

Volume 8, Issue 1, April 2026 p. 162-183

Istanbul / Türkiye

Article Information

Article Type: Research Article

This article has been checked for similarity.

This is an open access article under the CC-BY-NC licence

Article History:

Received

10/01/2026

Received in

revised form

11/03/2026

Accepted

30/04/2026



EXPLORING PRIMARY TEACHERS' LIVED EXPERIENCES OF THE NEW EVALUATION FRAMEWORK IN GREECE: A QUALITATIVE INQUIRY

Nikoula Floka¹

Panagiotis Mitropoulos²

Ioannis Mitropoulos³

Abstract

The issue of teacher evaluation stands as a central concern in contemporary Greek educational discourse. Following several decades of inactivity – during which various laws were enacted, multiple Presidential Decrees were issued, and systematic opposition and resistance from trade unions were observed, leading to the repeal or non-implementation of certain regulations – the Greek educational system is currently entering a new era. In this context, teacher evaluation is no longer a theoretical prospect but a concrete reality. This qualitative study explores teachers' views, perceptions and experiences in the new evaluation process, identifying challenges, benefits, and its impact on their professional lives. The findings indicate that while teachers acknowledge evaluation as a tool for improvement, they criticize its bureaucratic burden and superficial execution while identifying it with a stressful, performative process lacking objectivity and fairness.

Keywords: Teachers' evaluation in Greece, Subjectivity, Performativity, Evaluation culture.

¹ Dr., Experimental Primary School, University of Patras, Greece. ORCID: <https://orcid.org/0009-0007-8100-9072>. Email: nifloka@sch.gr.

² Assoc. Prof, University of Patras, Greece. ORCID: <https://orcid.org/0000-0002-6450-2670>. Email: pmitro@upatras.gr.

³ Prof., University of Patras, Greece. ORCID: <https://orcid.org/0000-0002-1522-1231>. Email: imitro@upatras.gr.

Introduction

International research highlights that teacher evaluation is a fundamental component of effective educational systems (OECD, 2020), offering opportunities to enhance teacher performance and the quality of education provided (Eurydice, 2021). Findings from the 2018 Teaching and Learning International Survey (TALIS), which included educators and school leaders from 48 countries, indicate that teacher evaluation is a common global practice, with only 7% of teachers reporting no evaluation experience. However, evaluation systems vary across countries in terms of purpose, criteria, implementing bodies, frequency, methods, and the use of results. For instance, teachers in Sweden and Switzerland are evaluated annually, while in Serbia, the process occurs every six years. In Hungary and Slovenia, evaluations inform promotion decisions, whereas in Romania, they are linked to salary increases. The responsibility for setting evaluation regulations lies with local communities in Spain, while in England and Germany, it falls under state authorities. In Luxembourg and Malta, school principals serve as evaluators, while in Belgium and Albania, external evaluators are also involved. In Japan and the UAE, evaluations consider various forms of evidence, such as classroom observations, teacher knowledge assessments, and student performance, whereas in Sweden, evaluations involve only an interview with the evaluator. In Greece, teacher evaluation has been a subject of negotiation, debate, and controversy among political leaders, education ministers, and unions for over half a century, often described as the 'Curse of Sisyphus' of Greek education (Kassotakis, 2016). Until recently, Greece was among the few countries worldwide where teachers were not evaluated for their work. This long-standing status quo changed with the introduction of Law 4823/2021, marking a new era for the Greek educational system. Accordingly, the school teachers are evaluated by both the educational counselor and the headmaster/headmistress of each school. Each one of the evaluators observes the evaluatee two times for a whole teaching hour. Each observation occurs at a predetermined day and time which is set after mutual agreement between the two parties. This is followed by a reflective discussion between the evaluator and the teacher. The final evaluation is compiled drawing on both the lesson observation and the discussion afterwards along with the

evaluee's self-evaluation data which he/she uploads to the relevant official platform.

Literature Review

This literature review examines key aspects of teacher evaluation, focusing on the teacher's visibility while being evaluated, the evaluator's qualifications and subjectivity, and the existent evaluation culture. Grounded in Foucault's concept of disciplinary power and empirical research, it explores the power dynamics of classroom observations and assessment practices. Additionally, it analyses evaluators' expertise and biases, assessing their impact on fairness and effectiveness. Finally, it considers the development of an evaluation culture, addressing both resistance and acceptance. By integrating these perspectives, this review enhances understanding of the complexities and implications of teacher evaluation in contemporary education.

The Visibility of Educators During Their Evaluation

The evaluation process, institutionalized through legislation with penalties for non-compliance, carries the authoritative power inherent in all mechanisms of observation and control. In this form, it transforms the evaluated individuals into objects of observation and of information collection, utilizing the visibility that it mandates and enforces. The visibility of those being evaluated emerges from a dialogical synthesis of the data gathered through various evaluative tools, among which classroom observation occupy a prominent role. This process integrates and embodies Michel Foucault's concept of power, specifically his notion of disciplinary power, as exercised by the state—in this case, the Greek government and the legislative framework governing teacher evaluation.

