the Standard-Based Assessment are multi-fold. It is consistent with the implementation of the 2009 revision of the curriculum that focuses more on creativity and character education. It expands the departmentalised classroom setting and helps create a customised educational environment that considers students’ achievement levels. It also seeks to relieve students from the stress of intense competition with their classroom and school peers, provide students with more reliable information about their aptitude and career plans when choosing school subjects and improve schools’ ability to select students based on their potential, talent and personal situation.

The Standard-based Assessment was introduced for all subjects in lower secondary schools and specialised subject in upper secondary schools in 2012. In 2014 it will be expanded to general subjects in upper secondary schools as well. This change is not free from complications as it changes a long-held tradition of grading and marking. The Korean government plans to strengthen teacher training on the development and implementation of the new framework to provide for a smooth transition. Parents and students will be informed of the backgrounds and goals of the plan to help them understand the new information conveyed by marks.

BEHAVIOURS REWARDED BY MARKS

Through school marks, teachers can reward different sets of student habits, attitudes and behaviours. These include, first and foremost, achievement, or the mastery of the skills and knowledge that students are expected to acquire by the time they reach agreed developmental and academic milestones. In the context of language courses, this includes abilities measured by the PISA reading assessment (e.g. the ability to extract relevant information from texts) as well as the ability to write extended texts, respond to poetry, and interpret different media, among others. Yet teachers may also reward students’ engagement and the use of particular types of learning strategies that may be beneficial for students’ future learning and overall well-being. Teachers may also reward behaviours and attitudes that are not directly related to learning per se, but that may determine the pace of instruction and classroom dynamics, or may be unrelated to what happens in school altogether.

This section explores the relationship between marks and student achievement, learning strategies, engagement, attitudes, relationship to teachers, and other attributes that are considered unrelated to learning. PISA does not have explicit data on what criteria teachers use to assign marks; it is thus impossible to have direct information as to which abilities, attitudes and behaviours are consciously rewarded by teachers. Nonetheless, statistical analyses can illuminate which behaviours and attitudes are best related to marks and indirectly identify which are being rewarded by teachers.

In all countries and economies that distributed the ECQ, marks are positively related to reading performance in PISA: students who score higher in the PISA reading assessment also tend to be rewarded with good marks in their language courses. In other words, in all countries and economies teachers tend to reward mastery of skills and knowledge and the ability to apply language to everyday situations, including through in-class assessments. This relationship is most pronounced in Hungary, Latvia, Poland and the Slovak Republic, where the correlation between marks and reading performance is greater than 0.5. The relationship between marks and reading performance is less strong, yet moderate, in the remaining countries (Figure 2.4). As will be discussed later, a more moderate relationship between marks and PISA reading scores may signal that marks are used to distinguish students *within* a school rather than across the whole student population.

