frontiers in Education

TYPE Original Research
PUBLISHED 04 August 2026
DOI 10.3389/feduc.2026.1811512

OPEN ACCESS

EDITED BY
Paulo Marinho,
University of Porto, Portugal

REVIEWED BY
Kätlin Vanari,
Tallinn University, Estonia
Luisa Grützmacher,
Technical University of Braunschweig, Germany

*CORRESPONDENCE
Filipa Seabra
filipa.seabra@uab.pt
Joana Sousa
jsousa@ie.uminho.pt

†These authors share first authorship

RECEIVED 14 February 2026
REVISED 12 July 2026
ACCEPTED 15 July 2026
PUBLISHED 04 August 2026

CITATION
Seabra F, Sousa J and Álvaro MH (2026) Private and vocational schools from oversight to insight: external evaluation of schools and self-evaluation in dialogue.
Front. Educ. 11:1811512.
doi: 10.3389/feduc.2026.1811512

Private and vocational schools from oversight to insight: external evaluation of schools and self-evaluation in dialogue

Filipa Seabra[1,2*†], Joana Sousa[2*†] and Maria Helena Álvaro[2]

[1]LE@D, Laboratory of Distance Education and eLearning, Universidade Aberta, Lisbon, Portugal, [2]CIEd, Research Centre on Education, Institute of Education, University of Minho, Braga, Portugal

Introduction: External School Evaluation and School Self-Evaluation are increasingly understood as complementary mechanisms within national regulatory architectures, yet the way this relationship unfolds in types of schools only recently incorporated into state inspection frameworks remains under-researched. This study examines that relationship within Portuguese private and vocational schools, both of which entered the third cycle of External School Evaluation in 2018.
Methods: The corpus comprised 47 inspection reports published between 2018 and 2025 (24 from private schools and 23 from vocational schools), analysed using quantitative document content analysis. Qualitative inspectorate judgements for the self-evaluation domain were converted into a discrete five-point numerical scale to perform non-parametric comparative tests, while specific institutional strengths and areas for improvement were coded inductively to trace longitudinal trends.
Results: The statistical findings reveal a steady upward trend and progressive convergence between global inspection averages and self-evaluation scores across the cycle, suggesting that external judgements operate as a comparative baseline that stabilises internal routines. Vocational schools achieved significantly higher scores on both indicators and relied predominantly on the European Quality Assurance Reference Framework for Vocational Education and Training, which correlated positively with evaluative outcomes. Conversely, private schools relied on self-designed models, displaying lower procedural systematicity but a more pronounced textual focus on curricular and pedagogical dimensions.
Discussion: Rather than establishing a direct causal link, the findings suggest that external evaluation operates as a discursive mechanism that consolidates internal administrative regularities, while the translation of evaluative evidence into concrete classroom practice persists as a complex institutional challenge. Reconciling managerial compliance with shared pedagogical reflection remains a critical condition for sustainable school development.

KEYWORDS
data-based decision making, external evaluation of schools, private schools, reinforcement data use, school evaluation, school improvement, school self-evaluation, vocational schools

1 Introduction

For more than two decades, the question of how schools should be held accountable, and to whom, has shaped one of the most contested debates in European education policy. The expansion of external inspection regimes, the institutionalisation of Schools' Self-Evaluation, and the growing centrality of data in school governance are not isolated phenomena. They belong to a wider movement in which the state, faced with the limits of direct administrative control, has come to rely on evaluation as a soft instrument of regulation. This movement, often described as a shift from government to governance (Lo, 2018), has produced converging mandates across very different national systems, even as local conditions continue to shape their actual translation into school life.

The Portuguese case offers a productive vantage point on this broader landscape. External School Evaluation (*Avaliação Externa das Escolas*) was formally established by Law 31/2002 (Portugal, 2002) and is currently coordinated by the General Inspectorate of Education and Science (*Inspeção-Geral da Educação e Ciência*, IGEC). The third cycle of external evaluation of schools, initiated in 2018, has been widely described as a turning point in comparison with previous models (Fialho et al., 2023; Seabra et al., 2022). This shift effectively displaced the predominantly output-oriented logic that had spread across Europe during the previous decade (Ehren et al., 2015), redirecting institutional attention towards processes, educational provision, and, more specifically, the quality of teaching practices (Fialho et al., 2014; Seabra et al., 2022). This reorientation aligns with a broader set of public policies reinforcing curricular flexibility and pedagogical relevance. Law-Decrees 54/2018 and 55/2018, on inclusive education and curricular autonomy respectively (Portugal, 2018a, b), reinforced this orientation by demanding that schools contextualise their action within their specific communities. Four core domains now structure the external evaluation framework: *Self-Evaluation* (SE); *Leadership and Management*; *Provision of Educational Services*; and *Results* (General Inspectorate of Education and Science [IGEC], 2023). Schools' self-evaluation is, therefore, and for the first time in the present model, recognized as a separate domain to be assessed by external evaluation, and given prominence as the first of the four domains of assessment, as it is considered pivotal to producing change at schools.

Historically focused on public school clusters, external evaluation of schools was extended in its third cycle to private and vocational schools, integrating these types of schools into national regulation mechanisms and seeking to ensure that the 'grammar of innovation' and quality standards are consistent across diverse institutional legal frameworks (Martins et al., 2025). Crucially, this regulatory integration relies upon a unified scope, principles and objectives (General Inspectorate of Education and Science [IGEC], 2019a), Reference Framework (General Inspectorate of Education and Science [IGEC], 2023), methodology (General Inspectorate of Education and Science [IGEC], 2019b), and classification scale (General Inspectorate of Education and Science [IGEC], 2019c), which dictates that public, private, and vocational institutions are subjected to identical inspection criteria, quality descriptors, and indicators. By utilizing a singular, non-differentiated framework across all types of schools, the state enforces a uniform standard of accountability that treats institutional variations not as grounds for separate evaluative rubrics, but as distinct contextual landscapes navigating a shared national baseline. According to the official inspectorate guidelines (General Inspectorate of Education and Science [IGEC], 2019c), the qualitative assessments recorded in the evaluation reports are structured around five ordinal descriptors: *Insufficient*, *Sufficient*, *Good*, *Very Good*, and *Excellent*.

This late-adoption of the external evaluation model by vocational and private school sets the scene for the current paper, aiming to analyse how the domain of self-evaluation has been assessed by external evaluation reports of private and vocational schools, and to identify patterns regarding the processes of self-evaluation adopted by those schools, through the lens of external evaluation.

This situation is not unique to the Portuguese landscape but reflects a systemic challenge visible across diverse European educational systems. In Ireland, Brown et al. (2021) examine whether school self-evaluation operates as a universal driver for school improvement or as a culturally situated, country-specific construction, demonstrating that the depth of internal evaluative routines varies extensively with national governance configurations and professional cultures. Similarly, Swedish research by Hult and Segerholm (2016) highlights the tensional, often contradictory dynamics that arise when external inspection encounters internal self-reflection, warning that intensive external monitoring risks reducing self-evaluation to a ritualised exercise of procedural compliance. Recent evidence from Estonia (Vanari et al., 2025) further underlines that the meaningful appropriation of data for school improvement is structurally mediated by local leadership capacity and professional cultures, rather than by the mere existence of standardized quality assurance frameworks. Consequently, understanding how late-entry types of schools internalise external quality mandates represents an essential locus of inquiry for international policy analysis. The focus of this paper is, therefore, not to compare these two late adopter typologies of schools, but rather to characterise how the adoption of the same external evaluation framework may have impacted self-evaluation in these distinct typologies, in interaction with their differences.

This is the wider debate to which the present study seeks to contribute. The question is not simply whether external inspection "works", but under which conditions schools move beyond compliance what we provisionally describe, borrowing from the Portuguese policy lexicon, as a movement from oversight to insight, and under which conditions they do not. Sousa and Pacheco (2019) remind us that evaluation is embedded in a global climate of accountability, where data serve as instruments of knowledge-based regulation. The tension between a managerialist logic, oriented towards results and efficiency, and a professional logic, oriented towards curricular and pedagogical improvement, is well documented and rarely resolved. When the former prevails, schools often produce evidence that "looks good" to the external eye while protecting the classroom from genuine scrutiny, inducing a culture of performativity in which the production of appearances may displace deep, shared reflection (Ehren et al., 2015).

It is within this dialectic that schools' self-evaluation acquires its strategic value. Conceptualised as an 'inside-out' reflexive

procedure led by internal agents to produce contextualised knowledge (Figueiredo, 2023), self-evaluation is nonetheless, as Sousa and Pacheco (2019) caution, frequently prone to becoming a mere 'preparation for external evaluation', losing its emancipatory potential and functioning as a ritualised response to external pressure in which the school's 'image' eclipses its 'substance'. To counter this risk, Figueiredo (2023) proposes that Self-evaluation be understood along three operative axes: from within, as it emerges from the specific characteristics and needs of the school context, designed to 'measure' the school's identity; by within, as it is developed by the educational actors themselves (teachers, leaders, students, and parents) who possess unique familiarity with the school's reality; and for within, as it serves as an instrument of self-regulation and reflection that informs strategic decision-making for the sake of development and improvement.

This conceptualisation finds significant resonance in wider developments regarding educational quality assurance. Gardezi (2024), in a comparative analysis of inspection systems, frames schools' self-evaluation as a systematic information-gathering procedure initiated by the school itself, aimed at assessing its functioning and the achievement of its educational objectives, anchored in local context and progressively independent from purely external impositions. The 'by within' axis is further reinforced by literature that emphasises the inherently collaborative nature of the process: self-evaluation must actively involve leaders, teachers, students, and parents who, sharing concerns and interacting regularly to improve their performance, sustain a sense of collegiality and ownership essential for the durability of intended changes (Brown et al., 2020; Figueiredo, 2023; Miranda et al., 2023). The 'for within' axis, in turn, consolidates schools' self-evaluation as a tool for self-regulation: by adopting models such as the Plan-Do-Check-Act (PDCA) cycle, schools mobilise the collected data to inform strategic decision-making and school improvement planning (Gardezi, 2024; Martins et al., 2025).

