



Academic Results and Educational Service Provision: An Analysis of External Evaluation Reports in Portuguese Schools

Resultados Académicos e Prestação do Serviço Educativo: Uma Análise dos Relatórios de Avaliação Externa em Escolas Portuguesas

Resultados Académicos y Provisión de Servicios Educativos: Un Análisis de Informes de Evaluación Externa en Escuelas Portuguesas

Diana Pereira
Universidade do Minho
dianapereira@ie.uminho.pt

Abstract

The promotion of educational success and particularly the assessment process are key aspects in the current educational research. This study is part of a larger research project “Assessment in Basic and Secondary Education: teachers’ practices and conceptions of internal and external assessment in Portuguese schools”. The main goal of this study is to analyse the external evaluation reports of 16 schools in order to examine the “results” domain concerning academic success; to examine the “provision of educational services” domain regarding the monitoring of assessment; and to relate the assessment obtained in each domain with the operationalisation of the assessment process in each context. The schools were selected based on the results presented in the 2019 ranking. Therefore, 8 basic schools were chosen (4 with the highest rankings and 4 with the lowest rankings), as well as 8 secondary schools (4 with the highest rankings and 4 with the lowest rankings). The main findings show that when the evaluators rate the results domain as “very good” or “good” are presented the reduction of retention rates, results above average, completion rate above the national average, positive internal and external assessment results and no or residual school dropout. Regarding the common categories evaluated as good and very good in the field of educational service provision are mainly the supervision and analysis of student’s results, the analysis and comparison of internal and external ratings for reflection on results, the establishment of common assessment criteria and public assessment system for all stakeholders. The common categories concerning the operationalisation of assessment are the use of different instruments to assess and the use of formative assessment as the main modality. This study provides implications for educators, policymakers, and educational stakeholders, emphasising the need for a comprehensive and reflective approach to assessment practices to enhance overall educational outcomes.

Keywords: Evaluation; External Assessment; General-Inspectorate of Education and Science.



Resumo

A promoção do sucesso educativo e, em particular, o processo de avaliação, são aspetos-chave na investigação educacional atual. Este estudo faz parte de um projeto de investigação mais amplo intitulado “Avaliação no Ensino Básico e Secundário: práticas e concepções dos professores sobre avaliação interna e externa nas escolas portuguesas”. O principal objetivo deste estudo é analisar os relatórios de avaliação externa de 16 escolas com o propósito de examinar o domínio “resultados” relacionado com o sucesso académico; analisar o domínio “prestação de serviços educativos” no que diz respeito à monitorização da avaliação; e relacionar a avaliação obtida em cada domínio com a operacionalização do processo de avaliação em cada contexto. As escolas foram selecionadas com base nos resultados apresentados no ranking de 2019. Assim, selecionaram-se 8 escolas básicas (4 com as classificações mais altas e 4 com as classificações mais baixas), bem como 8 escolas secundárias (4 com as classificações mais altas e 4 com as classificações mais baixas). As principais conclusões indicam que, quando os avaliadores classificam o domínio de resultados como “muito bom” ou “bom”, são apresentados indicadores como a redução das taxas de retenção, resultados acima da média, taxas de conclusão acima da média nacional, resultados positivos na avaliação interna e externa e ausência ou redução significativa do abandono escolar. No que diz respeito às categorias comuns avaliadas como boas e muito boas no campo da prestação de serviços educativos, destacam-se principalmente a supervisão e análise dos resultados dos alunos, a análise e comparação das classificações internas e externas para reflexão sobre os resultados, o estabelecimento de critérios de avaliação comuns e um sistema de avaliação público para todos os intervenientes. As categorias comuns relacionadas com a operacionalização da avaliação incluem a utilização de diferentes instrumentos de avaliação e a utilização da avaliação formativa como a principal modalidade utilizada (incluindo a avaliação diagnóstica). Este estudo sugere implicações para educadores, decisores políticos e intervenientes na área educacional, sublinhando a necessidade de uma abordagem abrangente e reflexiva às práticas de avaliação para melhorar os resultados educacionais globais.

Palavras-chave: Avaliação; Avaliação externa; Inspeção-Geral da Educação e Ciência.

Resumen

La promoción del éxito educativo y, en particular, el proceso de evaluación, son aspectos clave en la investigación educativa actual. Este estudio forma parte de un proyecto de investigación más amplio titulado “Evaluación en la Educación Básica y Secundaria: prácticas y concepciones de los profesores sobre la evaluación interna y externa en las escuelas portuguesas”. El objetivo principal de este estudio es analizar los informes de evaluación externa de 16 escuelas con el propósito de examinar el dominio de “resultados” relacionado con el éxito académico; analizar el dominio de “prestación de servicios educativos” en lo que respecta al seguimiento de la evaluación; y relacionar la evaluación obtenida en cada dominio con la operacionalización del proceso de evaluación en cada contexto. Las escuelas fueron seleccionadas según los resultados presentados en el ranking de 2019. Por lo tanto, se eligieron 8 escuelas básicas (4 con las clasificaciones más altas y 4 con las clasificaciones más bajas), así como 8 escuelas secundarias (4 con las clasificaciones más altas y 4 con las clasificaciones más bajas). Las principales conclusiones indican que cuando los evaluadores califican el dominio de resultados como “muy bueno” o “bueno”, se



