Is pupil assessment a mechanism for improving school performance?

Sheila González Motos

Assessment has traditionally been understood as the evaluation that teachers conduct regarding what their pupils have learned. Today, assessment is no longer a purpose-driven tool (for giving marks) and has become something more complex: another resource for learning. However, the breadth and variety of tools for assessments suggests that we investigate and reflect on which practices have a greater impact. This review of the evidence aims to provide information that can help us to find out if educational assessments are tools for improving pupils' performance in school, discover which assessment mechanisms are more effective in improving academic results and learn which pupil profiles benefit the most from implementing the different strategies of evaluation in the classroom.

"For too long, education has been based on inertia and tradition, and changes in educational intuitions or beliefs were unfounded. The 'what works' movement enters into the world of education with a clear objective: to promote evidence-based educational policies and practices. Ivàlua and the Jaume Bofill Foundation have joined forces to promote the movement in Catalonia."

ivàlua✔ | Institut Català d'Avaluació
de Polítiques Públiques

What Works
in Education?

Using Evidence to Improve Education

Is pupil assessment a mechanism for improving school performance?

Sheila González Motos
PhD in Political Science and Social Transformation. She has conducted research in the sphere of education and immigration, with special interest in topics of social inequality. She is currently a postdoctoral researcher and professor at the Sociology Department of the Autonomous University of Barcelona.

Motivation

Learning assessments are part of the work of every teacher and are becoming a central part of the educational system, given their implications for pupils to obtain certain qualifications, degrees and/or accreditations. At the same time, regardless of the tools used, it is agreed that assessments are one of the tasks that teachers spend the most time on, which suggests that we identify which evaluation methods are the most efficient and provide a better way to use the time dedicated to them. In addition, in a context like the current one, where educational approaches that question evaluations as a central part of the school system are gaining strength, it is interesting to reflect on the educational value of different assessment mechanisms.

Assessments have traditionally been understood as the evaluations that teachers conduct regarding what their pupils have learned at the end of a roughly broad period (teaching unit, semester or academic year), either through written exams, oral exams, individual work or group activities, as well as other tools. However, different educational trends have been incorporating new elements, turning assessments into something more complex: this does not always occur at the end of a period, but may also be done beforehand (diagnostic assessment) or progressively (continuous assessment). It is not always conducted by teachers, but can also be done by classmates (peer assessment) or by bodies outside the school (basic skill exams, PISA, etc.) and especially by the pupils themselves (self-assessment). Most importantly,

assessments are no longer a purpose-driven tool (for giving marks) and have become just another resource for learning.

It is in this vein that the OECD carried out extensive research on the impact of evaluation methods developed in most countries not only with purpose-driven aims, but as tools for learning [1]. Much has been written on this subject, but there are few rigorous studies that indicate a clear line of action in terms of assessment. The quality of the evidence is generally low and the conclusions drawn are not univocal [1] [2]. Nevertheless, some studies provide interesting data and reflections for the subject that concerns us. This review of the evidence focuses on assessment tools of a formative nature, meaning those that are not limited to giving marks, but that also become part of the learning process. It is therefore about addressing the impact on performance in school and not about finding the best mechanisms to verify what has been learned.

This review of the evidence focuses on assessment tools of a formative nature, meaning those that are not limited to giving marks, but that also become part of the learning process. It is therefore about addressing the impact on performance in school and not about finding the best mechanisms to verify what has been learned.

Which assessment methods are we talking about?

As we have pointed out, assessment is a very broad field that includes many tools and purposes. From the start, we can very generically identify three types of assessments. First is the formative assessment, used by teachers and pupils during the learning process in order to adjust the dynamics of teaching and study. Second is the summative assessment, which takes place at the end of a unit, semester or school year and is aimed at verifying that a certain level of knowledge has been achieved (based on different criteria). Finally, the standardized or interim assessment aims to provide educators and policy makers with comparable information between classes, schools, districts or countries regarding the pupils' level of performance or knowledge.

However, this does not mean that these evaluation rationales are mutually exclusive or incompatible, since both summative and interim assessments may be used for formative purposes. Conversely, mechanisms designed for formative assessments lose their formative value when teachers are limited to giving out marks or grades [3] [4].