As Gardezi (2024) further notes, the transition from a purely administrative process to a 'high trust' model requires the continuous development of the evaluative capacity of school leaders. Although schools' self-evaluation is primarily intended for internal development, its robustness allows the findings to also serve 'for outside', providing transparency and credibility to the institution before inspection bodies and the external community. The impact of evaluation, however, is mediated by the perception of fairness (Vanhoof et al., 2025). When the relationship between oversight (external evaluation) and insight (self-evaluation) is balanced, the external gaze provides the external anchor that legitimises change; when the accountability logic overrides the improvement logic, a significant translation gap persists, leading to what Sousa and Pacheco (2019) describe as a hybrid regulation, in which schools navigate between state-imposed standards and their own contextual needs. To fulfil its role 'for within' (Figueiredo, 2023), schools must resist self-evaluation being reduced to a bureaucratic tool and administrative workload that does not translate into better classroom support (Mouraz et al., 2014, 2019; Sousa and Pacheco, 2018). Evidence from Penninckx et al. (2016) suggests that, for schools' self-evaluation to effectively serve as an instrument for strategic decision-making, it must be sustained by a school's internal policy-making capacities, including innovative capacity, shared leadership, and supportive professional relationships. The robustness of these internal processes is, moreover, intrinsically linked to the perceived quality of the evaluation itself: when teachers perceive internal or external evaluation as having high quality, transparency, and respectful conduct, there is a significantly stronger conceptual and instrumental effect, with deeper reflection on school policy and a greater willingness to implement concrete improvement actions (Penninckx, 2017; Penninckx et al., 2016). Self-regulation, in this light, is not merely an administrative requirement but a professional learning culture that, when executed with rigour, provides the transparency necessary to satisfy external accountability while driving internal growth.

Seen from this angle, school self-evaluation is essential for school transformation: a meaningful approach identifies weaknesses and designs tailored actions for and within the institutional actors (Figueiredo, 2023; Sousa, 2020; Sousa and Pacheco, 2018). The practice of schools' self-evaluation articulates around three main pillars. Firstly, it serves a diagnostic function, producing a 'radiography' of the school by identifying strengths and weaknesses in organisational functioning and student outcomes (Figueiredo, 2023). Secondly, it follows an improvement logic, supporting informed decision-making to resolve problems and adjusting teaching practices (Ehren et al., 2013; Schildkamp et al., 2009). Thirdly, it fulfils an accountability and legitimacy function, providing transparent evidence of the school's quality to the external community and oversight bodies, thereby bridging the gap between internal insights and external expectations (Fialho et al., 2014; Vanhoof et al., 2025). Sousa (2020) argues that this dual and complex nature of self-evaluation for within and for the outside, places school evaluation between the poles of regulation and emancipation: on the one hand, it functions as a knowledge-based regulatory tool at the service of the state, enabling the monitoring of public policies at a local level; on the other, it constitutes a tool for school empowerment, enabling institutions to set their own agendas and assume responsibility for their improvement. The practice of school self-evaluation is, in addition, increasingly shaped by transnational organisations such as the OECD and the European Commission (Eurydice et al., 2015; Faubert, 2009; Sousa and Pacheco, 2019), which play a significant role in narrowing the gap between global agendas and local school realities (Fialho et al., 2014). This global influence often reinforces an accountability logic, which must be balanced both with an improvement logic oriented towards educational excellence and with an economic logic that frames schools' self-evaluation as a cost-effective alternative to external monitoring (Figueiredo, 2023).

While the literature and school evaluation practices traditionally focus on identifying gaps and correcting weaknesses, Aderet-German and Ben-Peretz (2020) highlight the importance of a balanced approach that also incorporates the intentional use of positive data. Contrary to the conventional focus on changing what is failing, these authors propose a complementary category, termed 'reinforcement data use', in which the identification of institutional strengths serves not only to enhance school ethos but also to provide a clear mandate for what must be sustained. Beyond the traditional instrumental,

conceptual, and symbolic uses of data (Alkin and King, 2016; Alkin and Taut, 2002; Cousins and Leithwood, 1986; Johnson et al., 2009; Shulha and Cousins, 1997), the systematic acknowledgement and celebration of strengths allow for the consolidation of practices that should be sustained. This process reinforces teacher commitment and team spirit, enhances the school ethos, and elevates the recognition of academic achievements and core values into strategic resources for innovation and organisational development. Unlike symbolic use, which often serves performative or merely celebratory purposes, reinforcement use provides a clear mandate for the continuity of successful practices, ensuring that they are institutionalised and shielded from organisational instability.

This strategic use of data is rooted in the concept of Data-Based Decision Making (DBDM), which draws its theoretical foundations from feedback and goal-setting theories. Visscher (2021) argues that, although trained educators are generally capable of interpreting student progress data, a significant translation gap often persists: schools frequently struggle to convert evaluative insights into concrete, tailor-made classroom instruction — a gap that no model, by itself, can close. This challenge resonates with the findings of Figueiredo et al. (2024), who note that school leaders in Portugal particularly value schools' self-evaluation when it provides rigorous and contextualised information that validates previously made decisions or planned strategic options, by grounding their actions in what they perceive as 'scientific' or 'solid' evidence, leaders gain a sense of security and sustainability in their institutional management. The effectiveness of such data-driven practices, however, is highly dependent on the institutional climate and the primary drivers of evaluation. Schildkamp and Datnow (2022) warn that, when data use is focused predominantly on external accountability and compliance rather than on internal improvement, the work of school data teams can be severely hindered. Without a clear shift towards a learning-oriented culture, the collection of evidence risks becoming a bureaucratic exercise that fails to stimulate the deep organisational changes required for sustained educational quality. As Vanhoof et al. (2025) emphasise, the impact of evaluation is mediated by the professionalism and credibility of the feedback source, which determines whether the data are accepted as a legitimate lever for change.

What remains analytically underdeveloped, however, is how this normative ideal is translated into practice. The literature on quality management in schools tends to assume that the adoption of a model — EQAVET, ISO 9001, CAF, EFQM, or proprietary frameworks — automatically generates institutional change. This assumption deserves scrutiny. Models do not institutionalise themselves. They become routinised through repeated use, through the accumulation of small organisational decisions, and through the work of intermediaries as self-evaluation teams, middle leaders, pedagogical coordinators, who translate abstract indicators into recognisable artefacts: an improvement plan, a meeting agenda, a dashboard, a paragraph in a report. Institutionalisation, in this sense, is less the outcome of a framework than the slow sedimentation of practices that, over time, come to be taken for granted. European Quality Assurance Reference Framework for Vocational Education and Training (EQAVET), for instance, can remain a bureaucratic exercise — a label periodically renewed by the National Agency for Qualification and Vocational Education and Training (ANQEP) — or evolve into a shared grammar through which schools interpret their own action. The difference does not lie in the model itself but in the density of the routines that grow around it, in the credibility attributed to the evidence it produces, and in the extent to which it informs decisions that matter to teachers and school leaders. This is where the notion of leadership for learning, such as presented by Hallinger (2011) comes into play. Leadership for learning refers to "approaches that school leaders employ to achieve important school outcomes, with a particular focus on student learning" (p. 126). School leaders, including but not limited to the school principal but rather the multiple school agents within the school that may be seen as leaders in a broader sense, have a pivotal role in the process of transforming information—such as that received through external evaluation reports, or by self-evaluation—into changes that can ultimately affect student outcomes. These are mediated by changes in academic structures and processes, but also in vision and goals for the school and capacity building. This vision aligns closely with Ehren and Shackleton's (2016) account of how inspections produce improvement in schools, which acknowledges that feedback can only produce change if it is accepted, if there is enough capacity, and if improvement actions are effectively implemented, but adds and highlights the role of educational leaders in promoting that acceptance, capacity and designing and implementing effective changes to school goals, structures and processes. The notion of translation of change into the classroom is also in line with the role of leadership for learning, insomuch as teachers can be seen as leaders in their own classrooms, and their agentic role in this process is essential for improvement.

The diverse institutional nature of the types of schools now integrated into external evaluation of schools is reflected in the variety of quality management models adopted to structure their internal processes. In the vocational schools, the EQAVET is the predominant reference, mandated by national regulation (European Quality Assurance Reference Framework for Vocational Education and Training [EQAVET], 2025). Grounded in the quality assurance and improvement cycle — planning, implementation, evaluation/assessment, and review/revision — echoing the renowned Deming Cycle (Plan-Do-Check-Act, PDCA) (Gardezi, 2024; Martins et al., 2025), EQAVET (2025) emphasises continuous improvement through specific indicators such as completion and placement rates, fostering institutional reflexivity and the routinisation of evaluative practices. Beyond EQAVET, other frameworks recur in school reports, notably the ISO 9001 international standard, which centres on administrative and procedural rigour (International Organization for Standardization [ISO], 2015), and the Common Assessment Framework (CAF), a quality management tool designed for public administration and mobilised as a Total Quality Management instrument to enhance organisational performance in public-sector contexts (European Institute of Public Administration [EIPA], 2020). Some institutions additionally implement the European Foundation for Quality Management (EFQM) model (Ruiz-Carrillo and Fernández-Ortiz, 2005) or have participated in regional initiatives such as the External School Evaluation

Programme (AVES). The AVES Programme (2000–2001), implemented by the Manuel Leão Foundation, constituted one of the first systematic school evaluation exercises in Portugal, anticipating the state model. Conceptually anchored in Stufflebeam's CIPP model—Context, Input, Process, Product—(Stufflebeam and Zhang, 2017) and in the currents of school improvement and learning organisations (Bolívar, 2000), AVES introduced a methodology based on the complementarity between self-evaluation and the mediating gaze of external consultancy, distancing itself from purely inspectoral and bureaucratic logics. Ultimately, it provided schools with validated instruments and external scientific support to monitor school climate and pedagogical effectiveness (Fundação Manuel Leão, 2024). Methodologically, both EQAVET and ISO 9001 are grounded in the PDCA cycle, ensuring systematicity across the phases of planning, implementation, evaluation, and review, although their focal points diverge: ISO 9001 emphasises general organisational efficiency and procedural standardisation, whereas EQAVET (2025) is specifically oriented towards the pedagogical dimension and VET-specific outcomes, such as completion rates (Indicator 4) and placement rates (Indicator 5). These models are widely regarded as complementary, and many schools combine ISO 9001 for structural efficiency with EQAVET for pedagogical results — a synergy that fosters institutional reflexivity and helps sustain evaluative practices in the face of personnel turnover.

The strategic hybridization of quality assurance models has emerged as a robust response to the complexity of educational governance. By combining ISO 9001, which targets structural and processual efficiency, with the EQAVET framework, which isolates and monitors pedagogical outcomes, institutions create a comprehensive evaluative matrix. Theoretically, this duality addresses the 'loose coupling' inherent to school organisations (Weick, 1976), balancing administrative compliance with educational relevance. Furthermore, this synergy fosters deep institutional reflexivity, transitioning from single to double-loop learning (Argyris and Schön, 1978).