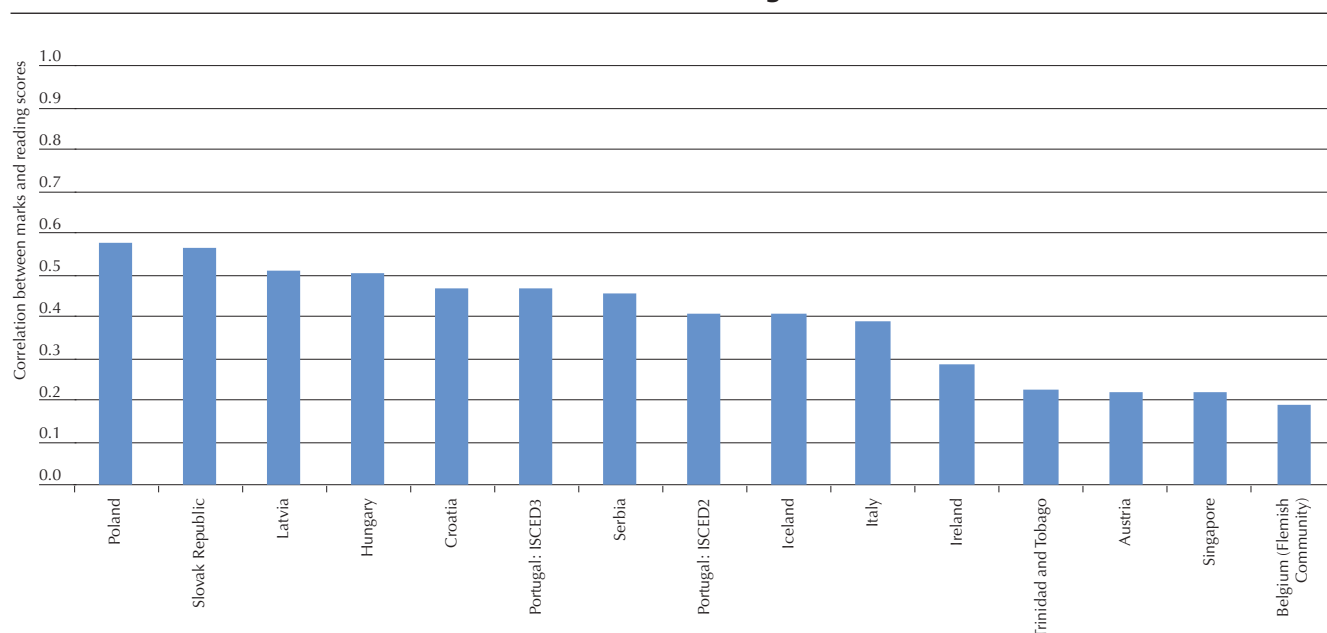
Students’ approaches to learning and learning strategies are also related to the marks they receive. In practically all countries and economies, students who are able to identify effective summarising and understanding and remembering strategies also have higher marks, indicating that teachers in most education systems reward the use of effective learning strategies. This is particularly the case in Hungary, Poland and the Slovak Republic, where the relationship between marks and the use of these strategies is the strongest.



Yet this is not the case in all education systems. In Singapore and Trinidad and Tobago, the relationship is comparatively weak (Table B2.3). For all countries and economies, these relationships remain stable after accounting for students' reading performance (Table B2.4).

■ Figure 2.4 ■

Correlation between the marks students received in their language-of-assessment course and their PISA reading score



Note: Macao-China, Mexico and New Zealand are omitted from this figure because they measured student marks as pass/fail as their marking systems differ by region and/or school.

Countries are sorted in descending order of the percentage of students with failing marks in their language courses.

Source: Table B2.3.

StatLink  <http://dx.doi.org/10.1787/888932733583>

Consistent with the “developmental” approach to teaching and assessment in language-of-assessment courses, teachers reward those students who enjoy reading with high marks. In seven countries and economies there is a moderate to strong relationship between students' enjoyment of reading and the marks they receive. The relationship is particularly strong in Croatia, Hungary, Latvia, Poland and Serbia. Teachers are comparatively less keen to give high marks to students who have good attitudes towards school. In all the countries and economies that participated in this study, the correlation between the index of attitudes towards school and student marks is lower than the correlation between the enjoyment of reading and student marks (Table B2.3). These results remain broadly stable after accounting for reading performance, with the exception of Iceland (Table B2.4).

Although there is a positive relationship in all countries and economies, the correlation between student-teacher relations and student marks is relatively weak. Teachers with a “keeping things moving” approach are more likely to reward these attitudes, as good relationships with the teacher keep the pace of instruction going; yet these rewards provide little direct benefit to developing students' learning potential. The correlation is weak (lower than 0.1) in Croatia, Serbia, Singapore and Trinidad and Tobago. Teachers reward good relationships the most in Hungary, Iceland and Portugal (ISCED 3), yet at very moderate levels (Table B2.3).

The correlation analysis suggests that, in general across participating countries and economies, teachers reward attitudes, behaviours and habits that directly benefit performance. This general trend, however, may mask a sizeable proportion of teachers and schools that reward attitudes, habits and behaviours that may be less beneficial for students' educational careers. For example, while teachers generally reward better reading performance with higher marks across all countries and economies that distributed the ECQ, on average, around 3% of schools have a significantly *negative* correlation between students' reading performance and the marks they receive in their language course. In the Flemish Community of Belgium, Hungary, Ireland and Portugal ISCED 2, more than 5% of schools show a negative relationship between marks and reading performance. There are practically no schools in Iceland, Poland, Serbia and the Slovak Republic where there is a negative relationship between marks and performance (Table B2.5).