presentan indicadores como la reducción de las tasas de retención, resultados por encima del promedio, tasas de finalización por encima del promedio nacional, resultados positivos en la evaluación interna y externa y la ausencia o reducción significativa de la deserción escolar. En lo que respecta a las categorías comunes calificadas como buenas y muy buenas en el campo de la prestación de servicios educativos, destacan principalmente la supervisión y el análisis de los resultados de los estudiantes, el análisis y la comparación de las calificaciones internas y externas para reflexionar sobre los resultados, el establecimiento de criterios de evaluación comunes y un sistema de evaluación público para todos los participantes. Las categorías comunes relacionadas con la operacionalización de la evaluación incluyen el uso de diferentes instrumentos de evaluación y la utilización de la evaluación formativa como la principal modalidad utilizada. Este estudio sugiere implicaciones para educadores, formuladores de políticas y partes interesadas en la educación, destacando la necesidad de un enfoque integral y reflexivo de las prácticas de evaluación para mejorar los resultados educativos generales.

Palabras clave: Evaluación; Evaluación Externa; Inspección-General de Educación y Ciencia.

Introduction

The dynamics of change that accompany educational systems, largely driven by inadequate and effective responses to societal demands, have established assessment as a field of interest. One of the most trends in research within this domain is the shift toward a more holistic and authentic approach. Assessment plays a predominant role in shaping the teaching strategies employed by educators, as well as in how students perceive, engage with, and allocate their time to pedagogical activities (Flores & Pereira, 2019; Biggs, 2003). Furthermore, it exerts influence over all stakeholders involved in the educational process (Snyder, 1971; Scouller, 1998). However, Pacheco and Maia (2019) argue that the school builds quality through assessment processes linked to external benchmarks, notably institutional assessment with a focus on academic outcomes. The assessment of learning is shaped by internal and external tests as well as the comparative method. In fact, and according to Flores et al. (2017), “assessment policies have been replaced by assessment as the primary policy in the field of education. More than a peripheral aspect of the functioning and management of educational systems and teaching practices, assessment currently assumes centrality in the educational field” (p.7). This centrality is driven by the effects of globalization and, in the words of Machado (2014), by the establishment of an *omniassessment* system. It becomes essential to understand what assessment systems tell us regarding curriculum development, particularly on the assessment process. Thus, this study is part of a larger research project “Assessment in Basic and Secondary Education: teachers’ practices and conceptions of internal and external assessment in Portuguese schools” funded by the Portuguese Foundation for Science and Technology (CEECIND/03157/2018) that has a strong empirical component in order to get to know the teachers’ conceptions and practices of assessment in Basic and Secondary Education and to understand to what extent the external and internal assessment



influences the quality of learning and how they relate to each other. In particular, the present study presents a critical analysis of external evaluation reports from 16 school clusters/schools regarding two major analytical axis: student academic outcomes and the provision of educational services, particularly with regard to the monitoring of assessment. The aim is to analyse the results assigned by the evaluators to these two axes, allowing the identification of the assessment practices valued and the areas for improvement in this field.

Internal and External Assessment

The promotion of educational success and particularly the assessment process are key aspects in the current educational research. However, research highlights different logics of promoting success through external evaluation, which allows for the monitoring and control of students' and schools' academic performance (Flores et al., 2017). In this sense, there is a materialization of *accountability* through the emphasis on results and standards, which narrow the curriculum and define a managerial profile for the teacher's role (Pacheco, 2019). Such a situation places, in many contexts beyond just the Portuguese context, an emphasis on results and rankings as the primary indicator of measuring the success of educational systems, leading to a narrowing of curriculum practices and standardized approaches (Marcondes, 2017; Flores, 2017; Flores et al., 2017). In Portugal, the external evaluation is a legally mandated activity, carried out by a government agency IGEC (General-Inspectorate of Education and Science) and supplemented by the involvement of external experts. The aim of IGEC is to ensure quality, equity and justice in education (nursery and pre-school to higher education and university) through the monitorization, control and evaluation of schools in order to ensure social trust and inform political decision-makers and public opinion. In this context, some recent studies in this field indicate that external evaluation has a more substantial impact on organizational development than on curricular and pedagogical development (Pacheco et al., 2023). Moreover, the analysis of 60 external evaluation reports conducted by Serra et al. (2023) proposes that the assessment of the results domain should encompass an integrated approach to various facets of educational action, extending beyond mere academic outcomes. In the Portuguese context what students learn is disseminated by rankings made on the basis of external assessments (standardized tests, national examinations and final exams) that allow comparisons between schools or regions by country. The curriculum of basic and secondary education assumes that schools are given autonomy for curriculum development in order to reinforce the dynamics of assessment by focusing on the diversity of tools enabling greater awareness of the effectiveness of the work carried out and monitoring at the first sign of difficulty in pupils' learning. The assessment process in these levels of schooling is divided into i) internal assessment that comprises the formative assessment; the summative assessment and the mobilization of diversified and adequate techniques, instruments and procedures and ii) the external assessment that establishes, depending on the nature of each educational and training offer, the performance of standardized tests; final tests of basic education and national examinations



(Decree-law n.55/2018). The recent introduction of recommendations, structural documents, and educational policies such as Decree-Law No. 54/2018, Decree-Law No. 55/2018, Student's Profile by the End of Compulsory Schooling (PASEO), Essential Learnings, and the National Education Strategy for Citizenship suggest a shift in the reorganization of the curriculum for basic and secondary education. With these new curriculum reorganization strategies, it is expected that the assessment process will become more flexible and innovative compared to traditional assessment practices.