Portuguese education provides a particular configuration of these conditions. The system remains highly centralised in curricular and regulatory matters, even after two decades of incremental moves towards school autonomy. The Ministry of Education sets national curricula, defines teacher recruitment rules, and authorises school networks, the General Inspectorate for Education and Science operates as a central inspectorate with national reach. At the same time, schools retain considerable margin in pedagogical organisation, internal structures, and educational projects, and the 2018 reforms reinforced this margin through curricular flexibility and inclusive education frameworks. The result is a hybrid arrangement in which centralised standards coexist with locally negotiated practices, a configuration that is likely to shape, in distinctive ways, how schools internalise external pressure and what counts, for them, as legitimate evidence. This structural condition matters analytically: it sets the limits of what self-evaluation can plausibly achieve and qualifies the transferability of our findings to systems with different governance configurations. It should be noted that, although schools' self-evaluation is mandatory, no specific model is mandated in general, as schools are free to adopt pre-existing models or structure their own models. The exception to this rule are professional schools, which are mandated to use EQAVET, but are still free to complement that model with others—this sets this type of school apart in the national landscape and makes it particularly relevant to study.

The relationship between schools' external evaluation and self-evaluation, in this context, is essentially dialectical. External evaluation often acts as a motor or lever that induces schools to systematise their internal self-evaluation practices (Vanhoof et al., 2025), but the effectiveness of this dialogue depends on the perception of the evaluation process, and school inspections function as catalysts for quality primarily when they are perceived as fair in terms of procedural and interactional justice (Vanhoof et al., 2025). When schools perceive external evaluation as a validation of their own self-reflection, the external gaze provides the sustenance, or external anchor, that legitimises internal change (Seabra et al., 2022). Conversely, a lack of alignment can generate ambivalence in teachers' perceptions, potentially leading to curricular isolation (Morgado et al., 2020; Sousa and Pacheco, 2018) — a challenge that becomes particularly acute when schools must value intangible educational dimensions, a need made evident by the COVID-19 pandemic (Morgado et al., 2020). In Portugal, schools' external evaluation has clearly driven the formalisation and institutionalisation of self-evaluation structures and procedures, especially via self-evaluation teams and external evaluation-aligned frameworks (Correia et al., 2015; Costa et al., 2025); indeed, the first cycles of schools' external evaluation were seen as the driving agent for the implementation of self-evaluation through soft-regulatory influences (Sousa and Terrassêca, 2015). It is unsurprising, therefore, that the impact of external evaluation on self-evaluation is one of the areas in which the influence of external evaluation is most consistently identified (Figueiredo, 2024; Seabra et al., 2022) and a critical pivot for global institutional change (Ehren and Shackleton, 2016; Ólafsdóttir et al., 2022).

Private and vocational schools occupy a particular position within this landscape. Both types of schools were exempted from external evaluation in Portugal until 2018, despite distinct trajectories in self-evaluation. Vocational schools have been required, since Decree-Law 92/2014 (Portugal, 2014), to align their internal processes with EQAVET (National Agency for Qualification and Vocational Education and Training [ANQEP], 2018). Many vocational schools also adopt ISO 9001, the CAF, or, less frequently, the EFQM model. Private schools, by contrast, are not mandated to follow any specific framework, and for many of them the third cycle of schools' external evaluation was the first occasion on which their self-evaluation practices were exposed to an external observer.

This combination of features makes the two types of schools particularly informative. The analytical rationale for studying them jointly is threefold. Firstly, they constitute late-entry cases: as schools that joined the external evaluation programme more recently, they allow us to observe early-stage institutionalisation processes that, in public schools, are already sedimented and therefore harder to read. Secondly, they offer an internal contrast between two distinct modes relating to formal quality management models: one strongly framed by a mandatory European reference (EQAVET), the other more dispersed and

identity-driven. Thirdly, despite their growing weight in the Portuguese school network, they remain comparatively under-researched in the international literature on school evaluation, which has tended to concentrate on public schools. As such, they offer a privileged territory for examining how schools' self-evaluation has developed in types of schools that, until recently, were not subject to the soft regulation of external evaluation, and whether their incorporation into the programme may generate discernible shifts in how these processes are evaluated (Amorim et al., 2021; Costa, 2024; Luís, 2022).

Building on this rationale, the present study pursues two specific objectives: first, to analyse how external evaluation of schools' reports assess school's self-evaluation processes in private and vocational schools between 2018 and 2024/25, taking into account the positive aspects that may contribute to reinforcement data use, as well as the areas identified as needing improvement; and second, to identify patterns and differences in school's self-evaluation processes between private and vocational schools across this period, as reflected in inspection reports. To guide this comparative inquiry, two operational research hypotheses were formulated based on the distinct regulatory trajectories of the two types of schools. First (H_1), it is hypothesised that vocational schools will display significantly higher quantitative evaluation scores in both global classifications and the self-evaluation domain than private schools, due to their extended administrative experience with the mandatory EQAVET quality assurance framework. Second (H_2), it is hypothesised that the formal adoption of structured quality management models will correlate positively with procedural systematicity in school reports, whereas the reliance on self-designed models will correlate with an increased frequency of identified areas for improvement in procedural scope. By examining the strengths and areas for improvement signalled by external inspectors, the study seeks to clarify how the dialogue between external oversight and internal quality management models may induce, or fail to induce, institutional change and contribute to a robust culture of self-regulation.

To understand how external monitoring induces internal change, it is necessary to examine the mechanisms through which formal quality management models become embedded within school organisations. Investigating private and vocational schools offers a valuable opportunity to analyze this dynamic. Having been exempted from external inspections until 2018, these types of schools function as distinctive late-entry cases, allowing researchers to observe early-stage institutionalisation processes that are already heavily sedimented in public school clusters. Furthermore, they provide an internal structural contrast: vocational schools operate under a mandatory, highly standardized European reference framework (EQAVET), whereas private schools rely on heterogeneous, identity-driven, self-designed models. Examining these types of schools clarifies how distinct administrative baselines shape a school's capacity to move beyond symbolic compliance towards internal self-regulation. The primary purpose of the comparison is to examine two types of schools with contrasting quality assurance traditions, and that are late entry cases to the model of external evaluation of schools, rather than contrast either of them with other types of schools already subjected to this model for decades.

2 Methods

The third cycle of school's external evaluation (2018–ongoing) so far produced a total 47 reports of non-vocational private schools (henceforth referred to as 'private' for conciseness) ($n = 24$) and vocational schools ($n = 23$), as of February 2026. These reports are publicly available through the Inspectorate's website. The distribution of reports by school type and academic year is presented in Table 1.

The 47 reports were downloaded and analysed using MaxQDA (v. 24), with statistical analyses conducted in SPSS (v. 30.0). This study is based on quantitative content analysis of documents (Bardin, 2016). The documents were systematically reviewed and coded to identify key patterns, themes, and structural elements and was structured in two complementary stages.

The first stage addressed the first objective, involved the quantitative analysis of the scores and evaluative judgements expressed in the external evaluation reports regarding the independent domain of SE and the consolidated global score. To facilitate statistical manipulation and permit robust non-parametric comparative analysis, the five qualitative descriptors established and attributed by the external evaluation teams (Insufficient, Sufficient, Good, Very Good, and Excellent) were converted into a discrete five-point numerical scale, where Insufficient maps to 1 and Excellent maps to 5. A singular numerical score was thus parameterised for the independent domain of self-evaluation. Concurrently, an overall score (Global Classification) was computed for each institution by calculating the arithmetic mean of the numerical values assigned across all four core domains of the inspection framework. These classifications were compared between types of schools using Mann–Whitney U tests to evaluate group differences and test H_1, while Spearman's rank correlation coefficients (r_s) supported the identification of longitudinal tendencies and model associations to test H_2 (Marôco, 2018).

The second stage addressed the second objective and focused on the content analysis of the specific segments within each report that explicitly detailed the strengths, areas for improvement, and quality management models in use. Unlike Stage 1, which converted the formal domain scores into numbers, Stage 2 applied an inductive textual coding technique to the narrative text of the reports using MaxQDA. Three *a priori* categories,

TABLE 1 Distribution of external evaluation reports by type of school and year of evaluation (2018–2024/25).

Year of evaluation	Private	Vocational	Total
Pilot	2	1	3
2018/19	3	1	4
2019/20	3	1	4
2021/22	7	1	8
2022/23	4	0	4
2023/24	3	7	10
2024/25	2	12	14
Total	24	23	47

namely strengths, areas for improvement, and models of self-evaluation, organised the coding tree, which was subsequently refined through inductively developed subcategories.

It is important to make explicit, at this point, where the analytical categories come from. The four *domains* of self-evaluation are not constructs developed by the authors: they are part of the official IGEC (2023) reference framework and structure every inspection report. What was coded inductively were the subcategories within the strengths and areas for improvement signalled by inspectors in the self-evaluation domain. The distinction matters for the credibility of the analysis, and we return to it below. Likewise, strengths and areas for improvement are identified by the external evaluation teams in a separate section of the report and are short textual descriptions of the aspects that these teams have deemed the most salient aspects of the school's work. Three *a priori* categories were established: strengths, areas for improvement, and models of SSE (Schools' self-evaluation). These organised the coding tree. The strengths and areas of improvement (which are identified as such in a section of the report comprising between 0 and 3 strengths and similar number of areas for improvement regarding each of the domains of external school evaluation) were then quantified and coded through inductively developed subcategories. Despite frequently encompassing more than one idea, only the main idea of each paragraph was coded. The coded segments were analysed, and contrasts were established between private and vocational schools, and tendencies throughout the time span of the third cycle of external evaluation were identified (Bingham, 2023).

Strengths and areas for improvement are explicitly identified as such in the external evaluation reports. They consist of short topics, organised according to the domain of the report they coincide with. Their brevity and conciseness advised using the topic itself as the registry and enumeration unit (Amado et al., 2014). Some examples of these topics are: "Development of mechanisms for monitoring improvement actions" as a strength, or "Deepening and clarification of the role of the Quality Observatory as a structure that consolidates and drives the self-evaluation process" as an area for improvement. As these are loose sentences/topics, organised in a table in the reports, there is no context unit available. We adopted exclusivity as a criterion for categorization and coding, whereby each unit of registry cannot belong to more than one category (Bardin, 2016). Therefore, only the primary analytical idea of each topic was coded. The absolute frequencies of these emerging subcategories were then tabulated and compared between private and vocational cohorts using Spearman's correlations to map out thematic shifts across the evaluation cycle. Coding was discussed among the co-first authors and agreed upon. This was particularly relevant in determining the most salient idea that should be coded for each topic, which happened in some slightly longer and more complex sentences, such as "Consultation with and participation of the educational community, as well as communication strategies to foster a better understanding and grasp of the self-evaluation process results".