With respect to the relationship between marks and both approaches to learning and metacognition strategies, in the Flemish Community of Belgium, more than 25% of schools show a negative relationship between marks and students' identification of metacognitive strategies and positive approaches to learning, while in Austria and Hungary, over 25% of schools show a negative



relationship with four of the five metacognitive strategies and approaches to learning that were evaluated in PISA. By contrast, a large proportion of schools in Latvia, Poland and Portugal show a positive relationship between beneficial metacognitive strategies and approaches to learning and student marks. In this latter set of countries, marks tend to be aligned with those non-cognitive attributes that are best related to student learning (Table B2.5 and B2.6).

TOWARDS A MODEL OF MARKS

School marks provide a strong signal to students about what behaviours, habits and attitudes are rewarded and which are discouraged. Ideally, marks should be positively related to academic performance and to habits that foster student learning, including positive approaches to learning, positive metacognitive strategies, and engagement with reading and school. Yet after accounting for these behaviours and habits – which, in most education systems, are positively associated with marks – other contexts and student characteristics may also be related to student marks.

In many countries, marks have important consequences for students' future educational and occupational prospects as they influence access to higher education and motivate in-school behaviour. Therefore, if teachers systematically favour, for example, socio-economically advantaged students, this inequality in grading practices may contribute to broader social disparities and reduce upward social mobility. On the other hand, marks may promote social mobility if teachers reward disadvantaged students or students from an immigrant background or those from particular social or academic contexts.

Across the countries and economies that distributed the ECQ, about a quarter of the variation in student marks is accounted for by students' performance, learning habits, engagement, background characteristics and contextual factors.⁴ These factors explain a particularly large share of the variation in student marks in Latvia, Poland and the Slovak Republic, where more than a third of the variation in marks is accounted for by these characteristics (Table B2.7).

After accounting for student background characteristics, students' reading performance and learning habits remain strongly associated with the marks students receive. Reading performance is significantly associated with better marks in all countries and economies, and especially so in Latvia, Poland, Portugal (ISCED 3) and the Slovak Republic. Attitudes, habits and behaviours related to engagement and learning habits are associated with better marks, even after accounting for students' performance. These are attitudes and behaviours that promote student learning not only in the classroom but throughout life. The use of control strategies is positively and significantly associated with better marks in 14 countries and economies. Similarly, in 16 countries and economies, especially Serbia, positive attitudes towards reading, as measured by the index of enjoyment of reading, are positively and significantly associated with better marks, after accounting for all other factors. In 13 countries and economies, students who believe that student-teacher relations in their schools are good are also more likely to receive better marks.

After taking into account students' learning habits and behaviours, some student background characteristics appear to be strongly related to marks. This indicates that, in classrooms, teachers reward attitudes and behaviours that are not directly measured in PISA but that are associated with students' background. Girls are more likely to receive better marks than boys in 13 countries and economies, even when they have similar reading performance and learning habits. The exceptions to this trend are Iceland, Portugal (ISCED 2), Singapore and Trinidad and Tobago, where girls and boys obtain similar marks when they have similar performances, attitudes and learning habits. Girls are especially more likely to receive higher marks in Poland and the Slovak Republic. Similarly, students from socio-economically disadvantaged backgrounds are more likely to receive lower marks in 15 countries and economies, with the exception of the Flemish Community of Belgium and Croatia. The association between marks and socio-economic background is strongest in Poland. After accounting for socio-economic status, performance and attitudes towards learning and school, only in Singapore do students with an immigrant background tend to receive lower marks, while in Austria, the Flemish Community of Belgium and Singapore, students who speak another language at home, but are similar in performance and attitudes to their peers, tend to receive lower marks.

The relationship between contextual factors and marks is another potential source of inequality, as students attending certain types of schools will receive lower (or higher) marks depending not on their own performance, effort and engagement, but on factors that are beyond a student's control. Students who attend higher-achieving schools tend to have lower academic self-concepts and receive lower marks (Espenshade et al., 2005; Kelly, 2008; Marsh and Hau, 2003; Marsh et al., 2008). Within the context of PISA 2009, in the majority of countries and economies, students who attend higher-achieving schools receive *lower* marks when compared to students with similar performance and learning habits who attend poorer-performing schools. In these countries and economies, marks are used *normatively*, meaning that students are evaluated in the context of their peers' achievement. This is particularly important because students are not distributed equally across different levels, programmes and school types, and this distribution influences the marks they receive. Thus, comparisons of the distribution of marks across school systems are, in part, determined by the distribution of students across grades, programmes and school types.