In fact, literature points to more sophisticated conceptions of assessment, founded on an assessment for learning paradigm, with the aim of promoting active student engagement, motivation, effective feedback (Black & Wiliam, 1998; Wiliam, 2010; Broadfoot et al., 2002), and the interdependence of the teaching, learning, and assessment process (Black, 2015). Similarly, other conceptions of assessment emerge, emphasizing assessment as a developmental process that supports students' metacognition, with a focus on their participation in self-assessment, thus directing their own learning (Earl, 2013; Lam, 2016). In this way, there is alignment between what research proclaims and what is established regarding the assessment in the Portuguese context, namely, the centrality of formative assessment. However, a perpetuation of assessment grounded in the principles of summative assessment persists, where the assessment's purpose is to classify what students have learned for the purpose of certifying their knowledge and comparing their performance (Earl & Katz, 2006; Perrenoud, 1999). This is accompanied by an enduring emphasis on students' academic outcomes through national tests and examinations in various countries (Flores et al., 2017). Also, an extensive body of research on assessment suggests that teachers' conceptions of assessment influence their assessment practices and how they relate these practices to the teaching and learning process (Fletcher et al., 2012; Samuelowicz & Bain, 2002; Brown & Hirschfeld, 2008; Pereira, 2016) and these conceptions may contribute to the perpetuation of summative assessment practices. Studies on this topic found that teachers' conceptions of assessment can be understood in terms of improvement of teaching and learning; school accountability; student accountability, or treating assessment as irrelevant (Brown, 2004) as well as teacher's conceptions reflect the differing practices of assessment by level of schooling (Brown et al., 2011).

Method

The main goal of this study is to analyse the external evaluation reports of 16 schools in order to:

- examine the "results" domain concerning academic success;
- examine the "provision of educational services" domain regarding the monitoring of assessment;
- relate the assessment obtained in each domain with the operationalisation of the assessment process in each context.



As a first step, the 16 school clusters/schools that are participating in the broader project “Assessment in Basic and Secondary Education: practices and conceptions of teachers about internal and external assessment in Portuguese schools” were selected. Among these 16 schools, 4 are ranked among the 7 best public secondary schools in Portugal (School A, B, C and D) and 4 are ranked among the 7 public secondary schools with the lowest academic results (School E, F, G and H). Furthermore, the other 4 schools are classified among the 7 best basic public schools in Portugal (School I, J, K and L), while the remaining 4 are classified as basic public schools with lower academic results (School M, N, O and P). The ranking analysed is from 2019, which is the ranking used in the broader project.

The website of the General-Inspectorate of Education and Science was accessed to analyse the external evaluation reports of the 16 selected schools. Since the external evaluation of schools has already been conducted in two assessment cycles (1st 2006-2011; 2nd 2011-2017), at the moment a third assessment cycle is underway. Thus, the reports from the 1st and 2nd assessment cycles were analysed in order to examine the reports of the selected schools. The reports from the 1st assessment cycle are structured into 5 evaluation domains, namely: 1) Results; 2) Provision of Educational Services; 3) Organization and School management; 4) Leadership; and 5) School Self-regulation Capacity and Improvement. In this first cycle, the evaluation scale included Very Good, Good, Satisfactory, and Insufficient. The reports of the 2nd assessment cycle are structured only into 3 evaluation domains, namely: 1) Results; 2) Provision of Educational Services; and 3) Leadership and Management. In the second cycle the Excellent was added to the scale of evaluation.

In this study, only the “results” dimension and the “provision of educational services” dimension were analysed since it aligns with the study’s objectives, particularly because it allows to know and understand the operationalisation of the assessment process and how it is evaluated in Portuguese schools.

Content analysis was used to identify the emerging categories in order to describe, in a systematic and articulated way, the meaning of the data collected (Schreier, 2012).

Regarding the ethical research issues, the anonymity of the institutions involved will be ensured and they will be identified as School Cluster A, School Cluster B, etc.

Results and discussion

As can be seen in table 1, regarding the results domain, in terms of academic success, it is possible to argue that when the evaluators rate this domain as “very good” or “good,” it generally exhibits the following characteristics: reduction of retention rates, results above average, completion rate above the national average, positive internal and external assessment results and no or residual school dropout (early school leaving). In this regard, what distinguishes “good” from “very good” in this domain is the significantly high rates of academic success and the absence of school dropout. As expected, all the schools ranked among the top 4 out of the 7 best secondary and basic schools in Portugal have achieved a



rating of “good” or “very good” in the external evaluation. On the other hand, 4 of the 7 secondary schools ranked with the lowest averages have achieved a rating of “sufficient” in the external evaluation. Surprisingly, among the 4 out of 7 lower-rated basic schools in rankings, received 1 “very good, 2 “goods” rating and 1 “sufficient.” This variation in evaluators rate is not aligned with the expected outcomes relative to their ranking positions. However, it’s worth noting that this result may be explained by temporal differences between the release of external evaluation reports and the publication of the rankings. Another finding that emerges from this study is the fact that the majority of the top-ranking national secondary schools (3 out of the 4 presented) are non-clustered schools. On the other hand, the 4 secondary schools with the lowest outcomes in ranking all belong to school clusters. The trend persists with regard to basic schools. 3 out of the 4 basic schools presented with the best results in the ranking are non-cluster schools. Similarly, the 4 basic schools with the lowest rankings belong to school clusters.