O'Leary (2016) emphasizes that teaching observation can serve as a tool for exercising power and control over what a teacher does and how his/her professional value is assessed and subsequently judged. The evaluator's role is analogous to that of the inspector in Bentham's Panopticon, capable of understanding and grasping institutional operations even when appearing suddenly and without warning. In the context of teacher evaluation, the 'Panopticon' is the classroom in which the teacher teaches, with everything becoming 'visible' to the evaluators from the moment they step into the room. The arrangement of desks, the visual stimuli adorning

the walls, and even the teacher's position and movement within the classroom provides valuable information about his/her teaching methods and techniques.

The teacher under evaluation becomes the primary subject of observation and knowingly enters a 'field of visibility' during the evaluative process. Simultaneously, the restrictions imposed by the prevailing power dynamics, which in this context are the criteria defining the characteristics his/her teaching must exhibit to achieve the highest possible evaluation score, 'play spontaneously upon himself' (Foucault, 1977, p. 202). Under such conditions, he/she internalizes the expectations regarding his/her choice of teaching methods and organizes his/her work accordingly, suppressing potential variations in daily educational routines, practices, and pedagogical approaches. The power wielded by discourse and the evaluation criteria's guidelines overshadows and overrides established teaching practices. Decoding the messages disseminated by legislative texts transforms recipients into both enforcers and subjects of their own compliance with the prescribed standards.

Thus, classroom observation functions as a highly effective mechanism, capable of penetrating someone's behaviour altering it according to the expectations and impositions of power structures (Foucault, 1977). This force intensifies feelings of pressure and oppression in the evaluated individuals and 'acts even before...mistakes...have been committed' as its power lies in the fact that 'it never intervenes, it is exercised spontaneously and without a noise... ..without any physical instrument...it acts directly on individuals; it gives power of mind over mind' (Foucault, 1977, p. 206).

The disciplinary power exerted through this process leads to the perception that evaluators can reliably assess overall teacher performance based on a single hour of classroom observation (Roberts, 2019). However, O'Leary (2016) cautions that such conditions may result in superficial judgments, deviating significantly from the broader purpose of the evaluative process. The entire process becomes 'a form of power...which categorizes the individual...imposes a law of truth on him which he must recognize and which others have to recognize in him' (Foucault, 1982, p. 781). The 'law of truth' referenced by Foucault can be equated with the evaluation result, which the teacher must accept as 'a form of power that makes individuals

subjects...[in the sense that they are subjected] to someone else by control and dependence' (ibid).

In such contexts, the qualifications of the evaluators, their scientific background, teaching specialization, and experience do not mitigate the influence of power dynamics during observation. Evaluators themselves are subject to similar forces, shaped by educational policy and relevant laws. Consequently, both the evaluated teachers and the evaluators are 'normalized to fulfil the accountability and improving standards agenda and self-govern the regimes of practice' (Roberts, 2019, p. 36). This highlights the multiple manifestations of authoritarian power wielded by a centralized, bureaucratically structured state, where individuals adhere to imposed rules and laws, normalizing everyday practices and behaviours. This power is perceived as 'an original right that is given up in the establishment of sovereignty [of the Greek state], and the contract [laws] as the matrix of political force, provide its points of articulation' (Foucault, 1980, p. 90).

Evaluator's Qualifications

According to Scriven (1971, 1972), evaluators must embody the analytical dexterity of anthropologists and sociologists, engaging effectively with diverse groups and interpreting complex social dynamics. Furthermore, they require acute linguistic sensitivity to ensure accurate and nuanced reporting of findings (Scriven, 1979). Stufflebeam (2000, 2003) expands on these competencies, emphasizing the need for systemic thinking, conflict resolution, consensus-building, report writing, and the dissemination of findings. Technical training in data collection, processing, and analysis is indispensable, as it underpins the validity and reliability of evaluative conclusions. Taylor and Tyler (2012a) highlight the evaluator's role in providing targeted feedback to educators. Their research demonstrates that constructive guidance and actionable recommendations lead to significant improvements in students' learning outcomes. They stress that evaluators must possess advanced educational expertise and critical thinking skills, as these are prerequisites for effective feedback and decision-making.

Subjectivity and the Observer's Judgement

The interpretive paradigm acknowledges the inherently subjective nature of evaluation (Greene & McClintock, 1991). All observations are influenced

by the evaluator's theoretical and value-based predispositions. As Scriven (1994) notes, evaluative judgments may sometimes reflect personal preferences or aesthetic biases. This subjectivity is particularly pronounced in qualitative tools, such as classroom observations and interviews, where multiple interpretations may arise. However, it is crucial to distinguish between subjective interpretation and personal prejudice. While the former is integral to the interpretive process, the latter undermines the validity and credibility of the evaluation. Evaluators must adopt a reflective approach, critically analysing their biases and striving for fairness in their judgments (Greene & McClintock, 1991).