Table 1.
Types of evaluations

Type of evaluation	Objective	Beneficiary	When
Formative assessment	(Re) guide the teaching-learning process	Pupils and teachers	During the learning process
Final or summative assessment	Identify if learning was achieved	Pupils	At the end of the process (unit, semester, school year, etc.)
Standardized or interim assessment	Compare achievements between classes, schools, districts, countries, etc.	Teachers and policy makers	Any time

Source: author

As stated above, this review focuses on formative assessment or formative feedback.[1] There is no single and complete definition of what we understand as formative assessment. The vague use of the term has generated an enormous literature, but also a lack of agreement with respect to its results. We can say that the vagueness of its definitions has led to a weakness in the empirical results and has made it difficult to identify good practices [3]. Nevertheless, there is a consensus to classify tools used in the educational sphere as formative assessment tools when they respond to a dual objective: enhancing knowledge and skill acquisition and boosting motivation for learning [5]. Assessment tools designed for this dual purpose mainly fit within this framework, but there is also space for other methods that develop their formative potential based on other aims more typical of interim or summative assessments.

There is a consensus to classify tools used in the educational sphere as formative assessment tools when they respond to a dual objective: enhancing knowledge and skill acquisition and boosting motivation for learning.

These two objectives (knowledge and motivation) may be achieved with different evaluation models that seem opposed at first glance, but are complementary in practice:

- The feedback may be **normative**, comparing a pupil's performance with their rest of the classmates or group, or **self-referenced**, comparing their performance with their abilities.

- The message that accompanies the feedback makes it possible to distinguish between **directive feedback**, meaning the pupil's indications about what he or she needs to review, and **facilitative feedback**, the set of comments and suggestions that help to guide their own process of improvement.

- In some cases, the assessment focuses on the exercise presented by a pupil so that the corrective information provided by teachers is addressed to each pupil

---

1   Although there are authors who identify differences between both concepts, most of the literature reviewed uses them indiscriminately as synonyms.

regarding their specific exercise. In other cases, however, the feedback is not produced individually, but in aggregate form, meaning that evaluations of the development of the group in general are presented to the class as a whole in an exam conducted on the individual level.

- A fourth category differentiates between a **learning-oriented assessment** that informs pupils how close or far they are from achieving the final aims laid out for their education and a **performance-oriented assessment** that is limited to informing them of the results of the exams conducted.

- Although most assessments in the school system are carried out by teachers, there are other formulas for evaluation such as **self-assessment** and **peer assessment**. The first kind of assessment is based on the pupil's ability to understand the teaching objectives, evaluate the tasks performed, detect mistakes and correct them. Peer assessment involves the review of one's work by one's peers based on a set of evaluation criteria previously discussed in a group in order to establish a common perspective for assessment.

Table 2.
Types of formative assessments

Types of assessments
Normative Assessment / Self-Referenced Assessment
Directive Assessment / Facilitative Assessment
Individual Assessment / Group Assessment
Learning-Oriented Assessment / Performance-Oriented Assessment
Standardized Assessment / Self-Assessment / Peer Assessment

Source: author

Apart from these models, it must be taken into account that there is a very wide range of subjects to evaluate on which formative feedback is based. Hattie and Timperley [6] identify three levels where formative feedback may be applied:

- **Specific task developed by the pupils**: this indicates the degree of the pupil's performance in a specific exercise, showing the right and wrong answers.

- **Learning process:** on this level, the assessment goes beyond evaluation of the specific exercise and is aimed at the learning process that connects the different tasks that must be performed, providing strategies for detecting mistakes and seeking solutions.

- **Self-regulatory capacity**: this includes in the assessment aspects linked to the pupil's commitment and organizational ability based on indications of how to improve autonomy, discipline and self-control in order to achieve the agreed learning objective.

The reviewed literature echoes this wide range of options with regard to assessment. An important part also incorporates into the analysis other variables with which the assessment interacts, such as the pupils' academic level, age or type of studies.

Questions that guide the review

Evaluation has traditionally been one of the organizational pillars of the school system, and therefore of teaching as well. In recent years, two parallel and antagonistic processes have been taking shape: on the one hand, international institutions and public authorities have been demanding more information on performance in schools by introducing new diagnostic exams; while on the other hand, educational movements that question the usefulness of both external and internal assessments have been gaining force. This review of the evidence aims to provide data to this debate, while responding to the following questions: 1) Can formative assessment be a tool for improving pupils' performance in school? 2) Which evaluative mechanisms are most effective in improving academic results and other skills? 3) Which profiles benefit the most from implementing different evaluative strategies in the classroom? and 4) In terms of public intervention, is it recommendable to invest in assessment? And if so, under which criteria?