The reference group to which the student is compared when marks are used normatively may be the classroom or school. This type of grading is discouraged as it fosters competition that may be detrimental to students' learning (see Boxes 2.1 and 2.2). In this



regard, teachers and schools may be inclined to use marks in alignment with their expectations of how students will perform in standardised, criterion-referenced examinations, so that marks are completely disassociated from students' relative position in the classroom, and are, instead, aligned with criteria defined at the national or regional level. Aligning marks with external criteria can, however, discourage low-performing students if their performance is below these benchmarks, or discourage high-performing students whose performance is far above these benchmarks. The alternative favoured by experts in the field lies at neither end: marks should be concise and clear in their scope and reflect students' achievement of clear pre-established and content-specific criteria. Teachers should thus target assessments to challenge and motivate students to learn, providing them with the information they need to enhance learning, without discouraging them.

Normative marking practices are especially notable in Austria, Poland and Portugal (ISCED 3), where students with similar performance receive marks that are almost one standard deviation lower than those in schools that perform 100 score points higher on the PISA reading assessment. In these countries, students receive marks in comparison to their peers, that is, marks measure relative performance rather than absolute performance. Students tend to be graded based on the distribution of performance within the school, so that, had they attended a poorer-performing school and maintained their performance levels, they would have obtained substantially better marks. In these education systems, and for individual students, marks are less of an indicator of their individual performance and their potential to succeed in their educational careers than an indicator of their relative standing within the school. Research on effective marking practices strongly advises *against* normative grading as it creates incentives for unhealthy competition among students and reduces the motivation to excel. The alternative is to use criterion-referenced marking, where marks reflect students' achievement in well-specified, fair and valid indicators (Guskey and Bailey, 2001; Marzano, 2000; O'Connor, 2002; Tomlinson, 2005) (Table B2.7).

GRADE INFLATION

Do teachers in all schools within a country give marks to students in a similar way? Or are teachers who teach in certain types of schools more likely to give students higher marks, not because of their actual performance but because of other factors? Do teachers in schools that benchmark student achievement with standardised tests give marks that better reflect student performance?

Grade inflation occurs when marks are systematically higher than the sets of acquired skills and abilities that they are measuring would normally merit. Under the assumption that a standardised assessment, such as PISA, provides an unbiased appraisal of the actual level of students' competencies, grade inflation can also be quantified by determining the width of the gap between marks and performance in the PISA test between groups of students.⁵

While anecdotal evidence on grade inflation abounds, studies on grade inflation in secondary schools are scarce. The existing evidence signals that grade inflation is common and that, at least in the United States, it has been increasing since the 1990s (Woodruff and Ziomek, 2004; Zirkel, 1999). Grade inflation may be problematic because in those schools where teachers systematically inflate marks, students receive inaccurate information about their performance and their potential educational careers, and high marks do not distinguish between what is excellent and what is merely good.

In PISA, grade inflation at the school level can be detected by observing the types of schools in which teachers systematically give higher marks to their students, after accounting for performance and other factors that predict successful educational careers. Teachers in certain types of schools may be more apt to inflating marks. Selective and private schools, and those schools that face stronger pressure from parents, may feel coerced into giving students higher marks because they must maintain an image of success for their stakeholders. Schools that do not administer standardised tests may be more likely to give marks to students using benchmarks that are unrelated to national or regional standards. However, as discussed below, there is no evidence to support the notion that grade inflation is pervasive across the countries that distributed the ECQ.

In this context, grade inflation is determined here by measuring whether, when predicting the marks students should receive, based on their reading performance in the PISA assessment, their metacognitive strategies, their approaches to learning, their attitudes towards reading, and school and student socio-economic status, the difference between the observed and the predicted mark is positive. Schools that have a systematic positive residual are schools that inflate marks.⁶ Grade inflation is most common in Austria, Italy, Poland and the Slovak Republic, where one out of eight students attends a school that give marks well above what would be predicted by students' performance, attitudes and behaviours (Table B2.8).