Evaluation Culture

The term 'culture' encompasses shared beliefs, perceptions, and behaviours. Erickson (1987) defines culture as an ideological construct shaped through social interactions and historical contexts. Within this framework, evaluation culture reflects collective attitudes and practices regarding evaluation processes in educational settings. The formation of an evaluation culture does not occur in isolation; it evolves through continuous negotiation and interaction within specific social and institutional contexts (Foucault, 1977). This dynamic process is influenced by power structures, which shape individuals' perceptions and actions, ultimately determining the scope of acceptable practices and behaviours (Hinde, 2004).

In educational contexts, evaluation is often perceived negatively, regarded as a threat to established norms and professional autonomy. This perception creates resistance among educators, who may view evaluation as a punitive mechanism rather than a developmental tool (Dounavis & Zmpainos, 2020). Scheerens (2004) emphasizes the importance of addressing these concerns by clarifying the purpose and objectives of evaluation, thereby reducing anxiety and fostering acceptance.

The integration of self-evaluation practices can play a pivotal role in shaping an evaluation culture. Koumentos (2021) introduces the concept of 'self-evaluative ethos', highlighting its dual application for students and educators. By engaging in self-assessment, educators can identify areas for improvement, refine their teaching strategies, and enhance student learning outcomes. This approach aligns with Scheerens' (2004) advocacy for higher-

order reflective practices, such as meta-reflection, which enable deeper analysis of evaluation processes.

Materials and Methods

This study aims to investigate the views and perceptions of primary education teachers regarding the evaluation process under Law 4823/2021. Specifically, it seeks to capture key aspects of each participant's evaluative experience and explore them through four investigating axes:

1. Views and perceptions about the evaluation process and its effectiveness.
2. Problems and difficulties encountered.
3. Perceived benefits
4. Suggested improvements for the evaluation process.

The above axes are deeply interrelated with and highly informed by the following research questions which this study aims to address:

- To what extent did the evaluation process serve its purpose in improving their performance and the quality of their work?
- What challenges and obstacles did they face during the evaluation process?
- What are the benefits they can identify?
- What areas for improvement do they suggest?

The choice of the research methodology is determined by the nature of the research questions and is grounded in specific philosophical beliefs about data acquisition and interpretation (Case & Light, 2011; Knox & Burkard, 2009). This study is rooted in the principles of qualitative research, following the 'interpretive paradigm' which views social reality as a construct of individuals' subjective experiences. The researchers are interested in human experience, its diversity, and complexity, aiming to uncover and to understand it.

The core of this qualitative research lies in the participants' individual perceptions as revealed through their interpretation of their evaluation experience and the personal meanings they attach to it. The investigation seeks to access 'subjective experience... subjective meanings... the

understanding and interpretation of these meanings' (Isari & Pourkos, 2015). The methodological design is based on the belief that reality is divergent, differing in how each individual perceives and describes it.

Six (6) primary education teachers from the Achaia region participated in this research. All of them were evaluated during the 2023-2024 school year, according to the provisions of Law 4823/2021. All participants have more than fifteen (15) years of experience, teaching in various schools both within and outside the region. All participants have postgraduate degrees and regularly participate in professional development seminars aimed at their ongoing professional improvement and growth. The selection of the participants was purposeful, aiming to fulfil the essential criterium of the prior evaluative experience.

Description of Data Collection Tool

The data for investigating the aforementioned research questions were collected using the qualitative tool of semi-structured interviews. Interviews encourage participants to share rich descriptions of phenomena, leaving the interpretation and analysis to the researcher (Warren & Karner, 2005). Additionally, they allow for the discovery or elaboration of information that may be significant for participants but may not have previously been considered relevant by the research team (Gill et al., 2008, p. 292). Moreover, they provide a deeper understanding of social phenomena than would be possible through quantitative methods, such as questionnaires (Silverman, 2000). The data collected contribute to the construction of theoretical knowledge derived from the meanings participants construct through their daily lived experiences (DiCicco-Bloom & Crabtree, 2006).

For the needs of this research, six (6) formal interviews, lasting between 30 to 45 minutes each, were conducted. Participants were informed in advance about the topic of the interview and received the specific research questions via email so they could review them and prepare their answers. This practice gave the participants an idea of what to expect from the interview, which increases the likelihood of honesty and is a fundamental aspect of the informed consent process (Gill et al., 2008, p. 292). However, during the interviews, additional questions arose that the researchers posed to seek further explanations and details regarding a participant's ideas or responses. The interviews were recorded using a special application on a

mobile phone. The recorded audio files were transcribed with the help of Microsoft Word's dictate tool. The transcriptions were sent via email to the participants to review and suggest any changes or additions to points they felt were inaccurately recorded.

Data Analysis Method

The data were analysed using thematic analysis, following the six-phase approach outlined by Braun and Clarke (2006). This method involves identifying and interpreting recurring patterns within the data set in a recursive and cyclical manner.

Phase 1: Familiarization and Transcription – The researchers engaged in repeated readings of the interview transcripts to identify initial patterns and themes.

Phase 2: Coding – Conceptual labels (codes) were systematically assigned to relevant excerpts, with some excerpts receiving multiple codes due to their complexity.