Regarding the provision of educational services concerning the monitoring of assessment, it is possible to conclude that only 2 schools received a rating of “sufficient” (one basic school and one secondary school with lowest academic results in ranking). 10 of the 16 schools were rated good and the remaining (4) were rated very good. It should be noted that one of the basic schools with the lowest results in the ranking was evaluated by inspectors as very good both in terms of service provision and in terms of results, which was not expected. However, after analyzing the report, it is possible to argue that this school presents low classifications in the external evaluation exams for basic education, but in secondary education it presents very good results.

The common categories regarding schools evaluated as good and very good in the field of educational service provision are mainly the supervision and analysis of student’s results (Pedagogical Council, Curriculum Departments and School Assembly), analysis and comparison of internal and external ratings for reflection on results, statistical data on student results, establishment of common assessment criteria and public assessment system for all stakeholders. The common categories concerning the operationalisation of assessment are: the use of different instruments to assess; the use of formative assessment as the main modality (including diagnostic assessment); the students’ self-assessment; the application of the same assessment instruments to diverse classes; the collaboration in development of tests or the application of the same test to different classes contributes to reinforcing confidence; proposals for support measures; joint matrix and test preparation and intermediate test implementation (GAVE); sharing assessment materials by teachers’ collaborative work; the adjustment of plans and mobilization of educational strategies and support measures; success goals by discipline, year and cycle; consistent collaborative work by teachers, in motivating students and in good management and pedagogical organization, particularly in the educational support provided; teachers application of the criteria regularly using common PASEO references; and the transparency in the assessment process. However, some aspects stand out as less positive that were attributed even to schools that obtained the classification of “good” and “very good” in this domain, such as no evidence of collaborative work; lack of



plan for identified learning difficulties; no culture of internal monitoring and supervision of teaching practices; lack of establishment of success goals; no systematic and common policies that establish coherence between teaching and assessment; no application of common classification matrices; analysis of results without a prospective dimension; and the lack of coherent and cohesive practices.

In this regard, an analysis of the school evaluation model carried out by Pacheco (2010) reveals that the *Results* dimension is privileged over others, although there is no application of factorial weight in the evaluation model for each component. For instance, the *Results* dimension, being one of our axes of analysis, particularly the academic success, appears to be central. The *Educational Service Provision*, the second axe of our analysis, roughly encompasses curriculum-related issues, particularly in this case (Pacheco, 2010) issues that are directly related to the operationalisation of the assessment process.

Overall, the findings point to an appreciation of the academic results achieved by students in external assessments. Schools that which have a high average in terms of academic results typically receive favorable evaluations as well as encompass completion rates and the absence of school dropouts. However, in addition to recognising the complexity and scope of assessments, it is crucial to take into consideration the various facets of academic success (Alves & Machado, 2008; Flores, 2017; Cabral & Matias, 2017) when analyzing data of this nature. Certain constraints, such as the schools' context variables or, for example, whether it presents itself as a TEIP school or not, certainly contribute to more unfavorable results. Findings also show that in schools high rated by inspectors, teachers use different assessment instruments, based on the assumptions of formative assessment where there is room for students to participate in the assessment process, either through self or peer assessment. Also, the dynamics of critical analysis of results obtained both through internal and external evaluation are valued. These dynamics are in line with what has been advocated in the literature and also with the recent assumptions of curricular autonomy and flexibility (Decree-law no. 55/2018) in the Portuguese curricular organisation of basic and secondary education.

However, the way educational leadership perceives the curriculum, the teaching and learning process, assessment, management, and school culture is crucial to implement this new curricular organisation. The study by Pereira et al. (2021) shows that most of school principals presents divergent discourses regarding of curricular autonomy and flexibility. Only a minority of principals reveals an adherence discourse. In addition to these aspects, it also requires the involvement of the entire educational community towards the same objective, particularly the commitment of leadership that advocates for pedagogical innovation, curricular coordination, the development of competencies suitable for current demands (Pereira et al., 2021) and a strong sense of collaborative work (Niklasson, 2017), contradicting the school organisation model based on individual work. From these results we can observe that the most formative practices are widely used by teachers at schools highly evaluated in the rankings. However, studies show that training practices sometimes focus too much on teachers, which implies a greater volume of work and may condition students' accountability for their learning (Fulmer et al., 2019).



Table 1. Analysis of external evaluation reports regarding results and provision of educational services.

School A Among the top public secondary schools in Portugal (ranking) Centre of Portugal, Urban			
Results Academic Success		Provision of Educational Services Monitoring of Assessment	
GOOD		GOOD	
+	Reduction of retention rates	+	Curriculum departments ensure the supervision and analysis of school results
-	Results fall short of the expected value (although they have very satisfactory success rates when compared at the national and regional levels)	+	Establishment of common assessment criteria and public assessment system for all stakeholders
		+	Application of the same assessment tool to diverse classes; calibration of tests and grading consistent with teaching practices
		+	Analysis of student's results (Pedagogical Council, Curriculum Departments and School Assembly) Analysis and comparison of internal and external ratings for reflection on results; statistical data on student results. Strategies and improvement plans
		+	Teachers consulted about training needs in this domain
		-	There is no evidence of collaboration with class councils to address and plan for identified learning difficulties; no culture of internal monitoring and supervision of teaching practices.
		-	Setbacks in the analysis of students' continuous assessment results, which allows evaluation to be designed as a formative and learning-regulating process
School B Among the top public secondary schools in Portugal (ranking) Centre of Portugal, Urban			
Results Academic Success		Provision of Educational Services Monitoring of Assessment	
GOOD		GOOD	
+	Personal and collective successes publicized by the school	+	Definitions of assessment criteria in all subjects; analysis of students' results and various success indicators by subject, class and year
+	Completion rate above the national average	+	Collaboration in development of tests or the application of the same test to different classes contributes to reinforcing confidence in the results obtained