A more detailed description of self-evaluation practices is also presented in the external evaluation reports. This more open account is nevertheless organized and structured by the external evaluation framework adopted by the General Inspectorate of Education and Science (2023). For the domain of Self-evaluation, this framework includes the following fields of analysis and referents (Not all indicators were presented for brevity) (Table 2).

The description of the self-evaluation process on the reports was also used to identify the models of self-evaluation implemented in these schools, where, for this category, emerging subcategories were not mutually exclusive, as some schools used more than one model simultaneously. This analysis is still fundamentally quantitative, despite being focused on emerging themes, and responds to the second objective. The categories and emerging subcategories are represented in the Figures 1–3, below. We should note that, since the identification of the models of self-evaluation present in these schools is based exclusively on the external evaluation reports, this analysis may not be exhaustive. Nevertheless, references to absence of a specific model were never inferred, but explicitly stated in the reports themselves.

A central methodological issue concerns the conceptual boundaries between subcategories. In effect, the proposed subcategories, although emerging from the data, are deeply related to the indicators of the framework that guides the external evaluation of schools, as this framework is determinant

TABLE 2 Framework for external evaluation of schools—domain of self-evaluation.

Domain	Fields of analysis	Referents	Indicators
Self-evaluation	Development	Organization and sustainability of self-evaluation	- Systematic procedures of school self-evaluation - Broad auscultation and participation by the educational community
		Strategic planning of self-evaluation	- Central role of teaching and learning processes - Strategies for communication and reflection about the results of self-evaluation with the educational community
	Consistency and impact	Consistency of self-evaluation practices	- Wide scope of the data gathering processes - Rigor of data analysis processes
		Impact of self-evaluation practices	- Evidence [of impact] on the organizational improvement of the school - Evidence [of impact] on curricular development - Evidence [of impact] on the teaching and learning processes …

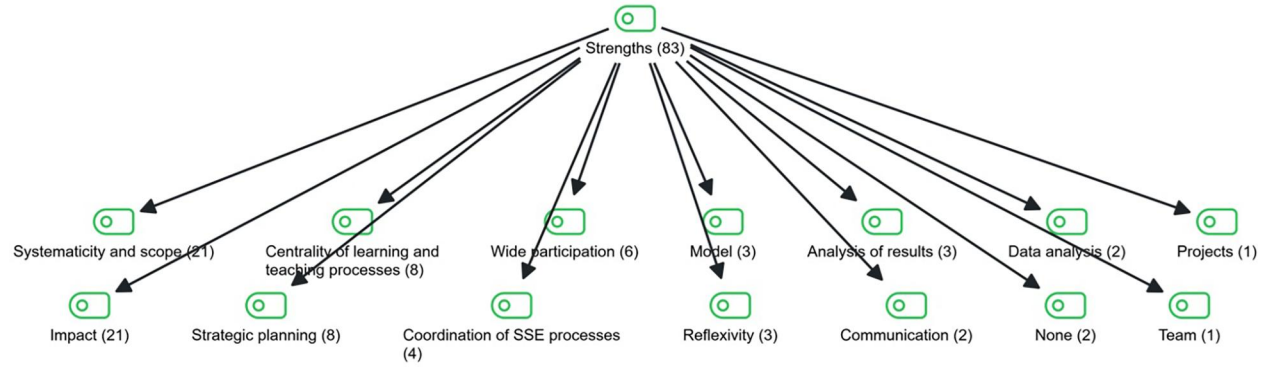

FIGURE 1
Category "strengths" in the SE domain and its subcategories.

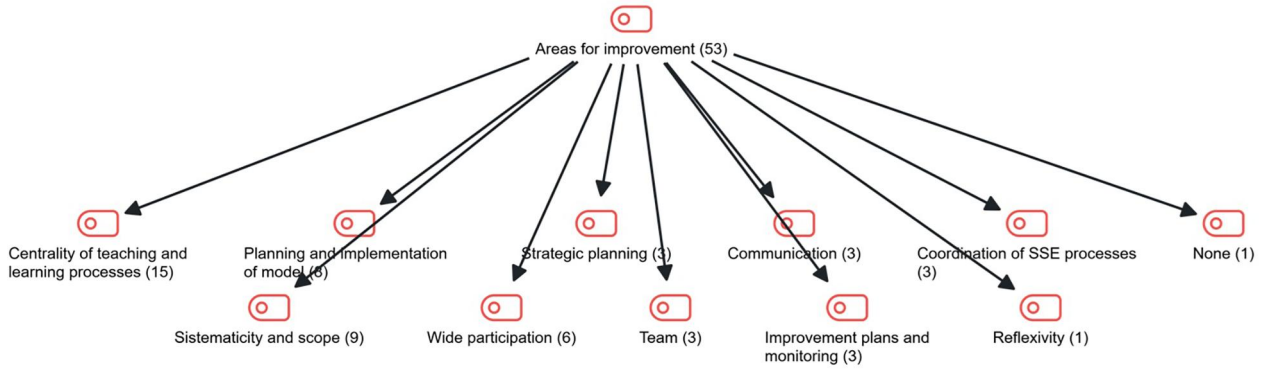

FIGURE 2
Category "areas for improvement" in the SE domain and its subcategories.

to what the external evaluation teams choose to report, and how they choose to report it. Therefore, just as some of the indicators presented in Table 1 somewhat overlap with each other (for example "broad auscultation and participation by the educational community" somewhat overlaps with "strategies for communication and reflection about the results of self-evaluation with the educational community") the same also happens with some subcategories of analysis. The authors, however, chose to always select a single category in each case; only the one that best expressed the main, and most salient aspect of each paragraph detailing a strength or an area for improvement.

A final note on ethical procedures is warranted. The corpus consists exclusively of documents publicly available on the website of the General Inspectorate of Education and Science, which makes the identification of any specific school technically possible for any reader who searches the database. To minimise the risk of inadvertent exposure in academic citation, the references to specific reports were revised throughout the manuscript: schools are now identified by a generic code (e.g., Vocational School A, Private School B), with the correspondence to the original reports kept by the authors. This procedure aligns the study with European data protection requirements.

3 Results

As previously mentioned, the first level of data analysis, directed at answering the first objective, is focused on quantitative analysis of the classifications attributed by the external evaluation teams to the domain of self-evaluation, and the global classification of the schools subjected to external evaluation (Figure 4).

The average classification of the SE domain has been consistently lower than the global average throughout the third cycle of external school evaluation. However, these two averages have been converging, and (excluding the 2018 pilot) both averages show a clear upward trend.

Significant differences were found between school types: private schools presented lower global and self-evaluation classifications than vocational schools (Figures 5, 6). Mann–Whitney U tests confirmed these differences for both global ($U = 125{,}5$, $p = .001$) and SE ($U = 124{,}0$, $p < .001$) classifications.

Stage 2 involved analyzing the strengths and areas for improvement identified in the reports. Following the coding process, code matrixes were generated. Consistent with the superior quantitative scores in this domain, vocational schools exhibited a higher frequency of identified strengths compared to private schools. Notably, in one school of each type, no

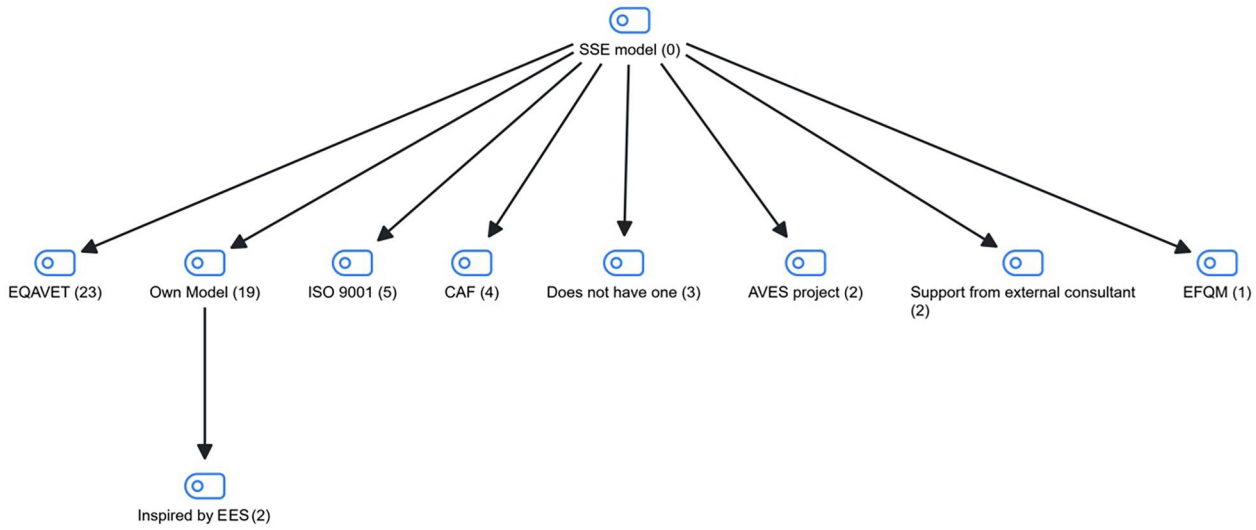

FIGURE 3
Category "model of SSE (schools' self-evaluation)" and its subcategories.