Phase 3: Searching for Themes – The generated codes were organized into broader themes based on their relevance to the research questions.

Phase 4: Reviewing Themes – The themes were refined for coherence and validity, ensuring internal consistency and meaningful distinction from one another. A two-week break was taken before a final review of the themes.

Phase 5: Defining and Naming Themes – Each theme was clearly delineated, analysed in detail, and, when applicable, structured into sub-themes.

Phase 6: Writing Up the Findings – The final analysis was composed to accurately represent participants' experiences and perceptions, integrating excerpts to substantiate the identified themes.

This rigorous approach ensured a thorough and credible analysis of the data, producing a well-structured thematic representation of the research findings.

Trustworthiness of the Study

To enhance the study's reliability and minimize bias, the researchers adhered to four key criteria: credibility, transferability, dependability, and confirmability.

Credibility was ensured through systematic data collection and verification processes. Participants reviewed and validated their interview transcripts, and clarifying questions were used to confirm interpretations. 'Member checks' were conducted to reduce researchers' bias (Lincoln & Guba, 1985). Additionally, the researchers' professional status as school teachers facilitated trust and rapport, enriching the data.

Transferability was supported through detailed descriptions of the research process and participants' experiences, aligning with Bitsch's (2005) emphasis on 'thick description'. Semi-structured interviews with participants who had prior evaluative experience ensured meaningful insights applicable to broader contexts.

Dependability was strengthened by employing a code-recode strategy with a two-week gap (Anney, 2014). This process allowed for reflective analysis, reinforcing consistency between coding iterations and enhancing the clarity of participants' narratives.

Confirmability was maintained by strictly following the research methodology and incorporating direct participant quotes to ensure authenticity. Thus, transparency and adherence to qualitative research standards were safeguarded against researcher bias (Tobin & Begley, 2004).

Accordingly, the researchers put systematic effort to ensure the study's reliability, validity, and ethical soundness maintaining their commitment to serve its purpose with integrity and honesty.

Results

The findings are organized into four thematic axes, each comprising distinct thematic units derived from the coding process. These axes interact dynamically, complementing each other through causal and explanatory relationships. Specifically, themes emerging from participants' perceptions of the challenges they faced during evaluation (Axis 2) and the benefits they gained (Axis 3) help interpret their views and perceptions about the process (Axis 1). Together, these three axes inform the themes of Axis 4, which focuses on participants' recommendations for improving the evaluation process.

*Axis 1: Views and perceptions about the Evaluation Process and its Effectiveness
Necessity but ineffectiveness*

All participants were favourably expressed about the necessity for teachers to be evaluated. As P1 noted, *'at some point, we needed to be evaluated in some way...all the public servants in this country are evaluated'*. Others emphasized the importance of evaluation for improving, with P3 stating that *'the purpose of the evaluation process is to improve performance and the quality of our work'*.

However, participants consistently highlighted the ineffectiveness of the evaluation process in achieving its stated goals, as outlined in Law 4823/2021. P1 raised concerns about the impartiality of evaluations, saying that *'the evaluators are our colleagues... it's difficult to imagine a colleague being deemed unsatisfactory'*. P2 echoed this scepticism, questioning the objectivity of a process where *'98% of teachers were rated as excellent... is that possible? Are we being fooled?'*

Additionally, evident was the disappointment with the evaluation's outcomes with P3 lamenting, *'this process... doesn't work properly, it doesn't clarify, it doesn't differentiate, it doesn't reward those who work hard'*. P4 criticized the evaluation as *'fictional'*, describing it as *'fabricated'*, while P5 questioned its usefulness, stating that *'I can't find anything positive in the evaluation... it should only provide positive feedback to help you grow'*.

Finally, P6 emphasized the lack of constructive feedback, observing that, *'it's not just about getting a good grade... there should be feedback like, 'You did well, but I noticed something here that you have struggled with''*. The consensus among teachers is that while evaluation is deemed necessary, the current process fails to foster meaningful improvement or align with the goals of enhancing education quality.

Axis 2: Problems and Difficulties Encountered

Time and Bureaucratic Burden

Teachers described the evaluation process as highly time-consuming and bureaucratic, diverting attention from actual teaching. P1 noted that *'the negative part was the time needed for preparation... we don't spend time documenting these things in our daily teaching routine'*. The pressure to meet deadlines exacerbated stress, with P2 lamenting that *'all my personal tasks were left behind... this stress with the bureaucracy... the platform where you had to upload documents from your entire professional life'*. P3 highlighted the unrealistic expectations by saying that *'to put so much*

effort for one lesson and to spend three days thinking about how I will present it... this doesn't happen in daily practice'. Others echoed similar frustrations, emphasizing excessive documentation, confusion about requirements, and technical issues with the platform.