Review of the evidence

Reviews and studies considered

Despite the enormous academic output about assessment in school environments that has been generated from different fields (education, psychology, sociology, etc.), there is little solid evidence to justify clear conclusions about it. This study is based on the reviews, meta-analyses and experimental studies collected in Table 3, primarily produced in the English-speaking world (even though many of the studies included have been developed in other contexts, such as Spain) and characterized by basing their contributions on solid methodological approaches.

The review is basically composed of nine reports that thoroughly review a high number of experimental or meta-analytic studies and five experimental studies that have made especially important contributions to the subject concerning us. Along with the evidence referring to compulsory stages of education, many of the reviews include research and meta-analyses about stages of higher education (university) without clearly differentiating the net effects in each stage. The inclusion of experimental studies in our review enables us to focus on primary and secondary education, which is its subject.

It must be taken into account that the methodology varies between the different studies reviewed. Although most of the literature we reviewed addresses the

comparison between purpose-driven assessment mechanisms and tools of formative feedback, in some cases there is a deeper understanding that lets us compare different types of formative feedback to identify those with greater impact. In some cases, the comparison also includes control groups without formative assessment or feedback in order to measure the differential impact of formative feedback and purpose-driven assessment compared to a lack of evaluation.

Table 3.
Studies reviewed

Authors	Type of document
Klute *et al.* (2017) [7]	Review
Elliot *et al.* (2016) [2]	Review
Dunn & Mulvenon (2009) [3]	Review
Hattie & Timperley (2007) [6]	Review
Shute (2007) [5]	Review
Ross (2006) [8]	Review
Dochy *et al.* (1999) [9]	Review
Black & Wiliam (1998) [10]	Review
Kluger & DeNisi (1996) [11]	Review
Meusen-beekman, *et al.* (2016) [12]	Experimental study
De Marcos, *et al.* (2010) [13]	Experimental study
Wiliam *et al.* (2004) [14]	Experimental study
McDonald & Boud (2003) [15]	Experimental study
Schunk (1996) [16]	Experimental study

Source: author

Can assessments improve pupils' educational outcomes?

The studies conducted on the impact of assessments on pupils' performance in school tend to identify a major improvement in results for reading, mathematics and the ability to retain information. While some research concludes that pupils subjected to quality formative assessments and feedback may boost their academic progress after around eight months compared to the performance achieved in an academic year, the most recent meta-analyses situate additional progress at around three months, and at around four months if the teachers have received specific training to conduct this formative assessment [17].

The studies conducted on the impact of assessments on pupils' performance in school tend to identify a major improvement in results for reading, mathematics and the ability to retain information.

↗

With greater or lesser intensity based on variables that we address in the sections below, formative assessment has an impact on performance and outcomes in school [7] [10]. Among the curricular areas considered, formative assessment shows the greatest effect on mathematics, while its impact on reading and writing is lesser, though still observable and significant [7].

The effect of self-assessment on education has received special attention in the literature, which mainly agrees that it has a positive impact, even though, in some cases, the way the studies are designed suggests that we be careful with the results [3] [16] [8]. In general, however, the data point to an improvement in the academic results of the groups that use self-assessment mechanisms [15], although they are inferior to teacher-led evaluations [7].

The results for peer assessment are also positive. This type of evaluation is often more intelligible for pupils, given that the language and expression of the corrector-pupil are more familiar than those of teachers. Additionally, these kinds of assessments not only enable pupils to evaluate the work of a peer, but they also provide information about their own performance through comparison, thereby having an impact on their own development. In this case, however, some studies have also detected negative effects in this type of assessment, like anxiety or stress, due to the tension stemming from evaluating others and from comparing their results with one's own [12].

Self-assessment and peer assessment have distinct effects depending on the skill worked on. Thus, while self-assessment and peer assessment are effective for working on mathematics, the results are not as clear for other subjects like reading or writing [7].

A lot of recent research has focused attention on what new technologies have brought to education (m-learning) in general and to assessment (especially self-assessment) in particular. Although most of it focuses on online learning contexts, some experimental studies have been developed in face-to-face educational environments. In these cases, digital tools are configured as a supplement to face-to-face learning, either by facilitating initial diagnostic exams, opening possibilities for self-assessment or expanding feedback among educators and pupils [13]. The data show that groups that use new technology systems for self-assessment get better final marks than groups that have no evaluation, with different levels of intensity according to the age and type of study, as we will discuss later.