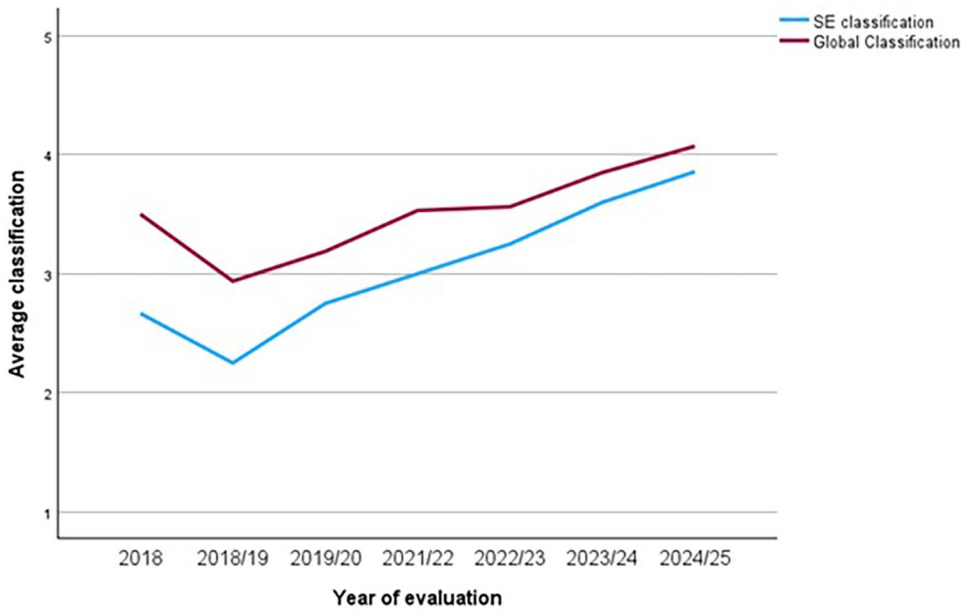

FIGURE 4
Evolution of the average overall classification and the average self-evaluation (SE) domain classification throughout the third cycle of external evaluation of schools (2018–2025).

strengths were identified in this domain. The most frequently reported strengths were the systematicity and scope of the school self-evaluation processes (e.g., "*Systematicity of the procedures adopted, adjusted to the reality of the School, covering both core and diversified areas*" (Private School A, p. 4) and their subsequent impact (e.g., "*Impact of SE on the development of strategic partnerships that enhance the diversification of learning contexts…*" (Vocational School A, p. 4).

These are followed by the central role attributed to the learning and teaching processes (e.g., "*The development of a SE microcycle, focused on the processes of teaching, learning, and assessment of classes and students, which enables their regulation and more rapid and effective educational intervention*" (Vocational School B, p. 4) and strategic planning of the school self-evaluation process (e.g., "*Elaboration of partial improvement plans that integrate strategic actions, result indicators and their monitoring*" (Private School B, p. 4) (Table 3).

Regarding the areas for improvement, the reverse is true: Private schools' reports identify a higher number of improvement areas in the domain of SE compared to vocational schools. Notably, the only school that presented no areas for improvement in this domain was a vocational school. The most frequent theme of the areas for improvement related to the centrality of teaching and learning processes in schools' self-

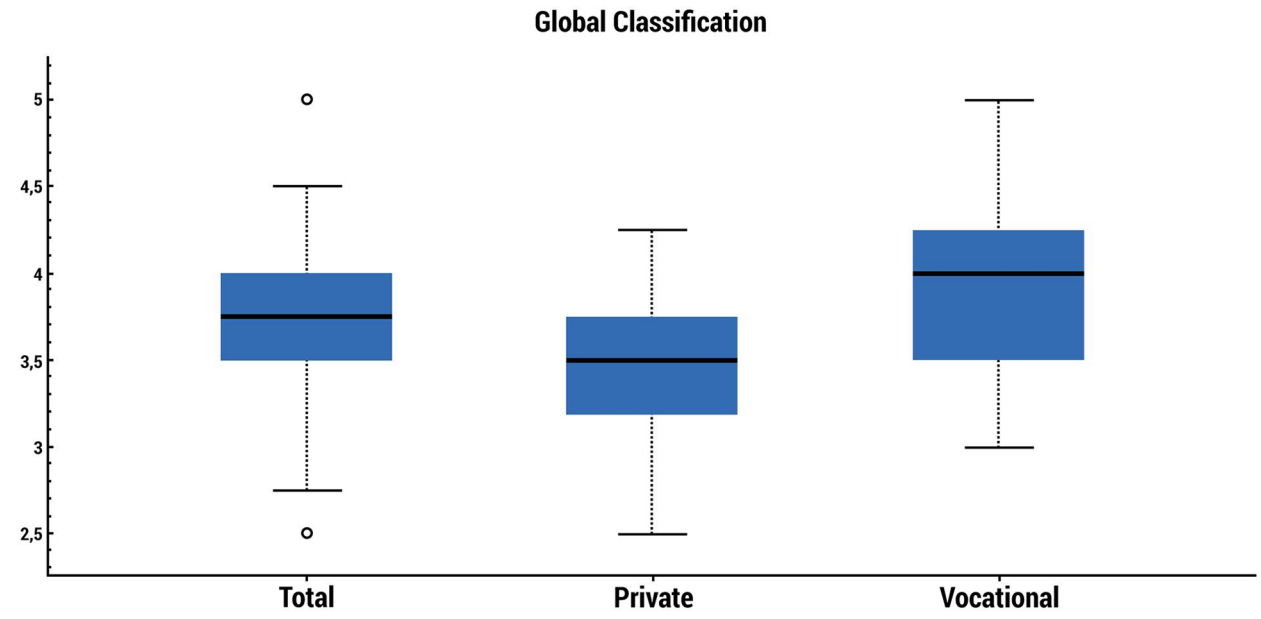

FIGURE 5
Global classification distribution, by type of school.

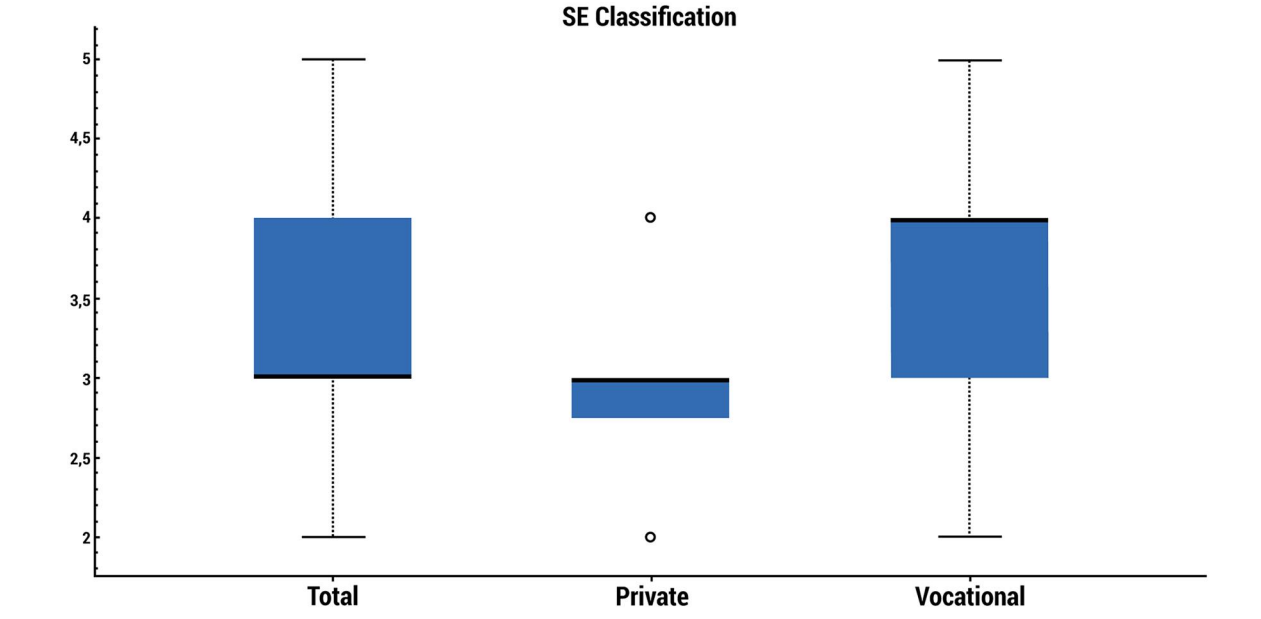

FIGURE 6
SE domain classification distribution, by type of school.

evaluation processes (e.g., "*Centrality of the teaching and learning process in SE procedures*" (Vocational School E, p. 5). Interestingly, this particular area for improvement is far more frequent among vocational schools than among private schools, and likewise, it has also been slightly more frequent as a strength among private schools. This theme is followed by the systematicity and scope of the school's self-evaluation (e.g., "*Implementation of a comprehensive and systematic SE process, consistent with the school's educational policy choices*" (Private School B, p. 4); an aspect only identified as an area for improvement in private schools, the planning and implementation of a model (e.g., "*To plan and implement a self-evaluation process tailored to the reality of the School that continuously supports decision-making and contributes to the improvement of student learning and outcomes*" (Private School C, p. 5), and the need for wide participation in self-evaluation (e.g., "*Expansion of the self-evaluation team to include other members of the educational community and the establishment of a communication and reflection plan regarding the results of self-evaluation practices*" (Private School D, p. 5) (Table 4).

TABLE 3 Distribution of strengths in the SE domain by type of school.

Strengths	Vocational	Private	Total
Systematicity and scope	14	7	21
Impact	12	9	21
Centrality of learning and teaching processes	3	5	8
Strategic planning	4	4	8
Wide participation	4	2	6
Coordination of SSE processes	3	1	4
Model	3	0	3
Reflexivity	1	2	3
Analysis of results	0	3	3
Communication	2	0	2
Data analysis	2	0	2
Projects	0	1	1
Team	0	1	1
Total	48	35	

TABLE 4 Distribution of areas for improvement in the SE domain by type of school.

Areas for improvement	Vocational	Private	Total
Centrality of teaching and learning processes	12	3	15
Systematicity and scope	0	9	9
Planning and implementation of model	2	6	8
Wide participation	2	4	6
Strategic planning	1	2	3
Team	0	3	3
Communication	2	1	3
Improvement plans and monitoring	1	2	3
Coordination of SSE processes	2	1	3
Reflexivity	1	0	1
Total	23	31	

TABLE 5 Distribution of quality models by type of school.

Model of SE	Vocational	Private	Total
EQAVET	21	2	23
Own model	4	17	21
ISO 9001	2	3	5
CAF	2	2	4
Does not have one	1	2	3
AVES project	0	2	2
Support from external consultant	0	2	2
EFQM	0	1	1

The models of self-evaluation described in the reports were also disproportionally distributed among types of schools, with vocational schools overwhelmingly and unsurprisingly implementing the EQAVET model (whether combined with other processes or not), and private schools having a higher representation of self-evaluation models developed by the schools themselves (Table 5).