+	Positive national exam results in all subjects (slightly below internal assessment)	+	The supervision and analysis of school results. The comparison of results with external referents (sociologically similar or national schools) or the analysis of the temporal evolution of the classifications of national exams are not part of the practices of internal reflection
+	No school dropout	+	Correction of tests by different teachers and comparison of internal assessment with the results obtained in standardised tests
-	Not have an honor roll	+	Proposals for support measures
-	Some enrollment cancellations (strategy to obtain better results in national exams)	-	Success goals are not set and interdisciplinary coordination takes place within the scope of informal relationships without assuming expression at the planning level
School C			
Among the top public secondary schools in Portugal (ranking)			
Centre of Portugal, Urban			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
VERY GOOD		VERY GOOD	
+	Positive internal and external assessment results	+	The results collected and processed statistically are analyzed systematically within the agency
+	9th year completion rate above average	+	Two-sided assessment method
+	Residual school dropout	+	Definition of general and specific standards, including practices and instruments: theories, oral, written, attitudes and values, etc.
+	Systematic dropout decrease	+	Comparative analysis of internal and external results.
		+	Joint Matrix and Test Preparation and Intermediate Test Implementation (GAVE)
		+	Cross-correction of tests
		+	Sharing assessment materials through collaboration
		+	Students perform self-assessment
School Cluster D			
Among the top public secondary schools in Portugal (ranking)			
Centre of Portugal, Urban			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
VERY GOOD		VERY GOOD	
+	Above average results	+	Known and understood assessment criteria
+	Completion rates above and well above average	+	Use of different assessment modalities
+	Nonexistent school dropout	+	Activity sheets that involve students in the systematic construction of learning



		+	Analysis of student results and forms help prevent situations of failure and demotivation
		+	Tests created in collaboration by teachers of the same subject even if they are not applied equally
		+	Results are assessed collectively
		+	Monitoring and evaluation of results by the administration, direction and management of educational coordination and pedagogical supervision structures
		+	Comparisons of results with classes of the same year and national values, with improvement plans being redefined and implemented
		+	Identification of areas of success and failure and formulation of improvement actions
School Cluster E			
Among the public secondary schools with lowest academic results in Portugal (ranking)			
South of Portugal, Urban			
Results Academic Success		Provision of Educational Services Monitoring of Assessment	
SUFFICIENT		GOOD	
+	Low abandonment rates	+	Diagnostic assessment used at all educational levels
-	National exams in the 9th year below average	-	There are no systematic and common policies that establish coherence between teaching and training practices
-	Mathematics always below average	-	Only some subjects use common assessment tools according to educational level
-	External assessment tests below the national average (except for the 3rd cycle)	-	Calibration of assessment instruments is only carried out by some departments
School Cluster F			
Among the public secondary schools with lowest academic results in Portugal (ranking)			
South of Portugal, Suburban			
Results Academic Success		Provision of Educational Services Monitoring of Assessment	
SUFFICIENT		SUFFICIENT	
+	Reduction in school dropout rates	+	Diversified practices and instruments (portfolios, oral presentations, reports) depending on the specificity of the area
+	Results in line and below with expected values	+	Intermediate tests
-	Completion rates below values	+	Predominance of formative assessment
		+	Use of different assessment modalities
		+	Analysis of results by structures. Systematic collection of data relating to this domain



	-	The assessment of the effectiveness of the measures adopted does not reflect a joint reflection leading to the diversification of practices with an impact on academic results	
	-	It is not practical to develop and apply common classification matrices	
School Cluster G			
Among the public secondary schools with lowest academic results in Portugal (ranking)			
North of Portugal, Rural			
Results	Provision of Educational Services		
Academic Success	Monitoring of Assessment		
SUFFICIENT		GOOD	
+	Completion rates in the 3 cycle above expected values and Portuguese exam results	+	Define general evaluation criteria and the Curricular Department defines the specific criteria
-	Secondary rates and results below expected value	+	Criteria known by teachers, students and parents
		+	Formative dimension of assessment (included diagnostic assessment)
		+	Adjustment of plans and mobilisation of educational strategies and support measures
		+	Diversified assessment instruments adjusted to levels, subjects and skills
		+	Development of common assessment matrices
		-	Self-evaluation practices at the end of each period but are not seen as self-regulation of the teaching and learning process
School Cluster H			
Among the public secondary schools with lowest academic results in Portugal (ranking)			
South of Portugal, Suburban			
TEIP School			
Results	Provision of Educational Services		
Academic Success	Monitoring of Assessment		
SUFFICIENT		GOOD	
+	1st cycle above expected values and completion rate	+	Investments in improving results are a priority outlined in the strategic planning document
-	Academic results fall short of expected values	+	Strengthening vertical curriculum management and peer intervention in the classroom context
-	Context with unfavorable variables	+	Implementation of various measures to promote success
		+	Analysis and reflection (educational coordination and supervision) of the evolution of learning and results with the aim of ensuring the improvement of educational success (detailed statistical information on external and internal performance in the different curricular areas)