As we pointed out above, despite the amount of research, the conclusions are general and the mechanisms for improvement are not always clear. Most of the literature

concludes that formative assessment and feedback lead to an increase in pupils' performance and learning [5] [7] [10], but there are also studies that do not identify any effect and some that even obtain negative effects [5] [11]. As exceptions to the general rule, these effects are explained by two arguments. First, they are explained by the existence of other variables that interfere in the effects of assessment, like the quality of the teaching or the pupil's previous level of learning [14]. In this regard, various studies have pointed to the interaction between previous levels of knowledge and what is taught in the classroom, finding that pupils with different previous levels of knowledge require distinct learning strategies and respond differently to structures and content. This is why it is important to be able to evaluate this previous knowledge as a conditioning factor for learning.

Second, mistakes in how the assessment is designed could explain its lack of effectiveness and especially its negative impact on learning. For example, there seems to be agreement in identifying assessments based on marks and not on qualitative feedback as detrimental, but unadvisable dynamics are also observed according to to different moments, ages or types of study, as we will discuss below [5].

What are the features of an effective assessment?

The wide range of assessment methods and approaches makes it difficult to establish a consensus on the overall impact of formative assessments on performance in school, yet at the same time, the large number of tools evaluated enables the reviewed research to identify some factors of assessment that help to improve it:

- **The different assessment tools have a differential impact on pupils' performance**: learning improves when basic correction (correct/incorrect) is accompanied by facilitative feedback, meaning by clues so pupils can find the right answer, since uncertainty in the face of failure can cause frustration and lower the motivation to learn. In this regard, explaining the specific mistake without using generic phrases or referring only to the general standard also noticeably improves learning. Moreover, tools such as verification (correct/incorrect) and repetition of the exercise until the correct answer is given make less of an impact or none at all. Feedback based on reward or punishment has little effect. The effects seem to vary based on the type of learning: when improving the same exercise, mechanisms of repetition are more effective, while the effect is transferred to the improvement of different tasks when the assessment tools enjoy greater support [5] [6].

- **Self-referenced feedback is more effective, since it places pupils within their own learning process**. Like all other normative feedback mechanisms, giving marks becomes a mechanism of lesser impact and often leads to negative results. In the specific case of giving marks, a reduction in the impact of the assessment is observed, since it focuses the pupil's concern on the mark or grade and not on the formative comments that come with the correction [2] [5].

Like all other normative feedback mechanisms, giving marks becomes a mechanism of lesser impact and often leads to negative results.

↗

- **Feedback must be adapted to the pupil to whom it is addressed**. When there is no consistency between the message and the recipient, feedback is not observed to have an impact on the pupil's performance. For example, high-level pupils benefit from lower levels of detail in feedback, while the most motivated perform better with more complex mechanisms of evaluation. Thus, low-performing pupils benefit more from immediate feedback or correction, while high-perform-ing pupils seem to benefit more from delayed feedback [5].

- **The research is not conclusive about the advisable level of preparation and complexity for giving feedback on the assessment**. The reviewed literature concludes that any evalua-tive feedback has a greater effect than simple correction and that the complexity of the mechanism used has a differen-tial effect depending on what is being assessed. However, the conclusions re-garding the level of complexity are not clear and some studies find that there is no differential effect based on the level of complexity [5].

> **The reviewed literature concludes that any evaluative feed-back has a greater effect than simple correction and that the complexity of the mechanism used has a differential effect depending on what is being assessed.** ↗

> **Feedback has a greater impact when the objectives of learning are very clear. The assessments therefore appear as a comple-ment to the formative dynamics developed in the classroom.** ↗

- **Feedback has a greater impact when the objectives of learning are very clear**. The assessments therefore appear as a complement to the formative dynamics de-veloped in the classroom, stressing their effects when the educational activity as a whole is organized according to some established aims [11].

- **The effects of assessments on performance in school appear higher when the com-plexity of the task is low**. The more complex an activity becomes, the lower the effect that formative assessments have on performance in school [11].