Performativity vs. Authenticity

The participants reported feeling pressured to perform rather than teach authentically, aligning with the concept of performativity. P1 observed that 'when it's evaluation time, we're pushed into stricter moulds... we measure time more tightly'. P2 emphasized the need for rigid goal setting by noting that 'you have set goals you must achieve... in evaluation, everything must be clear and measurable'. The pressure to showcase an idealized version of teaching was evident, with P2 further noting that 'you sit down, you work hard, you prepare the perfect lesson... you use tools you've never used before just to show something to the evaluator... not what you'd normally do in class'.

Subjectivity, Bias, and Personal Relationships

Participants contested the objectivity of the evaluation process, citing biases influenced by personal relationships. P1 questioned the impartiality of evaluators emphasizing that 'we were evaluated by colleagues... how objective can they be?'. The role of personal judgments is highlighted by P2 when she stated that 'it's not just the criteria set by the ministry of education... to what degree a teacher meets them depends on the evaluator's personal judgment'. P4 added on the above by noting that 'I don't know how there can be objectivity when I am evaluated by a school counselor with whom I might have had dinner the previous day'. This sentiment was reinforced by P5, who critiques the influence of personal connections reporting that 'there's always the case of someone not getting an 'excellent'... because of his/her connections... there is no impartiality'.

Evaluator's Skills

According to the participants in the present study, the effectiveness of the teacher evaluation largely depends on the competence and the approach used by the evaluators. P1 appreciated the constructive feedback she received noting that 'my counselor showed me the evaluation criteria... provided material... made great suggestions'. Conversely, P5 expressed frustration over the counselor's comments on her teaching by stating that

'after me having 20 years of teaching experience and holding two degrees and two master's, he questions whether I'll play a video first or use the school book...that's ridiculous'. Concerns about inadequate evaluator training were also raised, with P6 reporting that 'they haven't been trained or prepared... the people doing the evaluations should be trained for this task'.

Evaluative Culture

A lack of evaluative culture in Greek schools is seen as a fundamental issue. P1 highlighted the possibility of reluctance on behalf of the evaluatee to express his/her concerns due to fear of repercussions, 'the headmaster may keep in his memory my disagreement with him and then he may do something to get back to me when he evaluates me'. P2 noted that lack of evaluation culture nurtures resistance to negative evaluations as growth opportunities, 'we don't accept criticism...we don't have an evaluation culture meaning that we don't view a low mark as a positive think to work with so as to become better'. P3 and P4 argued that personal relationships compromise objectivity, while P6 advocated for clear evaluation goals and training noting that 'for a proper evaluation culture to be developed...a systemic effort is needed...some seminars, relevant training'. Overall, participants emphasize the need for a structured, growth-oriented evaluation system to enhance professional development. In conclusion, teachers perceived the evaluation process as bureaucratic, performative, and subjective. Addressing these issues through evaluator training, standardized criteria, and fostering a constructive evaluative culture is essential for meaningful professional growth.

Axis 3: Perceived Benefits

Feedback, Self-Evaluation, and Professional Growth

All participants emphasized post-observation feedback as the most valuable aspect of the evaluation process. P1 describes that 'the discussion after the observation helped me... the self-assessment questions made me reflect... it helped me realize I underestimate things I do well'. Similarly, P2 highlighted the introspective benefits she experienced by stating that 'it made me critique myself, reflect on my teaching, and also get a professional's perspective about my teaching'.

Other participants echoed the above sentiment. P3 found the feedback 'clear and constructive', while P6 appreciated its impact on her professional development, 'it made me pay more attention to details... like using more visual aids for weaker students'. She also emphasized on the fact that the feedback she received strengthened her self-confidence, 'when they say, "you did this well, but look into this", it builds my self-esteem'.

Evaluation as a Catalyst for Change

Several teachers viewed evaluation as a driver for professional improvement. P2 acknowledged its potential for change but she also stressed personal agency, 'you have to take the feedback positively and use it to improve... it's a choice'. Additionally, she argued for moving beyond outdated teaching methods stressing that 'evaluation could push us to modernize... but this hasn't happened yet'. P3 found that the process has 'activated' teachers, while P4 linked it to continuous professional development, 'it prevents professionals from becoming complacent'.

However, some participants were sceptical about its lasting impact. P5 criticized superficial adaptations made only for evaluation purposes, 'if you really learn and change, you should apply it permanently... but often, the changes don't stick for a long time'. P6 reinforced this, stating that real improvement depends on motivation, 'if you're not ready to change, feedback won't matter'.

Axis 4: Suggested Improvements for the Evaluation Process

Unscheduled Observations

Several participants suggested unscheduled evaluations as a mean to ensure and to enhance authenticity. P1 proposed 'the counselor could drop in unexpectedly and then give us feedback on our teaching'. P4 agreed by saying that 'it would be more genuine and transparent'. While P6 supported the idea, she stressed the need for constructive feedback, 'he should guide me...to tell me about my deficiencies...to tell me what I have to improve'.