PISA results do not provide evidence of pervasive grade inflation related to the type of school a student attends. In Italy, Portugal (ISCED 2) and Trinidad and Tobago, students in private schools tend to have better-than-expected marks, but this is not true in the other participating countries and economies. Furthermore, only in Italy is the magnitude of grade inflation in private schools quantitatively substantial. In Austria, Croatia and Latvia, selective schools tend to deflate marks relative to performance, possibly because of normative marking practices. Schools that, according to the school principal, are under constant pressure from parents to set high academic standards and have students meet them, and schools that use standardised tests are not systematically more likely to award higher-than-expected marks than schools where parents are less vocal and schools where the use of standardised tests is not common practice.



In a minority of countries and economies, there is evidence of grade inflation in schools that cater to a relatively disadvantaged student population. In these countries and economies, students in socio-economically disadvantaged schools tend to receive better marks than they would have had they attended more advantaged schools, given their reading performance and learning habits and behaviours, and even when students' own socio-economic status is considered when predicting marks. This is most likely due to normative marking practices rather than grade inflation. The exception to this trend is Singapore, where schools that cater to a socio-economically advantaged student body tend to give higher marks after taking into account students' reading performance and their attitudes and behaviours towards learning and school. In most countries, school attributes are not strongly related to grade inflation, indicating that grade inflation is not a pervasive phenomenon (Table B2.8).

CONCLUSION

In PISA, practically all students attend schools that assess students' work or progress through teacher-developed tests and/or assignments that the students return. In almost all countries and economies over 95% of students are assessed this way (see Table IV.3.10 in OECD, 2010). It is thus widely believed that the assessment of students' work in schools provides meaningful information to students. The marks students receive from these assessments are a constant source of information for students, and their parents, about their own achievement, progress in school and, more broadly, their alignment with the schools' expected values, attitudes and behaviours.

The marks students receive vary across countries and economies. Some countries do not differentiate among failing students, so these students do not know how far below the minimum satisfactory criteria they are; other schools allow students to receive marks that go far below the minimum criteria. The literature on the subject suggests that more efficient marking systems do not allow failing marks to fall too far below the minimum satisfactory mark. If the minimum possible mark is too far below the satisfactory mark, students who receive an unusually low mark may have a difficult time recovering from it in order to pass the course – even if they put in the effort and learn throughout the year. Moreover, students receiving a minimum mark at the beginning of the year may become discouraged and disengaged – precisely the opposite reaction that marks are intended to provoke (Brookhart, 2009; Guskey, 2004; Reeves, 2004). Some educators and researchers have sparked debate (“the minimum grading controversy”) by proposing that minimum marks should be established, particularly in school systems that allow the lowest mark to go far below the minimum passing mark. Opponents of minimum marks argue that setting them would promote grade inflation and automatic promotion into higher grades without meeting the academic standards; but research on this issue suggest that minimum marks do neither (Carey and Carifio, 2012).

For those students who do receive passing marks, some school systems differentiate between distinct categories, while others use a relatively continuous scale to locate their students' progress. Results from this analysis suggest that the choice of marking scheme has implications on the degree to which marks can differentiate students who have met the criteria satisfactorily. Although counterintuitive, school systems that attribute a fixed set of four or five clearly distinct marks (e.g. “satisfactory”, “good”, “very good” and “excellent”) describe performance differences among their students more clearly than school systems with relatively continuous marks for passing students (e.g. values ranging from 10 to 20). This is most likely due to the fact that the distinction among qualitatively different categories allows teachers to describe their students more accurately. When the scale is continuous, teachers may interpret different parts of the scale as representing different mastery levels and will tend to restrict themselves to marking around the midpoint, as they may be reluctant to give very high or very low marks (see Box 2.2 for further details on effective marking practices).

Countries and economies use different scales and are some countries are more strict than others with respect to the likelihood of students achieving the highest possible marks. Yet in all countries and economies, the marks students receive in their language-of-assessment courses are positively related to students' reading performance: in all countries and economies, higher reading proficiency is rewarded with better marks. But the information marks convey extends to other attitudes and behaviours. Marks constitute an important lever that teachers and schools can use to motivate students to work towards higher performance, but also towards deeper engagement and more effective attitudes towards learning. In many countries and economies that disseminated the ECQ, students with better control strategies are rewarded with better marks, even after accounting for their reading performance. Similarly, students who are more engaged – those who enjoy reading more – are more likely to obtain higher marks, even when comparing students of similar performance. Thus, teachers and schools reward student attitudes and behaviours that promote learning, in addition to rewarding learning, in and of itself.