		+	Approval of criteria, application of common assessment instruments and construction of equal tests in some years (greater equity and transparency in the assessment process)
		+	Shared correction
		+	Various types of assessment, diagnostic assessment (data reflected in the respective coordination and pedagogical supervision structures)
		+	Formative assessment appears important to systematize and generalize self-regulation procedures in order to constitute a strategy for reflection and readjustment of the teaching and learning process
		+	Success goals by discipline, year and cycle evaluated quarterly
		+	Initiatives to combat and prevent school dropout
		+	Creation of professional courses guarantees the completion of compulsory education
School I			
Among the top public basic schools in Portugal (ranking)			
South of Portugal, Urban			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
VERY GOOD		GOOD	
+	Increase in the success rate	+	Public presentations and quarterly hearings encourage monitoring of the work developed and allow students to self-evaluate
+	The good academic results achieved demonstrate the quality of the teaching provided	+	Carrying out masterclasses and the results obtained provide feedback on the quality of the work being developed, contributing to strengthening skills
		+	Pedagogical differentiation strategies contribute to the results achieved
		+	Comparison of results with national averages
		+	Analysis of results
		+	Diagnostic assessment to identify learning difficulties
		+	Assessment criteria by subject, year, cycle
		+	Uniform matrices for tests
School Cluster J			
Among the top public basic schools in Portugal (ranking)			
North of Portugal, Urban			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
GOOD		GOOD	
+	Results above the national average	+	Use of different assessment modalities and instruments



+	3rd cycle completion rates 100%	+	Assessment as an integral and regulating element of educational practice
+	Good results obtained in internal classifications of the 9th year	+	The evaluation criteria consider the specificity of the subjects
-	To improve success rates of professional courses	+	Formative assessment as the main modality
		+	Self-assessment that precedes the summative assessment
		+	Analysis of results by structures
		-	Diagnostic assessment is only used in the 7th and 10th year
		-	The results obtained by applying the diagnostic assessment do not have a prospective dimension
School K			
Among the top public basic schools in Portugal (ranking)			
North of Portugal, Urban			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
GOOD		GOOD	
+	Above average results	+	Systematic analysis of school results
+	Completion rate above national benchmark	+	Demand widely accepted and shared by students, teachers and guardians regarding school results
+	Decrease in dropout rate to residual levels	+	Strictly implemented evaluation criteria covering different evaluation modalities
		+	Teaching collaboration practices in the preparation and sharing of summative tests
		+	Teachers try to adapt planning and evaluation in Class Curricular Plans
		+	Agile processes regarding the identification of necessary support
School L			
Among the top public basic schools in Portugal (ranking)			
Centre of Portugal, Suburban			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
VERY GOOD		VERY GOOD	
+	Completion rates above expected value	+	Consistent collaborative work by teachers, in motivating students and in good management and pedagogical organisation, particularly in the educational support provided
+	Trend of good results	+	Analysis of student results that leads to planning and readjustment of teaching activities and identification of elements of success
+	No dropout rates	+	Learning improvement strategies



		+	Different assessment practices and instruments (tests, research and group work, oral participation and reports of practical and experimental activities)
		+	Defined evaluation criteria by subject, teaching and course and are known by parents and students
		+	Objective goals per year/subject and integrated into the educational project
		+	Differentiated educational offer - extracurricular projects
		+	Consistent impact above expected values in improving student learning and results
		+	Widespread and effective organizational practices
School Cluster M			
Among the public basic schools with lowest academic results in Portugal (ranking)			
South of Portugal, rural			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
VERY GOOD		VERY GOOD	
+	Academic success has been decreasing	+	The student evaluation process is the subject of permanent reflection in educational coordination bodies and structures. The implementation of different strategies (more class project; more academic success program; mathematics plan; national reading plan; accompanied study and systematic formative assessment) is not reflected in the results achieved
-	Below average classifications in national basic education exams and assessment tests	+	Assessment criteria debated, approved and published, by discipline
		+	Teachers apply the criteria regularly using common PASEO references
		+	Application of intermediate tests (GAVE) to ensure confidence in internal evaluation and identify difficulties
		+	The modalities of diagnostic, summative and formative assessment are combined in order to assume an effective function of regulating learning
		+	Students recognise transparency in the assessment process (adoption of relevant procedures in the various disciplines)
		+	Use of various formative assessment tools
		+	Use of self and peer assessment
		+	Insertion in the tests of question quotations and a space reserved for report difficulties in solving them



		+	Frequent collaborative work in terms of designing assessment instruments
School Cluster N			
Among the public basic schools with lowest academic results in Portugal (ranking)			
South of Portugal, Urban			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
GOOD		GOOD	
+	Eradication of school dropouts	+	Systematic procedures for analysing and monitoring academic results involving different structures of educational coordination and pedagogical supervision
+	Positive evolution	+	Reflection on practices and use of mechanisms to improve results (tutoring, support, review of evaluation criteria)
		+	Definition of evaluation criteria
		+	Development of different assessment instruments
		+	Diagnostic assessment as a regular practice that is used to construct curricular and group/class projects and to adapt planning at different levels and cycles
		-	It is not common practice to construct common matrices for tests nor to jointly correct tests, which compromises the measurement of criteria
School Cluster O			
Among the public basic schools with lowest academic results in Portugal (ranking)			
South of Portugal, suburban			
Results		Provision of Educational Services	
Academic Success		Monitoring of Assessment	
GOOD		GOOD	
+	Constant success in line with expected value	+	Systematic analysis of student results
+	Decrease in school dropout rates	+	Diagnostic assessment at all educational levels
-	Given the economic context, more favorable results were expected	+	Diversified recording and observation instruments
		+	Self-assessment of students from a formative perspective
		+	Some tests and matrices are prepared together
		+	The correction criteria have been checked based on what is used in the external assessment tests
		-	In the 2nd and 3rd cycle, the analysis of evaluation practices focuses on results and not so much on sharing and collaborative work
		-	The determining factors of success and failure essential to the teaching and learning process are not identified, with consequences for the implementation of improvement actions