Scepticism About the Future of Evaluation

Many participants doubted the long-term success of the evaluation process. P1 argued that 'the evaluators can't differentiate teachers—everyone is 'excellent'... what kind of evaluation is this?' P2 remained cautiously optimistic by noting that 'with clearer guidelines, evaluation will

progress... but if all teachers are rated 'excellent,' nothing will change'. Concerns also extended to political challenges. P4 stated that 'the state is not ready to work with teacher who didn't do well during their evaluation...they cannot be fired... it's more about compliance than improving education'. P6 concluded that 'it's just a well-organized show... what's the real plan to help teachers improve?'

Conclusion

The participants in this study recognize the necessity of the evaluation process. However, they consider its implementation problematic, as it often becomes a formalistic and procedural exercise that ultimately fails to achieve its primary objective—namely, the improvement of the quality of Greek education. Concerns regarding the evaluative process's effectiveness as a tool for professional development have also been expressed by participants in the studies of Strikas (2018), Collins (2004), Flores (2012), Krana & Vryonidis (2017), Al Maktoum & Al Kaabi (2024) and Tzitzika & Stavakou (2024). Conteras et al. (2024) provide their perspective on the matter by arguing that 'if an assessment is not used for the purposes it was designed for, its validity is compromised' (p. 4).

Repeated concerns about the lack of objectivity and impartiality emerged from the collected data, with many participants attributing these issues to personal relationships between evaluators and evaluatees. As a result, the objectivity of the process is called into question, as evaluators often struggle to maintain professional detachment. This perceived bias generates frustration, dissatisfaction, and disappointment among dedicated and hardworking educators who expect the evaluation process to introduce mechanisms that differentiate and reward their efforts in contrast to colleagues who may not demonstrate the same level of conscientiousness and commitment.

These findings align with the research of Weisberg et al. (2009), who identified this phenomenon as the "widget effect," emphasizing that evaluation systems' inability to recognize and appropriately respond to varying levels of teacher effectiveness not only causes dissatisfaction but are also perceived as demeaning by those being evaluated. Similar concerns regarding the objectivity of evaluators' judgments have also been expressed

in studies by Kreki (2012), Geladaki & Trimi-Kyrou (2014) and Strikas (2018).

Participants also attributed evaluator bias to inadequate training and specialization in skills essential for conducting objective and impartial assessments. Concerns regarding evaluators' qualifications and experience have been similarly highlighted in the studies of Kosma (2008), Larsen (2009), Weisberg et al. (2009), Krekis (2012), Flores (2012), Petousi (2013), Goula (2020), Rallis (2020), and Adamou (2023). Additionally, issues related to fairness and impartiality in the evaluation process have been raised by participants in studies conducted by Larsen (2009), Defarana (2014), Miari (2018), and Politidou (2021) and Tzitzika & Stavakou (2024).

For the participants in the present study, the evaluation process represents a significant bureaucratic burden that demands considerable preparation time, diverting them from their daily teaching responsibilities. The process involves various time-consuming administrative tasks, such as document submission and data entry, which shift the focus away from substantive instructional improvement. Similar concerns have been expressed by Portuguese teachers in the study conducted by Flores (2012), Canadian educators in Larsen's (2009) research and Greek teachers in Tzitzika's & Stavakou's qualitative research study (2024).

Teachers experience particularly high levels of stress during the preparation and implementation of their evaluations, which negatively affects their ability to effectively manage their daily teaching responsibilities. The intense focus on preparing observation lessons—sometimes even involving pre-preparation of students—leads teachers to prioritize evaluation success over other instructional goals. Similar findings have been reported by Larsen (2009) and Ball (2003).

A particularly noteworthy concern is the preparation of idealized lessons designed exclusively for the scrutiny of evaluators, which significantly deviates from the teachers' actual daily instructional practices. This phenomenon reflects a performativity-based approach in which lessons are carefully structured to align with pre-established evaluation criteria and maximize scores, rather than authentically representing everyday teaching. The process resembles a theatrical performance, where teachers and students assume roles to satisfy evaluators, who function as an audience assessing an educational scenario that is ultimately artificial and

disconnected from real classroom dynamics. These findings are consistent with observations made by Ball (2003), Collins (2004), Larsen (2009), Edgington (2013), Bradford & Braaten (2018), and Roberts (2019).

Despite its shortcomings, one aspect of the evaluation process that is highly valued by participants is the post-observation feedback and self-evaluation component. Teachers appreciate the opportunity to reflect on their practices, identify weaknesses, and recognize areas for growth. Furthermore, positive feedback and constructive advice from evaluators enhance their self-confidence, supporting their professional development and continuous improvement. Conteras et al. (2024) use the term 'actionable feedback' which they define as 'feedback that is specific, timely, useful, effective and makes sense to the recipient' arguing for its fundamental role 'for an evaluation programme to achieve its desired outcomes in terms of professional development and learning on the part of teachers' (p.19). The role of evaluation as a tool for self-improvement has also been highlighted by educators in the study conducted by Svarna (2022). Conversely, the evaluation process has been heavily criticized by participants in the studies of Collins (2004), Weisberg et al. (2009), Flores (2012), and Al Maktoum & Al Kaabi (2024) for its superficial and perfunctory nature.