Spearman's rank correlations coefficients were calculated to assess the relationship between the self-evaluation models implemented and the SE domain classifications. The use of the EQAVET showed a significant positive correlation with SE classification ($r_s = .353$, $p < .01$). Conversely, the absence of a specific model was found to have a significant negative correlation with the SE domain scores ($r_s = -.318$, $p < .05$). Regarding areas for improvement, the adoption of the EQAVET yielded a significant positive correlation with the centrality of teaching and learning being reported as an area for improvement ($r_s = .334$, $p < .05$). Conversely, it showed a significant negative correlation with the systematicity and scope of the processes as an area for improvement ($r_s = -.476$, $p < .001$). No significant correlations were found between the adoption of the EQAVET model and the most frequently reported strengths. However, the use of a self-designed school model was found to correlate negatively and significantly with reporting strengths related to the systematicity and scope of self-evaluation ($r_s = -.276$, $p < .05$). Consistently, this internal modeling also showed a significant positive correlation with identifying systematicity and scope as an area for improvement ($r_s = .268$, $p < .05$).

An analysis of the evolution of strengths over time reveals significant trends. While the most prevalent strengths persist throughout the entire cycle of external evaluation of schools, several key themes have only emerged as strengths in recent years. These include the centrality of teaching and learning, strategic planning, the coordination of diverse self-evaluation processes, the evaluation models themselves, reflexivity, and data analysis. Notably, the reports that failed to identify any strengths regarding self-evaluation were exclusively found in the early years of the present cycle of external evaluation. However, any longitudinal analysis must account for the non-uniform distribution of reports across the study period. As detailed in Table 6, the academic years 2021/22, 2023/24, and 2024/25 provided the largest volume of data, representing the most significant contributions to the current *corpus*.

A longitudinal analysis of areas for improvement reveals a trend similar to that observed in institutional strengths, with the centrality of teaching and learning processes gaining increased prominence. Conversely, systematicity and scope of the self-evaluation process has dwindled as an area for improvement and increased as a strength, signaling a transition in the nature of the self-evaluation processes implemented over time. Other categories do not evidence clear patterns across time, possibly due to a smaller number of references (Table 7).

4 Discussion

The longitudinal analysis of the evolution of classifications signals that the relationship between external school evaluation and school self-evaluation is dialectical in nature, with the

TABLE 6 Evolution of strengths over time.

Strengths	Pilot	2018/19	2019/20	2021/22	2022/23	2023/24	2024/25	Total
Systematicity and scope	0	1	1	5	2	6	6	21
Impact	1	1	2	2	1	6	8	21
Centrality of learning and teaching processes	0	0	1	0	2	2	3	8
Strategic planning	0	0	1	0	1	3	3	8
Wide participation	0	1	0	1	1	2	1	6
Coordination of SSE processes	0	0	0	1	0	1	2	4
Model	0	0	0	0	0	0	3	3
Reflexivity	0	0	0	1	1	0	1	3
Analysis of results	0	1	1	1	0	0	0	3
Communication	0	0	0	0	0	1	1	2
Data analysis	0	0	0	0	0	1	1	2
None	1	1	0	0	0	0	0	2
Projects	1	0	0	0	0	0	0	1
Team	1	0	0	0	0	0	0	1

TABLE 7 Evolution of areas for improvement over time.

Areas for improvement	Pilot	2018/19	2019/20	2021/22	2022/23	2023/24	2024/25	Total
Centrality of teaching and learning processes	0	0	1	1	1	5	7	15
Systematicity and scope	0	2	2	3	1	1	0	9
Planning and implementation of model	0	1	0	3	0	1	3	8
Wide participation	1	0	1	1	1	0	2	6
Strategic planning	0	0	0	1	1	1	0	3
Team	0	0	0	2	0	0	1	3
Communication	1	0	0	0	0	1	1	3
Improvement plans and monitoring	1	1	0	1	0	0	0	3
Coordination of SSE processes	0	0	0	0	0	3	0	3
Reflexivity	0	1	0	0	0	0	0	1
None	0	0	0	0	0	0	1	1

former functioning as a structural mechanism that induces the systematisation of internal practices. This phenomenon is inferred from the upward trend and convergence between the global average and the SE domain average throughout the third cycle (Figure 4). Such a trajectory corroborates the perspective of Vanhoof et al. (2025), suggesting that the external gaze, when perceived as legitimate, provides an external anchor that validates and drives internal formalisation. Furthermore, this convergence suggests that external evaluation results may provide a source of legitimacy for school leadership. In line with recent research on Portuguese school leaders (Figueiredo et al., 2024), these findings indicate that schools' self-evaluation is increasingly valued when it provides rigorous and contextualized information that validates strategic decisions or planned institutional options. By grounding their actions in what is perceived as solid, external evidence, leaders gain a sense of security and institutional stability in their management routines —as we have previously addressed in the introduction of this paper, information can only be translated into practice if it is first accepted (Ehren and Shackleton, 2016), and then mobilized by leaders for learning (Hallinger, 2011) to introduce changes in institutional and classroom practices and support capacity. The evidence of convergence in results over the course of the third cycle of external evaluation suggests that school leaders may be using the information provided by the documents produced by the General Inspectorate for Education and Science and the reports from the schools that have already been evaluated externally to inform their own practices.

The observed transition from identifying 'systematicity and scope' as an area for improvement to its progressive affirmation as a structural strength does not necessarily imply a linear evolution in the underlying data literacy of educators, a construct not directly operationalised within this research framework. Rather, within the boundaries of the text analysed, this shift indicates that private and vocational schools are increasingly successful in embedding evaluative requirements

into formal organizational routines that survive staff turnover. As institutions accumulate experience with the third cycle of external evaluation, these administrative regularities seem to become more resilient, structuring the bureaucratic presentation of data within the school's public profile. By adopting the concept of 'reinforcement data use' (Aderet-German and Ben-Peretz, 2020), these textually documented strengths may be interpreted not merely as celebratory markers, but as an index of how inspectorate expectations become internalised as a strategic mandate for continuity, serving to stabilise specific institutional routines against organisational instability.

The significant disparities found between vocational and private schools suggest the impact of quality management models. Vocational schools, by implementing the EQAVET model, benefit from a now well-established experience with a structural framework that correlates positively with the classifications obtained ($r_s = .353$, $p < .01$). This result supports the thesis that models based on the PDCA (Plan-Do-Check-Act) cycle may foster shared reflexivity. However, a nuanced reading of the data reveals that while vocational schools excel in procedural and global metrics, private schools frequently demonstrate a higher focus on the 'centrality of teaching and learning processes'. This suggests that private schools, despite showing less systematicity in their often 'self-designed' models, may maintain a more 'tailor-made' and emancipated pedagogical focus, less constrained by the standardized indicators of the EQAVET framework. In contrast, private schools show a greater dependence on self-designed models, which, when less robust, show a negative correlation with how SE domain is assessed by the external evaluation teams ($r_s = -.318$, $p < .05$). This evidence engages with the concerns of Sousa and Pacheco (2019) regarding the risk of 'hybrid regulation' that can leave schools navigating between state impositions and contextual needs without a clear methodological compass.

The emergence of the 'centrality of teaching and learning processes' as a dominant theme, which appears simultaneously as a strength and an area for improvement, is one of the most analytically significant findings of this study. According to Figueiredo (2023), for self-evaluation to operate authentically as a 'For Within' process, it must serve as a self-regulation instrument that actively informs and shapes strategic decision-making. The increasing frequency of textual references to this curricular and pedagogical focus suggests an institutional effort by private and vocational schools to mitigate the 'translation gap' conceptualised by Visscher (2021), wherein educational institutions frequently struggle to convert evaluative data into concrete classroom interventions. It also suggests that this is an aspect which is particularly valued and given centrality in the external evaluation model, and by the external evaluation teams themselves, shedding light into how the prescribed guidelines for evaluation are applied in practice by external evaluation teams. When situated within the broader landscape of contemporary educational leadership scholarship, this dualism directly engages with the foundational tenets of instructional leadership and the leadership for learning paradigm (Hallinger, 2011; Robinson et al., 2008). This paradigm establishes that shifting the primary axis of school governance from administrative, managerialist machinery into the instructional core of the classroom represents the most complex, hard-to-surface, and contested transition in any school system. Rather than signifying an isolated procedural failure, the persistence of this theme as an area for improvement, particularly within the highly rationalised framework of vocational institutions, offers a novel contribution to ongoing international discussions on how leadership mediates accountability. It suggests that as schools deepen their structural engagement with quality frameworks, their evaluative lens becomes more granular and demanding. Consequently, school leaders and evaluation teams internalise these criteria not as a static compliance ritual, but as a mechanism that renders the instructional core more transparent, sharpening their capacity to recognise that the curricular and pedagogical core remains the most resilient to transformation. The simultaneous presence of this domain across both evaluative poles is therefore an indicator of evaluative maturation, reflecting known systemic tensions where prioritizing teaching and learning inevitably exposes the deepest institutional vulnerabilities.

Ultimately, this comparative advantage of vocational schools over private ones in terms of the SE domain scores invites a cautious reading rather than an uncritical endorsement of procedural systematicity. This divergence can be understood as an indication of how highly structured, knowledge-based regulation patterns reward technical compliance over more organic forms of pedagogical reflection and how external evaluation may continue to reinforce this pattern. As suggested by Sousa and Pacheco (2019), a risk persists that intensive alignment with standardized frameworks induces a performative orientation, wherein statistical effectiveness and the production of legible evidence for the external gaze overshadow contextualised, critical reflection on the broader educational project. Consequently, the superior scores achieved by vocational schools regarding SE, through the lenses of external evaluation teams, may manifest a greater institutional dexterity in navigating the discursive demands of accountability and performativity, rather than indicating inherently superior curricular and pedagogical practices. The empirical results thus illuminate the ongoing tension between the managerialist and professional logics in school governance, leaving questions still unanswered which can only be addressed by knowing the contexts directly and listening to the actors involved. Nevertheless, the prevalence of strengths related to 'systematicity,' 'impact,' and 'strategic planning' denotes that these schools are attempting to negotiate a complex equilibrium, using the external mandate to anchor internal regularities, managing to balance the need to produce evidence for the external eye with shared reflections. As advocated by Penninckx et al. (2016), when an evaluative architecture is perceived as possessing rigour and legitimacy, the demands of accountability do not automatically nullify professional autonomy, instead, they can operate as a discursive mechanism that structures a professional learning culture, providing the institutional stability necessary to sustain internal development.