School Cluster P Among the public basic schools with lowest academic results in Portugal (ranking) South of Portugal, Urban TEIP School			
Results Academic Success		Provision of Educational Services Monitoring of Assessment	
SUFFICIENT		SUFFICIENT	
+	Reduction of school dropouts	+	The results are analysed in the different educational coordination and pedagogical supervision structures
-	Completion rate below expected value	+	Identified internal causes related to the difficulty in implementing strategies that promote parental participation in students' school life (lack of work habits, study method, lack of subject matter, poor attendance)
-	Results below the national average	+	Strategies for articulating cycles and supporting improved learning (without producing desired results)
		+	Diagnostic used assessment at all levels of education
		+	Collaboration in preparation and correction of summative assessment tests that are not common
		+	In the 1st cycle, the exchange of assessment instruments is established
		+	Different use of formative assessment instruments, as main assessment method
		+	Clear assessment criteria defined in relation to the domains of assessment (knowledge and attitudes). As the instruments to be used in the evaluation and respective quantification
		+	Learning recovery and monitoring plans
		-	Lack of coherent and cohesive practice
		-	They have not yet found the most effective and appropriate strategies to bring parents closer to the lives of their students



Conclusions

This study allows us to conclude that, based on the examined reports, formative, reflective, and collaborative practices are being carried out in schools. These indicators demonstrate that we are moving in the right direction by prioritising assessment practices that place the student at the center of learning and allow for the regulation of the learning and teaching process. However, there is still a path to be pursued, particularly concerning the effectiveness of these practices. Are the assessment practices used genuinely formative? Are the results of external assessment tests taken into account to regulate the learning process? Another question that arises is the implication of these new demands and the new organization of curriculum in the context of initial teacher training. Since teachers are subject to a set of tensions in practice in the context of a performative and accountability culture (Pratt, 2018), it is urgent to find a way to complement the potential of internal and external assessment under a formative logic (Flores et al., 2017). According to Fernandes (2019) a transition towards a new era of external assessment grounded in socially constructed principles and focused on students' learning is needed. However, more recent studies point to the need for a broader and comprehensive investigation into the effectiveness of external evaluation in improving the quality of student learning (Machado et al., 2022; Sousa, 2020). Machado et al. (2022) suggests that the desired shift from a "certification culture" to a "formative culture" has not yet materialized, leading to changes that serve more political-ideological assumptions than the need to ensure effective improvement in student learning. In fact, the external evaluation has a more substantial impact on organisational development than on curricular and pedagogical development (Pacheco et al., 2023). As clues for future investigations, it would be important to analyse the leadership and management dimension alongside the results and provision of the educational service to analyse whether there are tensions or convergences between the 3 domains that are taken into consideration by the General-Inspectorate.

Funding

This work was financed by National Funds through the FCT - Portuguese Foundation For Science and Technology under the project "Assessment in Basic and Secondary Education: teachers' practices and conceptions of internal and external assessment in Portuguese schools" (CEECIND/03517/2018) and under the CIEC (Research Centre on Child Studies) project with reference UID/CED/00317/2019.

References

- Alves, P., & Machado, E. A. (2008). *Avaliação com sentido(s): contributos e questionamentos*. De Facto Editores.
- Biggs, J. (2003). *Teaching for Quality Learning at University*. SHRE and Open University Press.
- Black, P. (2015). Formative Assessment: An Optimistic but Incomplete Vision. *Assessment in Education: Principles, Policy and Practice*, 22(1), 161–177.