Teachers also reward students who believe that student-teacher relations in the school are good, even after accounting for student performance, and may reward attitudes and behaviours that are not directly related to student performance and successful educational careers but that promote a calm learning environment that allows teachers to “keep things moving”.

Similarly, in many countries and economies, marks tend to be higher for girls and socio-economically advantaged students, and are also sensitive to the academic context of the school, even after accounting for individual students' performance, attitudes and behaviours towards learning. The fact that marks are sensitive to factors that are unrelated to students' performance, engagement and learning habits signals that teachers may reward aspects that they feel are important but are not measured directly by PISA and



that are strongly related to students' backgrounds. Teachers may also reward behaviours that are valued in the labour market and in other social environments. As marks constitute one of the most reliable and consistent indicators of students' own performance and potential, systematic inequalities in the allocation of marks may contribute to systematic inequalities in educational expectations, as discussed in the following chapter. Box 2.1 on Korea's initiative to reform marks earlier on this chapter signals the potential changes a reform to marking systems may bring to a school system if aligned to a wider reform project.

This chapter has highlighted how countries and economies that distributed the ECQ use marks. As will be shown in the following chapter and as underlined by the academic literature on the subject, marks may constitute an important source of information for students about their future prospects in education. This and the following chapter do not, however, provide information on how teachers and students use marks and the information conveyed by marks. In effect, this chapter works under the assumption that marks are most beneficial when they convey useful information to students; but as important as the usefulness and legitimacy of the information conveyed by marks is how students and teachers actually use marks. Box 2.2 develops this point inasmuch as effective marking practices depend not only on the clarity and usefulness of the information conveyed by marks, but on the assumption that this information is used for the purpose of enhancing and improving learning. At the system level, assessments and how marks are used should be consistent with the broader policies on student assessment.

Box 2.2 Effective marking practices

What are effective marking (grading) practices? Ideally, they should communicate clear, useful information *for the purpose of enhancing learning*. As such, marks should motivate students to put forth more effort, to identify their strengths and encourage them to work on their weaknesses. In general, marks should provide accurate, specific and timely feedback. They should be a final judgment about a student's achievement. Marks should be clear and concise in their scope. They should be based on clearly-specified criteria (criterion-based), measuring student achievement against a set of clear, pre-established and content-specific goals. Students should not be evaluated relative to their peers (normative grading, or grading on a curve), as normative grading spawns unhealthy competition and diminishes motivation (Marzano, 2000; O'Connor, 2007; Reeves, 2008).

Moreover, marks should not be a synonym for assessment – the gathering of information about students' achievement for the purpose of making instructional or career-path decisions. The evidence used to gather marks should be valid, measuring what they are intended to measure and not other factors (e.g. neat hand-writing, good behaviour, or pace of work in assessments that measure reading comprehension). The instrument to evaluate student achievement should reduce measurement error as much as possible, such that the task directions are clear and the "test" is aligned with its goals. Marks should reflect achievement; and if they are to reflect behaviour, behaviour and achievement should be reported separately, with an emphasis on intrinsic motivation. In fact, once achievement and behaviour are explicitly decoupled from marks, both achievement and behaviour tend to improve (Guskey and Bailey, 2001; Marzano, 2000; O'Connor, 2002; O'Connor, 2009; Tomlinson, 2005).

Many common grading practices have been labelled as detrimental to student learning. Considering that one of the goals of grading is to motivate students to learn, using marks as punishment, such as using the minimum mark as punishment for late or missing work, is strongly discouraged. The use of an average of all the marks given during the semester or school year should be evaluated, inasmuch it presumes that learning at the beginning of the term is equally important as learning at the end of the semester. Students should be evaluated with respect to their understanding at the end of the semester. Yet the use of a one-and-only evaluation (a final test, essay or research project) reduces student motivation to work throughout the entire semester (Guskey, 2004; Reeves, 2004; Reeves, 2008).