5 Conclusions

The evidence collected in this study suggests that, within the parameters of the analysed inspection reports, external school evaluation has functioned as a significant mechanism for structuring institutional development within the private and vocational types of schools that entered this evaluation process recently. The longitudinal analysis of 47 external evaluation

reports reveals a progressive alignment between external standards and internal administrative routines, characterized by an upward trend and convergence in evaluation scores. This trajectory suggests that the external gaze, when perceived as legitimate, provides an external baseline that validates and structures internal practices, reinforcing the formalisation of self-evaluation routines across successive inspection cycles, as leaders for learning progressively use this information to steer their practices and develop capacity.

Within this dynamic, schools' self-evaluation emerges with the potential to be more than merely an accountability exercise, but as a fundamental instrument of validation and security for school leadership. In line with research on Portuguese school leaders (Figueiredo et al., 2024), self-evaluation results are valued as a source of contextualized information used to legitimate strategic decisions and provide sustainability to institutional management. One of the contributions of this work is the application of the concept of 'reinforcement data use' to the analysis of external evaluation of schools' reports. By moving beyond mere weakness remediation, through the focus on identifying the strengths valued by external evaluation teams regarding self-evaluation processes, the findings suggest a trend where the identification of institutional strengths can serve as a strategic mandate for continuity. This approach potentially reinforces school ethos and encourages a more proactive attitude toward evaluative processes, contributing to the stabilization of successful routines against institutional instability.

However, the disparity found between vocational and private establishments offers a potential insight into the impact of quality management models. While vocational schools benefit from the procedural rigor of the EQAVET model, which shows a significant positive correlation with SE domain success, there is an inherent risk of this systematicity becoming a formalised exercise of compliance. Nevertheless, the longitudinal data seem to partially contradict this concern, as reflexivity and impact progressively emerge as recognised strengths in more recent reports, suggesting that the standardization of quality criteria may induce a deeper engagement with institutional evidence. Conversely, while private schools maintain a more customized focus on pedagogical and curricular dimensions through self-designed models, they present less systematic processes, which frequently translate into increased difficulty structurally demonstrating the impact of their practices to an external observer.

In both types of schools, the critical challenge remains to bring these processes closer to the core of the school: pedagogical and curricular issues, which again highlights the relevance of the role of leaders for learning in translating any evaluation, internal or external, into effective change and improvement. Despite the growing maturity of monitoring practices, the centrality of teaching processes, curricular development, and learning remain the most complex dimensions to permeate and transform through self-evaluation. To effectively bridge the identified translation gap (Visscher, 2021), future self-evaluation practices must shift from producing administrative evidence for external accountability towards working authentically for internal school improvement. Reconciling managerial systematicity with shared curricular and pedagogical reflection constitutes a decisive condition for school development, confirming that the persistence of this challenge is a reflection of wider international tensions in the leadership for learning paradigm, where transforming the instructional core remains the most complex milestone of educational reform.

Regarding limitations, this study acknowledges its exclusive focus on the perspective of inspection reports. It is imperative that future research incorporates other voices from the school community to mitigate the 'translation gap' between statistical data and the concrete reality of the classroom. For example, the fact that the identification of self-evaluation models used by the schools is based on what is reported by the external evaluation teams may mean that that information is not complete. Furthermore, the fact that this paper is exclusively based on the external evaluation reports makes it impossible to leave out the possibility of alternative interpretations of the data. For example, private and vocational schools have different social mandates, institutional identities and relationships with stakeholders, which may also have played roles in the ways these institutions approach self-evaluation. This is an important concern that will guide subsequent research on how these two late-entry cases have navigated self-evaluation and external evaluation of schools.

Data availability statement

Publicly available datasets were analysed in this study. This data can be found here: https://igec-aee.site/index.php.

Author contributions

FS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. JS: Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MÁ: Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This manuscript is supported by national funds through the FCT—Fundação para a Ciência e a Tecnologia, I.P., in the scope of project through the project UID/04372/2025, https://doi.org/10.54499/UID/04372/2025 and under the Pluriannual Funding of CIEd—Research Centre on Education of the University of Minho (UID/01661/2025). The funders had no role in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Conflict of interest

JS and FS currently collaborate with the General Inspectorate of Education and Science as the elements teams involved in the

external evaluation of schools. No financial incentives affected the research process.

The remaining author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author FS declared that they were an editorial board member of Frontiers at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aderet-German, T., and Ben-Peretz, M. (2020). Using data on school strengths and weaknesses for school improvement. *Stud. Educ. Eval.* 64, 100831. doi: 10.1016/j.stueduc.2019.100831

Alkin, M. C., and King, J. A. (2016). The historical development of evaluation use. *Am. J. Eval.* 37 (4), 568–579. doi: 10.1177/1098214016665164

Alkin, M. C., and Taut, S. M. (2002). Unbundling evaluation use. *Stud. Educ. Eval.* 29 (1), 1–12. doi: 10.1016/S0191-491X(03)90001-0

Amado, J., Costa, A. P., and Crusoé, N. (2014). "A técnica de análise de conteúdo," in *Manual de Investigação Qualitativa em Educação*, Coord. J. Amado (Coimbra: Coimbra University Press), 301–352. doi: 10.14195/978-989-26-0879-2

Amorim, C., Bidarra, M. G., and Barreira, C. (2021). Autoavaliação das escolas públicas e das escolas particulares e cooperativas de ensino artístico: perspetiva dos professores. *Revista Investigar em Educação. 2* (13), 61–77. doi: 10.34626/214t9j28

Argyris, C., and Schön, D. (1978). *Organizational Learning: A Theory of Action Perspective*. Reading, MA: Addison-Wesley.

Bardin, L. (2016). *Análise de Conteúdo*. Lisboa: Edições 70.

Bingham, A. J. (2023). From data management to actionable findings: a five-phase process of qualitative data analysis. *Int. J. Qual. Methods. 22*, 1–11. doi: 10.1177/16094069231183620

Bolívar, A. (2000). *Los Centros Educativos Como Organizaciones que Aprenden: Promesa y Realidades*. Madrid: La Muralla.

Brown, M., Gardezi, S., Blanco, L. D. C., Simeonova, R., Parvanova, Y., McNamara, G., et al. (2021). School self-evaluation an international or country specific imperative for school improvement? *Int. J. Educ. Res. Open* 2, 100063. doi: 10.1016/j.ijedro.2021.100063

Brown, M., McNamara, G., Cinkir, S., Fadar, J., Figueiredo, M., Vanhoof, J., et al. (2020). Exploring parent and student engagement in school self-evaluation in four European countries. *Eur. Educ. Res. J.* 20 (2), 159–175. doi: 10.1177/1474904120961203

Correia, A., Fialho, I., and Sá, V. (2015). A autoavaliação de escolas: tensões e sentidos da ação. *REIPE*. 10, 100–105. doi: 10.17979/REIPE.2015.0.10.535

Costa, L., Henriques, S., and Abelha, M. (2025). Impacto da avaliação externa de escolas na melhoria do processo de autoavaliação—estudo de caso. *Rev. Meta Aval.* 17 (56) 339. doi: 10.22347/2175-2753v17i56.5020

Costa, M. C. (2024). *Ensino secundário profissional: motivações e expectativas—um estudo com alunos de escolas profissionais privadas em Portugal*. (Tese de doutoramento). Lisbon: Universidade Nova de Lisboa. Repositório da UNL. Available online at: http://hdl.handle.net/10362/167228

Cousins, J. B., and Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Rev. Educ. Res.* 56 (3), 331–364. doi: 10.3102/00346543056003331

Ehren, M. C. M., Altrichter, H., McNamara, G., and O'Hara, J. (2013). Impact of school inspections on improvement of schools - Describing assumptions on causal mechanisms in six European countries. *Educ. Assess. Eval. Account*. 25 (1), 3–43. doi: 10.1007/s11092-012-9156-4

Ehren, M. C. M., Gustafsson, J. E., Altrichter, H., Skedsmo, G., Kemethofer, D., and Huber, S. G. (2015). Comparing effects and Side effects of different school inspection systems across Europe. *Comp. Educ.*. 51 (3), 375–400. doi: 10.1080/03050068.2015.1045769

Ehren, M. C. M., and Shackleton, N. (2016). Mechanisms of change in Dutch inspected schools: comparing schools in different inspection treatments. *Br. J. Educ. Stud.*. 64, 185–213. doi: 10.1080/00071005.2015.1019413

European Institute of Public Administration [EIPA]. (2020). *CAF 2020: Common Assessment Framework. The European Model for Improving Public Organisations Through Self-Assessment*. Maastricht: European Institute of Public Administration. Available online at: https://www.eipa.eu/wp-content/uploads/2020/11/CAF2020.pdf (Accessed February 12, 2026).

European Quality Assurance Reference Framework for Vocational Education and Training [EQAVET]. (2025). EQAVET European Reference Framework for quality assurance in vocational education and training: Improving Quality Assurance in Vocational Education and Training across Europe through common principles of quality management. Available online at: https://employment-social-affairs.ec.europa.eu/policies-and-activities/skills-and-qualifications/working-together/eqavet-european-quality-assurance-vocational-education-and-training/about-eqavet_en (Accessed February 12, 2026).

Eurydice, A., De Coster, I., Birch, P., Czort, S., Delhaxhe, A., and Colclough, O. (2015). *Assuring Quality in Education: Policies and Approaches to School Evaluation in Europe*. Luxembourg: Publications Office of the European Union. Available online at: https://data.europa.eu/doi/10.2797/678

Faubert, V. (2009). *School Evaluation: Current Practices in OECD Countries and a Literature Review. OECD Education Working Papers 42*. Paris: OECD Publishing. doi: 10.1787/218816547156

Fialho, I., Saragoça, J., Gomes, S., Silvestre, M. J., and Correia, A. P. (2014). "Relatórios internacionais e nacionais. Um olhar sobre a avaliação externa de escolas," in *Avaliação Externa de Escolas: Quadro Téorico-conceptual*, Org. J. A. Pacheco (Porto: Porto Editora), 147–180.

Fialho, I., Sousa, J., and Pacheco, J. A. (2023). "Modelo do 3.° ciclo da avaliação Externa das escolas," in *Avaliação Externa das Escolas—Mecanismos de Mudança nas Escolas e na Inspeção*, eds. I. Fialho, J. Saragoça, S. Gomes, M. J. Silvestre, and A. P. Correia (Vila Nova de Famalicão: Editora Húmus), 15–43.