- **Learning-oriented assessments that are not limited to evaluating performance have a greater impact on each exam**. Assessments and feedback that help to im-prove performance focus on learning as a whole and not when they are restricted to correcting performance in each exam. In the latter case, the impact may be neg-ative or inconsistent [5].

- **Assessments and feedback can be immediate or delayed**. Some research finds that immediate feedback prevents mistakes from being memorized, while others think that it can interfere prematurely, without giving the pupils time and space to process and correct the mistake themselves. In addition, several studies recommend varying the time when the correction is made based on the pupil's profile: while low-performing pupils benefit more from immediate correction, higher-performing pupils improve most clearly when the feedback is delayed [5].

- **There is agreement that formative assessment provided continuously throughout the learning process is the type of evaluation with the greatest impact**. While final assessments and diagnostic assessments are also im

> **Though self-assessment, peer assessment and teacher-led assessment are effective, there is a greater impact on perfor-mance in school when it is the teachers who conduct and monitor the pupils' progress.** ↗

portant, several studies on evaluating progress or continuous assessment indicate a significant reduction in the impact of individual factors on performance based on the completion of progress exams with the corresponding feedback through the school year [9].

- Though self-assessment, peer assessment and teacher-led assessment are effective, **there is a greater impact on performance in school when it is the teachers who conduct** and monitor the pupils' progress [7].

Box 1.
Self-assessment training and educational impacts: experiment in Barbados

Betty McDonald and David Boud [13] conducted an experiment in 10 secondary schools on the island of Barbados. In order to evaluate the impact of self-assessment skills on the pupils' performance in school, they designed a self-assessment training program for pupils and teachers.

They selected 10 schools according to the results obtained in external assessment exams administered each year by the Caribbean Examinations Council and secondary education entrance exams that every pupil must complete in order to choose a school (Barbados Secondary Schools Entrance Examination). These 10 schools can be categorized as high-, mid- and low-performing according to these results.

Two class-groups from the final year of compulsory secondary education were selected for each school, with between 30 and 40 pupils. School after school, the researchers reviewed the previous years' assessments of the pupils of both classes to ensure that there was no grouping by levels that could distort the sample with bias. In each school, each of the two classes was selected randomly as either the control group or the test group. Thus, a total of 256 pupils were trained in self-assessment skills, while the control group (259 pupils) was not given such training.

The test group was trained in self-assessment during the school year as part of the ordinary school curriculum based on 12 modules designed by the researchers. In the same way, the teachers were instructed to provide pupils with this training. The control group followed an ordinary curriculum that did not include this training.

At the end of the year, all the pupils took the Caribbean Examinations Council (CxC). The results achieved by both groups differed significantly. The test group (which was trained in self-assessment) obtained higher averages in all the curricular subjects evaluated (Economics, Humanities, Science and Technology).

The researchers say that it is necessary to take into account the unpredicted effects that self-assessment training may have had on other variables that can also explain performance in school, such as the motivation of the teachers involved in the process or the motivation of the pupils in the test group.

For further information:

McDonald, B. and Boud, D. (2003)." The Impact of Self-assessment on Achievement: The effects of self-assessment training on performance in external examinations", *Assessment in Education. Principles*, Policy and Practice, 10(2), pp. 209-220..

Which pupils are most sensitive to formative assessment?

Rather than indicating which pupils are most sensitive to formative assessment, we should ask ourselves which profiles benefit the most from the different formative assessment tools. In other words, what effects do these tools have on different types of pupils?

- High-performing pupils achieve better results with less intense mechanisms (with simple verification of right or wrong). Low-performing pupils, however, learn more with more intense mechanisms of facilitative feedback, which include detailed explanations of mistakes and give pupils the chance to correct them. Mid-level pupils do not show any major differences in their performance based on the feedback mechanism used. In any case, all three pupil profiles show substantial improvement compared to those who received no feedback at all [5].

  High-performing pupils achieve better results with less intense mechanisms (with simple verification of right or wrong). Low-performing pupils, however, learn more with more intense mechanisms of facilitative feedback. ↗

- Low-performing pupils benefit more from immediate feedback or correction, while high-performing pupils benefit more from delayed feedback [5].

- Low-performing pupils are more sensitive to the type of feedback that they receive. When the feedback is normative (compared to the class as a whole), low-performing pupils tend to attribute their results to a lack of ability and do not expect to get better results in subsequent exams, which has a negative effect on their motivation. However, self-referenced feedback (compared to the pupil's own abilities) directly links the pupil's results to the effort made, forcing him or her to focus on their personal progress [5].