Because these effective grading practices may be hard to achieve, some researchers argue that the only effective grading practice is replacing marks by a individualised narrative assessment for each student. Eliminating the traditional notion of marks is a necessary, but not sufficient, step to promote deep thinking and engagement in learning. Marks, they contend, diminish students' interest in learning, create a preference for the easiest possible task, and reduce the quality of students' thinking. Schools can promote better learning environments by eliminating marks gradually and, if final grades are required, arrange for collaborative grading, where students mark themselves at the end of the year upon the evaluation of a portfolio and negotiate this mark with the teacher who has the final say (Kohn, 2011). The elimination of marks requires a gradual process and teachers' willingness and ability to adopt a change in incentives, behaviours and dynamics in the classroom.

Sources: Guskey and Bailey (2001); Guskey (2004); Kohn (2011); Marzano (2000); O'Connor (2002); O'Connor (2007); O'Connor (2009); Reeves (2004); Reeves (2008); Tomlinson (2005).



Notes

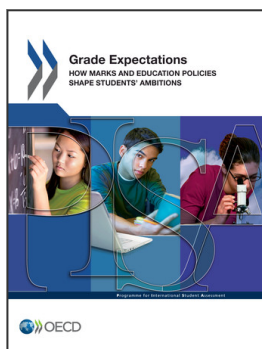
1. Among all PISA countries and economies, 21 countries implemented the optional Educational Career questionnaire. Of these countries and economies, Australia, Hong Kong-China, Korea and Slovenia chose not to include the question that asked students about their marks. The results in this chapter correspond to the 17 countries and economies that did include this question. Different marking systems exist in Portugal (in ISCED level 2 programmes marks range from 1 to 5, and in ISCED level 3 programmes marks range from 1 to 20), so results for Portugal are disaggregated by ISCED level to accommodate the different marking schemes.
2. In Ireland, teachers often assign letter grades rather than points on a 0-100 point scale to students. These letter grades are then converted to the 0-100 point scale (e.g. A = 85-100, B = 70-84, C = 55-69, D = 40-54, E = 25-39, etc.).
3. These analyses exclude Macao-China, Mexico and New Zealand because they asked students about whether they received a passing or failing mark and not the specific mark. It is thus impossible for analyses to infer the distribution of passing marks in these school systems. This exclusion and the conversion of the data are necessary to render comparable information that is specific to each school system.
4. These results are based on a regression model predicting students' marks in their language courses, considering their reading performance, learning habits and strategies, engagement in reading and school, student-teacher relations, student background characteristics, and school contextual factors. For education systems that use a pass/fail marking scheme, a logistic regression predicting whether students received a passing mark is used instead.
5. Grade inflation is typically understood as the increase in grades over time with no corresponding change in the underlying abilities that these marks measure. PISA does not have information on student marks to evaluate the change in marks over time compared with the change in test scores over time. However, grade inflation is a comparison between two groups – students at one moment in time compared with students at another moment in time. This report uses the comparison of groups implicit in grade inflation to identify the phenomenon by comparing residuals cross groups. This alternative approach to grade inflation is not inflation with respect to the past, but with respect to another theoretically relevant group.
6. Schools that have an average residual greater than 0.5 in a regression predicting student marks over reading scores, metacognitive strategies, approaches to learning, attitudes towards reading and school, and student socio-economic status, are considered schools that inflate grades.