Figueiredo, C. (2023). Autoavaliação de escolas: o quê? Como? Com quem? E depois? *Ensaio: Aval. Pol. Públ. Educ.* 31 (120) 1–24. doi: 10.1590/S0104-40362023003103545

Figueiredo, C. (2024). External school evaluation feedback and school self-evaluation: what feedback is provided? *Res. Educ. Adm. Leadersh.* 9 (1), 123–164. doi: 10.30828/real.1300512

Figueiredo, C., Sampaio, M., Marinho, P., Dihl Moraes, A., and Fernandes, P. (2024). School self-evaluation in decision making: perspectives from Portuguese school leaders. *Educ. Sci.* 14 (12) 1275. doi: 10.3390/educsci14121275

Fundação Manuel Leão. (2024). *AVES: Programa de Avaliação Externa de Escolas. Referencial Genérico*. Vila Nova de Gaia: Fundação Manuel Leão. Available online at: https://www.fmleao.pt/wp-content/uploads/2024/05/Referencial_generico_maio_2024_FINAL.pdf

Gardezi, S. (2024). The development and Status of self-evaluation as a component of school inspection and quality assurance: a comparison across four countries. *Heliyon* 10 (21) e39321. doi: 10.1016/j.heliyon.2024.e39321

General Inspectorate of Education and Science [IGEC]. (2019a). *Avaliação Externa das Escolas: Âmbito, Princípios e Objetivos [External Evaluation of Schools: Scope, Principles and Objectives]*. Lisbon: IGEC. Available online at: https://www.igec.mec.pt/upload/AEE3_2018/AEE_3_Amb_princ_objetivos.pdf (Accessed February 12, 2026).

General Inspectorate of Education and Science [IGEC]. (2019b). *Avaliação Externa das Escolas: Metodologia [External Evaluation of Schools: Methodology]*. Lisbon: IGEC. Available online at: https://www.igec.mec.pt/upload/AEE3_2018/AEE_3_Metodologia_I.pdf (Accessed February 12, 2026).

General Inspectorate of Education and Science [IGEC]. (2019c). *Avaliação Externa das Escolas: Escala de Avaliação [External Evaluation of Schools: Classification Scale]*. Lisbon: IGEC. Available online at: https://www.igec.mec.pt/upload/AEE3_2018/AEE_3_Escala_avaliacao.pdf (Accessed February 12, 2026).

General Inspectorate of Education and Science [IGEC]. (2023). *Avaliação Externa das Escolas: Quadro de Referência—terceiro Ciclo [External Evaluation of Schools: Reference Framework—Third Cycle]*. Lisbon: IGEC. Available online at: https://www.igec.mec.pt/upload/AEE3/AEE_QR_2023.pdf (Accessed February 12, 2026).

Hallinger, P. (2011). Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49 (2), 125–142. doi: 10.1108/09578231111116699

Hult, A., and Segerholm, C. (2016). The process of juridification of school inspection in Sweden. *Utbild. Demokr.* 25 (2), 95–118. doi: 10.48059/uod.v25i2.1059

International Organization for Standardization [ISO]. (2015). *ISO 9001:2015 Quality Management Systems—Requirements*. Geneva: International Organization for Standardization.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., and Volkov, B. (2009). Research on evaluation use: a review of the empirical literature from 1986 to 2005. *Am. J. Eval.* 30 (3), 377–410. doi: 10.1177/1098214009341660

Lo, C. (2018). "Going from government to governance," in *Global Encyclopedia of Public Administration, Public Policy, and Governance*, ed. A. Farazmand (Cham: Springer). pp. 1–5. doi: 10.1007/978-3-319-31816-5_3282-1

Luís, M. S. R. (2022). *Avaliação externa de escolas: estudo a partir das perceções e expectativas dos diretores de um estabelecimento de ensino profissional privado.* (Dissertação de Mestrado em Administração e Gestão Educacional). Lisboa: Universidade Aberta. Available online at: http://hdl.handle.net/10400.2/11626

Marôco, J. (2018). *Análise Estatística com o SPSS Statistics*. 7th Edn. Pêro Pinheiro: ReportNumber.

Martins, N., Palmeirão, C., and Alves, J. M. (2025). Innovation in Portuguese schools: what is the grammar of its conceptualization? *Front. Educ.* 10, 1–14. doi: 10.3389/feduc.2025.1476880

Miranda, J., Seabra, F., and Pacheco, J. A. (2023). Avaliação externa das escolas, regulação por pares, trabalho colaborativo e qualidade educativa: qual a relação? *Práxis Educ.* 18, 1–18. doi: 10.5212/PraxEduc.v.18.21863.065

Morgado, J. C., Sousa, J., and Pacheco, J. A. (2020). Educational transformations in pandemic times: from social confinement to curriculum isolation. *Práxis Educ.* 15, 1–10. doi: 10.5212/PraxEduc.v.15.16197.062

Mouraz, A., Fernandes, P., and Leite, C. (2014). Influências da avaliação externa das escolas no desenvolvimento de uma cultura de autoavaliação. *Rev. Port. Invest. Educ.* 14, 67–96. doi: 10.34632/investigacaoeducacional.2014.3400

Mouraz, A., Leite, C., and Fernandes, P. (2019). Between external influence and autonomy of schools: effects of external evaluation of schools. *Paidéia* 29, 1–11. doi: 10.1590/1982-4327e2922

National Agency for Qualification and Vocational Education and Training [ANQEP]. (2018). Guia para o alinhamento com o Quadro EQAVET. Available online at: https://aedas.edu.pt/wp-content/uploads/2022/11/Guia_Alinhamento_QuadroEQAVET.pdf (Accessed February 12, 2026).

Ólafsdóttir, B., Jónasson, J., Sigurðardóttir, A., and Aspelund, T. (2022). The mechanisms by which external school evaluation in Iceland influences internal evaluation and school professionals' practices. *Nord. J. Stud. Educ. Policy* 8, 209–224. doi: 10.1080/20020317.2022.2076376

Penninckx, M., Vanhoof, J., De Maeyer, S., and Van Petegem, P. (2016). Explaining effects and Side effects of school inspections: a path analysis. *Sch. Eff. Sch. Improv.* 27 (3), 333–347. doi: 10.1080/09243453.2015.1085421

Penninckx, M. (2017). Effects and Side effects of school inspections: a general framework. *Stud. Educ. Eval.* 52, 1–11. doi: 10.1016/j.stueduc.2016.06.006

Penninckx, M., Vanhoof, J., De Maeyer, S., and Van Petegem, P. (2016). Explaining effects and Side effects of school inspections: a path analysis. *Sch. Eff. Sch. Improv.* 27 (3), 333–347. doi: 10.1080/09243453.2015.1085421

Portugal. (2002). *Law No. 31*/2002. Diário da República n.° 294/2002, Série I-A de 2002-12-20. Available online at: https://diariodarepublica.pt/dr/detalhe/lei/31-2002-405486

Portugal. (2014). *Law-decree No. 92*/2014. Diário da República. Diário da República n.° 117/2014, Série I de 2014-06-20. Available online at: https://diariodarepublica.pt/dr/detalhe/decreto-lei/92-2014-25676935

Portugal. (2018a). *Law-decree No. 54*/2018. Diário da República: I série, n.° 129. Available online at: https://www.dge.mec.pt/sites/default/files/EEspecial/dl_54_2018.pdf

Portugal. (2018b). *Law-decree No. 55*/2018. Diário da República: I série, n.° 129. Available online at: https://files.dre.pt/1s/2018/07/12900/0292802943.pdf

Robinson, V. M., Lloyd, C. A., and Rowe, K. J. (2008). The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674. doi: 10.1177/0013161X08321509

Ruiz-Carrillo, J. I. C., and Fernández-Ortiz, R. (2005). Theoretical foundation of the EFQM model: the resource-based view. *Total Qual. Manag. Bus. Excell.* 16 (1), 31–55. doi: 10.1080/1478336042000309857

Schildkamp, K., and Datnow, A. (2022). When data teams struggle: learning from less successful data use efforts. *Leadersh. Policy Sch.* 21 (2), 147–166. doi: 10.1080/15700763.2020.1734630

Schildkamp, K., Visscher, A., and Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *Sch. Eff. Sch. Improv.* 20, 69–88. doi: 10.1080/09243450802605506

Seabra, F., Henriques, S., Mouraz, A., Abelha, M., and Tavares, A. (2022). Schools' strengths and areas for improvement: perspectives from external evaluation reports. *Front. Educ.* 7, 14. doi: 10.3389/feduc.2022.868481

Shulha, L. M., and Cousins, J. B. (1997). Evaluation use: theory, research, and practice since 1986. *Eval. Pract.* 18, 195–208. doi: 10.1177/109821409701800302

Sousa, J. (2020). External evaluation of school in Portugal: essay on rationality and meaning. *Rev. Educ. Polít. Soc.* 5 (2) 155–170. doi: 10.15366/reps2020.5.2.008

Sousa, A. B. P., and Terrasêca, M. M. A. (2015). Processos de autoavaliação de escolas ou avaliação interna do sistema? *Arquivos Analíticos de Políticas Educativas*. 23 (115). doi: 10.14507/epaa.v23.2085

Sousa, J., and Pacheco, J. A. (2018). Teachers' perceptions on the external evaluation of schools in Portugal. *Roteiro.* 43 (1) 237–258. doi: 10.18593/r.v43i1.15852

Sousa, J., and Pacheco, J. A. (2019). External evaluation of schools: policy rationale of institutional evaluation. *Estud. Aval. Educ.* 30 (74) 536–556. doi: 10.18222/eae.v30i74.5860

Stufflebeam, D. L., and Zhang, G. (2017). *The CIPP Evaluation Model. How to Evaluate for Improvement and Accountability*. New York City: The Guilford Press.

Vanari, K., Reinaru, D., and Eisenschmidt, E. (2025). Data use for school self-evaluation—the process, factors, and leadership practices in Estonian schools. *Stud. Educ. Eval.* 85, 101465. doi: 10.1016/j.stueduc.2025.101465

Vanhoof, J., Buvens, R., and Van Petegem, P. (2025). School inspections as a lever for educational quality: case studies on processes and effects. *Res. Pap. Educ.* 40 (5), 663–683. doi: 10.1080/02671522.2025.2464740

Visscher, A. J. (2021). On the value of data-based decision making in education: the evidence from six intervention studies. *Stud. Educ. Eval.* 69, 100899. doi: 10.1016/j.stueduc.2020.100899

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Adm. Sci. Q* 21 (1), 1–19. doi: 10.2307/2391875