  Low-performing pupils are more sensitive to the type of feedback that they receive. When the feedback is normative, low-performing pupils tend to attribute their results to a lack of ability and do not expect to get better results in subsequent exams, which has a negative effect on their motivation. ↗

- Research focused on pupils with disabilities draws clearer conclusions than when ordinary pupils are involved. Therefore, despite the disagreement between some studies regarding the differential impact compared with the population as a whole, there is unanimous agreement among studies indicating that special education pupils benefit intensively from the introduction of formative assessment methods [3] [10].

- The pupils' socio-economic levels have scarcely been addressed in the reviewed literature, but studies focusing on schools in the most impoverished social settings identify impacts above the average for the studies as a whole, which could indicate an improvement in these pupils' performance higher than that of the whole [3] [10]. Thus, the notion that poor pupils are clearly sensitive to the effects of assessment is confirmed.

- Some research points to the fact that formative feedback only has a greater impact on the youngest pupils. Even so, identification of these ages varies depending on the study. However, there is agreement that formative assessment requires learning so that the sooner it is started, the greater the impact will be. Moreover, for assessments based on digital methods, the youngest pupils are found to be more

FUNDACIÓ JAUME BOFILL

ivàlua | Institut Català d'Avaluació de Polítiques Públiques

familiar with new technologies and this familiarity is a motivating factor that boosts their impact [3] [13].

In addition to performance, what other skills can improve assessments?

Research on formative assessment and feedback focuses on how these tools help to improve learning in the present and the future, and not just results in school. In this regard, there are three aspects that are frequently addressed by the reviews and studies analyzed: the motivation to learn, the pupil's ability to self-regulate or self-organize what is learned and the pupil's perception of self-efficacy [10] [15]. We assert that the literature only finds a positive impact on formative assessment in the first two.

- **Motivation to learn**: attitudes towards learning improve with formative assessments. The confidence that pupils gain in their ability to evaluate leads to a rise in their motivation to learn. It also must be pointed out that the motivation to learn increases among the youngest groups through m-learning mechanisms, certainly because they are more familiar with new technologies, while variations are not observed in older pupils [13].

Attitudes towards learning improve with formative assessments. The confidence that pupils gain in their ability to evaluate leads to a rise in their motivation to learn.

↗

- **Ability to self-organize learning**: the self-regulation of learning is not a simple task and requires many skills, from the ability to plan and set goals to the ability to reflect and implement the right strategies for learning.[2] Traditionally evaluated in higher levels of education (secondary and university education), pupils' ability to self-manage and self-organize is increasingly becoming an objective of primary education as well. The search for strategies that help pupils to organize their work during the school year has become part of the work of primary education teachers and formative assessment is being established as a tool among these strategies. Even though there is no set age to develop the ability of self-management, research shows that, with methodologies adapted to early stages, self-regulation strategies for learning during primary school have positive effects on subsequent levels of education. Formative assessment seems to help to develop these abilities, especially when this perspective is included in the classroom activities as a whole [12]. The value of formative assessment lies in the information it provides on the gap between the pupils' real performance and expected performance, but also and especially on the tools that it offers for improving performance, while indicating the best strategies to follow. This is where self-regulation is found to improve, since pupils receive a wide array of options to incorporate and develop in the future that help them to take responsibility for their

The search for strategies that help pupils to organize their work during the school year has become part of the work of primary education teachers and formative assessment is being established as a tool among these strategies.

↗

2 For more details about the concept of self-regulation and about the effectiveness of educational efforts that try to encourage it, please see the article by Gerard Ferrer-Esteban in issue number 5 of the series *What Works in Education*: https://goo.gl/4Xvz5T.

**What Works
in Education?**

FUNDACIÓ
JAUME
BOFILL

ivàlua | Institut Català d'Avaluació
de Polítiques Públiques

14

education [14]. The impact on self-regulation is greater when self-assessment is the evaluative mechanism, even though peer assessment also boosts pupils' capacity for self-regulation [12].

- **Perception of self-efficacy**: the pupil's perception of his or her ability to carry out tasks and/or achieve the goals established is an important part of the learning process. Even though it is an aspect very close to and often related to the capacity for self-regulation, our research does not observe significant effects between the method of evaluation and the pupil's perception of his or her abilities [12].