- Black, P., & William, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Broadfoot, P. M., Daugherty, R., Gardner, J., Harlen, W., James, M., & Stobart, G. (2002). *Assessment for learning: 10 principles*. University of Cambridge School of Education.
- Brown, G. (2004). Teachers' conceptions of assessment: implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301-318.
- Brown, G., & Hirschfeld, G. (2008). Students' conceptions of assessment: links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3-17.
- Brown, G., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27(1), 210- 220.
- Cabral, I., & Matias, J. (2017). *Da construção do sucesso escolar*. Fundação Manuel Leão.
- Decree- Law 55/2018, July 6th – Autonomia e Flexibilidade Curricular
- Decree-Law 54/2018, July 6th – Educação Inclusiva
- Earl, L. (2013). Assessment for learning: Assessment as learning; changing practices means changing beliefs. *Assessment and learning (Issue 2)*. The Government of the Hong Kong Special Administrative Regions, Education Bureau.
- Earl, L., & Katz, S. (2006). *Rethinking Classroom Assessment with a Purpose in Mind*. http://www.edu.gov.mb.ca/k12/assess/wncp/rethinking_assess_mb.pdf
- Fernandes, D. (2019). Para um enquadramento teórico da avaliação formativa e da avaliação sumativa das aprendizagens escolares. In M.I. R. Ortigão, D. Fernandes, T. V. Pereira, & L. Santos (Orgs.), *Avaliar para aprender em Portugal e no Brasil: Perspectivas teóricas, práticas e de desenvolvimento* (pp.139-164). CRV. <https://doi.org/10.24824/978854443463.5>
- Flores, M. A. (2017). *Práticas e discursos sobre currículo e avaliação. Contributos para aprofundar um debate*. De Facto Editores.
- Flores, M.A., Alves, M. P., & Machado, E. (2017). *Avaliação das aprendizagens e sucesso escolar: perspectivas internacionais*. De Facto Editores.
- Flores, M. A., & Pereira, D. (2019). Revisão da literatura. In M. A. Flores (Org.), *Avaliação no ensino superior: conceções e práticas* (pp.23-48). De Facto Editores.
- Fletcher, R., Meyer, L., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and Students Conceptions of Assessment in Higher Education. *Higher Education*, 64(1), 119- 133.
- Fulmer, G., Tan, K., & Lee, I. (2019). Relationships among Singaporean secondary teachers' conceptions of assessment and school policy contextual factors. *Assessment in Education: Principles, Policy and Practice*, 26(2), 166-183.
- Lam, R. (2016). Assessment as learning: examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies in Higher Education*, 41(11), 1900-1917.
- Machado, E. A. (2014). Entre a omniavaliação e a confiscação: contributos para uma avaliação dialógica sustentável. *Avaliação, Campinas; Sorocaba*, 19(2), 333-343.
- Machado, E. A., Flores, M. A., Pereira, D., Fernandes, E., & Costa, L. (2022). Políticas e práticas de avaliação externa dos alunos na perspectiva dos diretores: o caso das provas de aferição. *Revista Portuguesa De Investigação Educacional*, 23, 1-26. <https://doi.org/10.34632/investigacaoeducacional.2022.10958>



- Marcondes, M. (2017). Desenvolvimento curricular e materiais padronizados no contexto brasileiro. In M. A. Flores (Org.), *Práticas e discursos sobre currículo e avaliação. Contributos para aprofundar um debate*. De Facto Editores.
- Niklasson, L. (2017). Implementação de novas diretivas de avaliação: desafios para os professores das escolas na Suécia. In M.A Flores, E. Machado, & M. Alves (Orgs.), *Avaliação das aprendizagens e sucesso escolar* (pp.167-186). De Facto Editores.
- Pacheco, J. A. (2010, July 13). Conferência realizada no Seminário “Avaliação Externas das Escolas: Modelos, Práticas e Impacto”, realizado na Universidade do Minho.
- Pacheco, J. A., Fialho, I., Barreira, C., & Seabra, F. (2023). Perceções de atores políticos, inspetivos e avaliativos sobre a avaliação externa das escolas como política, procedimento e experiência. In I. Fialho, et. al. (Org.), *Avaliação Externa das Escolas. Mecanismos de Mudança nas Escolas e na Inspeção* (pp. 45-87). Editora Humus.
- Pacheco, J. A., & Maia, I. (2019). Avaliação das aprendizagens no contexto de políticas curriculares de accountability. *Linhas Críticas*, 25, e23778, Enero-Diciembre. <https://doi.org/10.26512/lc.v25.2019.23778>
- Pereira, D. (2016). *Assessment in Higher education and quality of learning: perceptions, practices and implications*. (Unpublished doctoral dissertation). Universidade do Minho.
- Pereira, D., Flores, M. A., & Machado, A. (2021). Autonomia e flexibilidade curricular na perspetiva dos diretores: entre o ceticismo e a adesão. *Indagatio Didactica*, 13(3), 57-76. <https://doi.org/10.34624/id.v13i3.25509>
- Perrenoud, P. (1999). *Avaliação, da excelência à regulação das aprendizagens: entre duas lógicas*. ARTMED.
- Pratt, N. (2018). Playing the levelling field teachers’ management of assessment in English primary schools. *Assessment in Education: Principles, Policy and Practices*, 25(5), 504-518.
- Samuelowicz, K., & Bain, J. D. (2002) Identifying academics’ orientations to assessment practice. *Higher Education*, 43(2), 173-201.
- Serra, L., Alves, J. M., & Soares, D. (2023). A avaliação externa das escolas: entre a conformidade e a inovação. *Revista Portuguesa De Investigação Educacional*, 25, 1-25. <https://doi.org/10.34632/investigacaoeducacional.2023.15913>
- Schreier, M. (2012). *Qualitative content analysis in practice*. Sage Publications
- Scouller, K. (1998). The influence of assessment method on students’ learning approaches: Multiple choice question examinations versus assignment essay. *Higher Education*, 35(4), 453-472.
- Snyder, B. (1971). *The Hidden Curriculum*. Knopf.
- Sousa, J. (2020). Avaliação externa das escolas em Portugal: Ensaio sobre racionalidade e significado. *Revista Educación, Política y Sociedad*, 5(2), 155-170. <https://doi.org/10.15366/reps2020.5.2.008>
- William, D. (2010). The role of formative assessment in effective learning environments. In H. Dumont, D. Instance & F. Benavides (Eds.), *The nature of learning; using research to inspire practice* (pp. 134-159). OECD Publishing.