References

- Bonesrønning, H.** (1999), "The Variation in Teachers' Grading Practices: Causes and Consequences", *Economics of Education Review*, No. 89, pp. 89-105.
- Brookhart, S.** (1993), "Teachers' Grading Practices: Meaning and Values", *Journal of Educational Measurement*, Vol. 30, No. 2, pp. 123-142.
- Brookhart, S.** (2009), *Grading*, Merrill, New York, New York.
- Carey, T. and J. Carifio** (2012), "The Minimum Grading Controversy: Results of a Quantitative Study of Seven Years of Grading from an Urban High School", *Educational Researcher*, Vol. 41, No. 6, pp. 201-208.
- Covington, M.** (1984), "The Self-Worth Theory of Achievement Motivation: Findings and Implications", *The Elementary School Journal*, Vol. 85, No. 1, pp. 4-20.
- Docan, T.** (2006), "Positive and Negative Incentives in the Classroom: An Analysis of Grading Systems and Student Motivation", *Journal of Scholarship of Teaching and Learning*, Vol. 6, No. 2, pp. 21-40.
- Espenshade, T., L. Hale and C. Chung** (2005), "The Frog Pond Revisited: High School Academic Context, Class Rank, and Elite College Admission", *Sociology of Education*, Vol. 78, No. 4, pp. 269-293.
- Fredricks, J., P. Blumenfeld and A. Paris** (2004), "School Engagement: Potential of the Concept, State of the Evidence", *Review of Educational Research*, Vol. 74, No. 1, pp. 59-109.
- Guskey, T.** (2004), "Zero Alternatives", *Principal Leadership: High School Edition*, Vol. 5, No. 2, pp. 49-53.
- Guskey, T. and J. Bailey** (2001), *Developing Grading and Reporting Systems for Student Learning*, Corwin, Thousand Oaks, California.
- Kelly, S.** (2008), "What Types of Students' Efforts Are Rewarded with High Marks?", *Sociology of Education*, Vol. 81, No. 1, pp. 32-52.
- Kohn, A.** (1993), *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise and other Bribes*, Houghton Mifflin Co., Boston, Massachusetts.
- Kohn, A.** (2011), "The Case against Grades", *Educational Leadership*, Vol. 69, No. 3, pp. 28-33.
- Marsh, H. and K. Hau** (2003), "Big Fish Little Pond Effect on Academic Self-Concept: A Crosscultural (26 Country) Test of the Negative Effects of Academically Selective Schools", *American Psychologist*, Vol. 58, No. 5, pp. 364-376.



- Marsh, H.**, et al. (2008), "The Big-Fish-Little-Pond-Effect Stands up to Critical Scrutiny: Implications for Theory, Methodology, and Future Research", *Educational Psychology Review*, Vol. 20, No. 3, pp. 319-350.
- Marzano, R.** (2000), *Transforming Classroom Grading*, Association for Supervision and Curriculum Development, Alexandria, Virginia.
- McMillan, J., S. Myran and D. Workman** (2002), "Elementary Teachers' Classroom Assessment and Grading Practices", *The Journal of Educational Research*, Vol. 95, No. 4, pp. 203-213.
- O'Connor, K.** (2002), *How to Grade for Learning: Linking Grades to Standards*, Skylight, Arlington Heights, Illinois.
- O'Connor, K.** (2007), *A Repair Kit for Grading: 15 Fixes for Broken Grades*, Educational Testing Service, Portland, Oregon.
- O'Connor, K.** (2009), "Reforming Grading Practices in Secondary Schools", *Principal's Research Review*, Vol. 4, No. 1, pp. 1-7.
- OECD** (2010), *PISA 2009 Results: What Makes a School Successful? Resources, Policies and Practices (Volume IV)*, PISA, OECD Publishing.
- Reeves, D.** (2004), "The Case Against Zero", *Phi Delta Kappan*, Vol. 86, No. 4, pp. 324-325.
- Reeves, D.** (2008), "Effective Grading Practices", *Teaching Students to Think*, Vol. 65, No. 5, pp. 85-87.
- Rosenbaum, J.** (2001), *Beyond College for All: Career Paths for the Forgotten Half*, Russell Sage Foundation, New York, New York.
- Stiggins, R. and N. Conklin** (1992), *In Teachers' Hands: Investigating the Practices of Classroom Assessment*, State University of New York Press, Albany, New York.
- Tomlinson, C.** (2005), "Grading and Differentiation: Paradox or Good Practice", *Theory into Practice*, Vol. 44, No. 3, pp. 262-269.
- Woodruff, D. and R. Ziomek** (2004), *High School Grade Inflation from 1991 to 2003*, ACT Research Report Series 2004-4.
- Zirkel, P.** (1999), "Grade Inflation: A Leadership Opportunity for Schools of Education", *Teachers College Record*, Vol. 101, No. 2, pp. 247-260.



From:

Grade Expectations

How Marks and Education Policies Shape Students' Ambitions

Access the complete publication at:

<https://doi.org/10.1787/9789264187528-en>

Please cite this chapter as:

OECD (2012), “What Behaviours Do Teachers Reward?”, in *Grade Expectations: How Marks and Education Policies Shape Students' Ambitions*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264187528-5-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.