Summary

Though there are many approaches to assessment, this review has only focused on those with an explicit formative intent, although at some point the studies used purpose-driven techniques rather than formative ones in the control group. Our review of the evidence has shown that formative feedback has a significant impact on academic performance, as well as on broader educational issues like the capacity for self-organization and the motivation to learn. However, pupils' perceptions about their own abilities do not seem to be sensitive to the methods of evaluation.

The breadth and variety of assessment tools make it difficult to obtain generalized results, but help to identify some practices with major impact at the same time. Simple verification (correct/incorrect) and dynamics of reward and punishment are seen as tools with little impact on pupils' performance. However, combining correction with facilitating feedback, self-assessment by the pupils themselves or peer assessment are considered an interesting option for improving performance.

Combining correction with facilitating feedback, self-assessment by the pupils themselves or peer assessment are considered an interesting option for improving performance.

↗

Moreover, it should be noted that this is a sphere with clearly differential impacts according to the pupils' profile. Thus, while low-performing pupils benefit from more complex and immediate formative feedback systems, more advanced pupils get better results with feedback that is simpler and delayed. Both profiles also differ in their sensitivity to how the feedback is oriented. As a whole, the most vulnerable pupils, whether due to their skills or social class, seem to improve their performance in school more clearly when formative assessment tools are used and especially when that formative assessment is adapted to their characteristics.

Table 4.
Arguments for and limitations of formative assessment tools

Points in favor	Limitations
• Formative assessments have greater effects on performance than purpose-driven ones.	• Assessments must be formative and not be limited to evaluating the learning process in order to have effects on performance.
• Facilitating assessments contribute more to learning than directive ones.	• Assessments must avoid comparisons with other pupils and focus on each pupil's potential.
• A clear connection must be established between the evaluation criteria and the objectives of the teaching units in order to turn the assessment into a formative tool.	• Giving marks reduces the potential formative effect of the assessment.
• The identification of mistakes must include clues so they can get rectified and improved.	• De-motivating feedback centered on mistakes must be avoided and recommendations or support for improvement must be stressed.
• Assessments must be directed at what is learned as a whole.	• Formative assessments may not be limited to evaluating performance in a specific exam.
• Feedback must be organized in units that pupils can understand.	• Assessments may not be complex, since they must be understandable to pupils and provide them with tools to rectify their previous mistakes.
• In order to encourage general learning not limited to the issue being evaluated, delayed feedback has better effects.	• Delayed feedback is not recommended for very complex tasks or for pupils with learning difficulties.

Source: author

Implications for practice

The information presented here demonstrates the important impact of formative assessment on pupils' performance, but it also shows its fragility before many factors that may reduce or emphasize these effects. In other words, the benefits of the evaluation are not guaranteed under any circumstances, but different requirements are needed, some of which depend on policy makers. In this regard, some recommendations are provided below:

• It is important that teachers' assessment efforts are as efficient and effective as possible. To that end, teachers must be trained to use formative assessment tools and support schools in designing their evaluation plans in order to boost pupils' performance.

• Context has proven to be very important for the success of formative assessments. They must be raised in the context of a learning process that is not purpose-driven, meaning one that is oriented towards ongoing education. Much of an assessment's success stems from the identification of clear objectives for the pupils in such a way that together with the criteria for correction, the learning objectives must be clarified so they may make the tasks and assessments understandable.

This is a long process that must begin with planning the activity and continue until evaluation and feedback. Therefore, teachers must be given time to plan the assessment as part of the learning process.

Teachers must be given time to plan the assessment as part of the learning process.

- The impetus of permanent or ongoing assessment must enable the expansion of objectives beyond performance and towards other spheres of learning. In this regard, self-assessment or peer assessment in primary school seems to have important effects on pupils' capacity for self-regulation and may lead to improvement in transitioning to secondary school, where teachers' levels of control decrease and more self-management skills are required. Self-assessment and peer assessment require special planning, resources for development and training for teachers so they can be implemented.

- In addition, given the importance that previous knowledge has on explaining pupils' performance and the certainty that this knowledge shapes future learning, it seems appropriate to provide schools with diagnostic assessment tools in order to fine tune the design of the learning process. In this regard, resources and training are needed to correctly identify previous knowledge. Therefore, it would be desirable for the public administration to use this initial evaluation as a mechanism of control in its analysis of the assessment of teachers and schools.