A widespread issue identified in this study is the absence of a well-established evaluation culture, which serves as a major obstacle to utilizing the process as a meaningful tool for improvement and professional growth. As a result, teachers often reject negative feedback, which could otherwise serve as a catalyst for change, innovation, and progress. This resistance effectively reduces the entire process to a bureaucratic obligation carried out merely to comply with legal requirements. Consequently, evaluation remains a theoretical concept with little practical implementation. The necessity of cultivating an evaluation culture among educators to ensure the process's meaningful and effective application has also been emphasized by Svarna (2022) and Tzanidou (2024).

Participants proposed several improvements to enhance the evaluation process. The primary recommendation was to reduce bureaucratic demands, which are perceived as excessively time-consuming and administratively burdensome. Additionally, there was a widespread call for ensuring objectivity by introducing external evaluators. Another key suggestion involved the implementation of unannounced classroom

observations to mitigate the performativity-driven approach, in which teachers prepare idealized lessons solely for evaluation purposes rather than as part of their everyday instructional practice.

The conclusions of this study pertain exclusively to the six participants. However, its findings can be transferred to future research involving different respondents. This research initiative represents an initial attempt to explore the experiences, emerging view and perceptions of evaluated teachers following their assessment under Law 4823/2021. It serves as a starting point for further investigation into a process being implemented in the Greek educational landscape for the first time. Therefore, future studies are recommended to include a larger sample of teachers from a variety of specializations, representing both primary and secondary levels of the Greek education system.

References

- Adamou, S. (2023). *Investigation of the Perceptions of Newly Appointed Teachers and Experienced Teachers in Primary Education Regarding Their Educational Evaluation*. Master's Thesis, University of Peloponnese.
- Al Maktoum, S. B., & Al Kaabi, A. M. (2024). Exploring teachers' experiences within the teacher evaluation process: A qualitative multi-case study. *Cogent Education*, 11(1), 1-20.
- Anney, V. N. (2014). Ensuring the quality of the findings of qualitative research: Looking at trustworthiness criteria. *Journal of Emerging Trends in Educational Research*, 5(2), 272-281.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of education policy*, 18(2), 215-228.
- Bitsch, V. (2005). Qualitative research: A grounded theory example and evaluation criteria. *Journal of Agribusiness*, 23(1), 75-91.
- Bradford, C., & Braaten, M. (2018). Teacher evaluation and the demoralization of teachers. *Teaching and Teacher Education*, 75, 49-59.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Case, J. M., & Light, G. (2011). Emerging research methodologies in engineering education research. *Journal of Engineering Education*, 100(1), 186-210.

- Collins, A. B. (2004). Teacher performance evaluation: A stressful experience from a private secondary school. *Educational Research*, 46(1), 43-54.
- Contreras, J., Martín, C.S. & Muñoz, S. Uses and consequences in the theory of action of high-stakes teacher evaluation in Chile. *Educational Assessment, Evaluation and Accountability* (2024), 1-23. <https://doi.org/10.1007/s11092-024-09448-3>.
- Defarana, V. (2014). *The New Evaluation Framework: Views of Principals & School Counselors on the Evaluation of Primary Education Teachers and Their Work*. Master's Thesis, University of Thessaly.
- DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical education*, 40(4), 314-321.
- Dounavis, G., & Zmpainos, D. (2020). *Teacher Evaluation: Investigation and Comparison of Teachers' Views in Schools in Greece and the USA*. *Educational Evaluation*, 6, 1-23.
- Edgington, U. (2013). Performativity and affectivity: Lesson observations in England's Further Education colleges. *Management in education*, 27(4), 138-145.
- Erickson, F. (1987). Conceptions of school culture: An overview. *Educational Administration Quarterly*, 23(4), 11-24.
- European Commission/EACEA/Eurydice (2021). *Teachers in Europe: Careers, Development and Well-being*. Eurydice report. Luxembourg: Publications Office of the European Union. https://eurydice.eacea.ec.europa.eu/sites/default/files/teachers_in_europe_2020_chapter_3.pdf.
- Flores, M. A. (2012). The implementation of a new policy on teacher appraisal in Portugal: how do teachers experience it at school?. *Educational Assessment, Evaluation and Accountability*, 24, 351-368.
- Foucault, M. (1977). *Discipline and Punish, The Birth of the Prison*. New York: Vintage Books.
- Foucault, M. (1980). *Power knowledge: selected interviews and other writings, 1972-1977*. New York: Pantheon Books.
- Foucault, M. (1982). The subject and power. *Critical inquiry*, 8(4), 777-795.
- Geladaki, S. & Trimi-Kyrou, K. (2015). Old and New Forms of Evaluation in Greece: How Secondary Teachers Perceived their Recent