It seems appropriate to provide schools with diagnostic assessment tools in order to fine tune the design of the learning process.

- The public administration requires data and information to improve the design of educational processes. To this end, assessments become very important. However, they must be complete and allow the greatest number of variables that intervene in the process to be captured (family instructional capital, motivation, teaching quality, etc.). These assessments must also be made available to teachers and pupils alike so they may be used for educational purposes.

- Finally, in light of the ambiguity of some results, research must be promoted that goes into greater depth to assess the impact of different assessment tools, the intermediation of other variables like the quality of teaching staff and, last but not least, the effects on low-performing pupils.

Bibliography

[1]   OCDE (2013). *Synergies for Better Learning.* OECD Publishing.

[2]   Elliot, V.; Baird, J.-A.; Hopfenbeck, T.; Ingram, J.; Thompson, I.; Usher, N.; Zantout, M.; Richardson, J. and Coleman, R. (2016). "A marked improvement? A review of the evidence on written marking".

[3]   Dunn, K. E. and Mulvenon, S. W. (2009). "A Critical Review of Research on Formative Assessment : The Limited Scientific Evidence of the Impact of Formative Assessment in Education". *Pract. Assessment, Res. Eval.*, vol. 14, no. 7, pp. 1–11.

[4]   Bell, B. and Cowie, B. (2000). "The Characteristics of Formative Assessment in Science Education". *Sci. Educ.*, vol. 85, pp. 536–553.

[5]   Shute, V. J. (2007). "Focus on Formative Feedback", no. March.

[6]   Hattie, J. and Timperley, H. (2007). "The Power of Feedback". *Rev. Educ. Res.*, vol. 77, no. 1, pp. 81–112.

[7]   Klute, M.; Apthorp, H.; Harlacher, J. and Reale, M. (2017). "Formative assessment and elementary school student academic achievement : A review of the evidence".

[8]   Ross, J. A. (2006). "The Reliability, Validity, and Utility of Self-Assessment". *Pract. Assessment, Res. Eval.*, vol. 11, no. 10, pp. 1–11.

[9]   Dochy, F.; Segers, Mi. and Buehl, M. M. (1999). "The Relation Between Assessment Practices and Outcomes of Studies : The Case of Research on Prior Knowledge". *Rev.*, vol. 69, no. 2, pp. 145–186.

[10] Black, P. and Wiliam, D. (1998). "Assessment and Classroom Learning". *Assess. Educ. Princ. Policy Pract.*, vol. 5, no. 1, pp. 7–74.

[11] Kluger, A. N. and DeNisi, A. (1996). "The Effects of Feedback Interventions on Performance : A Historical Review , a Meta- Analysis , and a Preliminary Feedback Intervention Theory". *Psychol. Bull.*, vol. 119, no. 2, pp. 254–284.

[12] Meusen-Beekman, K. D.; Brinke, D. J. and Boshuizen, H. P. A. (2016). "Studies in Educational Evaluation Effects of formative assessments to develop self-regulation among sixth grade students : Results from a randomized controlled intervention". *Stud. Educ. Eval.*, vol. 51, pp. 126–136.

[13] De Marcos, L.; Hilera, J. R.; Barchino, R.; Jiménez, L.; Martínez, J. J.; Gutiérrez, J. A.; Gutiérrez, J. M. and Otón, S. (2010). "An experiment for improving students performance in secondary and tertiary education by means of m-learning auto-assessment". *Comput. Educ.*, vol. 55, no. 3, pp. 1069–1079.

[14] Wiliam, D.; Lee, C.; Harrison, C. and Black, P. (2004). "Teachers developing assessment for learning : impact on student achievement". *Assess. Educ. Princ. Policy Pract.*, vol. 11, no. 1, pp. 49–65.

[15] Mcdonald, B. and Boud, D. (2003). "The Impact of Self-assessment on Achievement : The effects of self-assessment training on performance in external examinations". *Assess. Educ. Princ. Policy Pract.*, vol. 10, no. 2, pp. 209–220.

[16] Schunk, D. H. (1996). "Goal and Self-Evaluative Influences During Children's Cognitive Skill Learning". *Am. Educ. Res. J.*, vol. 33, no. 2, pp. 359–382.

[17] Education Endowment Foundation (2017), "Teaching and Learning Toolkit: Feedback".