- Evaluation. *International Journal of Interdisciplinary Social Sciences: Annual Review*, 9, 43-52.
- Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: interviews and focus groups. *British dental journal*, 204(6), 291-295.
- Goula, I. (2020). *Teachers' and Primary School Principals' Views on the Institution of Teacher Evaluation: The Case of Serres Prefecture*. Master's Thesis, University of Macedonia.
- Greene, J. C., & McClintock, C. (1991). The evolution of evaluation methodology. *Theory into practice*, 30(1), 13-21.
- Hinde, E. R. (2004). School culture and change: An examination of the effects of school culture on the process of change. *Essays in education*, 12(3), 1-12.
- Isari, F., & Pourkos, M. (2015). *Qualitative Research Methodology: Applications in Psychology and Education*. Hellenic Academic Electronic Textbooks and Aids.
- Kassotakis, M. (2016). *Modern International Trends in the Fields of School and Teacher Evaluation and the Case of Greece*. Proceedings of the 1st Scientific Conference of the Hellenic Society for Educational Evaluation: *Evaluation of Schools and Teachers: Modern Trends, Dilemmas, and Prospects*, 43-75.
- Knox, S., & Burkard, A. W. (2009). Qualitative research interviews. *Psychotherapy Research*, 19(4-5), 566-575. <https://doi.org/10.1080/10503300802702105>
- Kosma, Th. (2008). *Teacher Evaluation and Their Evaluation by School Counselors*. Master's Thesis, University of Ioannina.
- Koumentos, I. (2021). *Suggestions and Texts for a Multimodal and Creative School*. Diinekis Publications: Athens.
- Krana, D., & Vryonidis, M. (2017). *Secondary Education Teachers' Perceptions of the Necessity of Their Evaluation*. *Education Sciences*, 2017(1), 93-113.
- Krekis, D. (2012). *Teachers' Views on Their Evaluation and the Evaluation of Educational Work. The Case of Primary Education in the Prefecture of Preveza*. In: *Proceedings of the Hellenic Institute of Applied Pedagogy and Education (ΕΛΛ.Ι.Ε.Π.ΕΚ.)*, 6th Panhellenic Conference, October 5-7, 2012, http://www.elliepek.gr/documents/60_synedrio_eisigiseis/31_Krekis.pdf.
- Larsen, M. A. (2009). Stressful, Hectic, Daunting: A Critical Policy Study of the Ontario Teacher Performance Appraisal System. *Canadian Journal*

- of *Educational Administration and Policy*, 95, 1-44.
<https://files.eric.ed.gov/fulltext/EJ863211.pdf>.
- Law 4823/2021. *School Upgrade, Teacher Empowerment, and Other Provisions*.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Miari, E. (2018). *Primary Education Teachers' Views on the Evaluation of Their Educational Work and Their Evaluation by Their Students*. Master's Thesis, Piraeus University of Applied Sciences.
- O'Leary, M. (2016). *Reclaiming lesson observation: supporting excellence in teacher learning*. London: Routledge.
- OECD (2020), "Education Policy Outlook in Greece", *OECD Education Policy Perspectives*, No. 17, OECD Publishing, Paris. <https://doi.org/10.1787/f10b95cf-en>, ανακτήθηκε 06.08.24.
- Petousi, K. (2013). *Investigation of the Views of Primary School Principals in the Ionian Islands Educational Region on Teacher Evaluation*. Master's Thesis, University of Ioannina.
- Politidou, E. (2021). *Teacher Evaluation: Views of Special Education Teachers in Public Secondary Education in Attica*. Master's Thesis, National and Kapodistrian University of Athens (NKUA).
- Rallis, D. (2020). *Educational Work Evaluation: A Case Study of Secondary School Principals and Teachers in Ilia Prefecture*. Master's Thesis, University of Patras.
- Roberts, D. (2019). Lesson observation: what is the problem it's solving? *Research in Teacher Education*, 9(2), 34-38.
- Scheerens, J. (2004). The evaluation culture. *Studies in Educational Evaluation*, 30(2), 105-124.
- Scriven, M. (1971-1972). Monograph Supplement: Curriculum Evaluation: Potentiality and Reality. *Curriculum Theory Network*, 182-192. <https://www.jstor.org/stable/i250376>.
- Scriven, M. (1979). Interview: Michael Scriven: Viewpoints on Education Evaluation. *Educational Evaluation and Policy Analysis*, 1(2), 66-72.
- Scriven, M. (1994). Evaluation as a discipline. *Studies in Educational Evaluation*, 20(1), 147-166.
- Silverman D. (2000). *Doing qualitative research*. London: Sage Publications.
- Strikas, K. (2018). *Evaluation: Yes, but... How will it provide feedback and upgrade education? Scientific Step of the Teacher*, 22, 191-202.

- Stufflebeam, D. L. (2000). Foundational models for century program evaluation. *Evaluation models: Viewpoints on educational and human services evaluation* (32-83). Dordrecht: Kluwer Academic Publishers.
- Stufflebeam, D. L. (2003). The CIPP model for evaluation. In *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 279-317). Dordrecht: Springer Netherlands.
- Svarna, K. (2022). *Teacher Evaluation and Their Professional Development*. Master's Thesis, Hellenic Open University.
- Taylor, E. S., & Tyler, J. H. (2012a). The effect of evaluation on teacher performance. *American Economic Review*, 102(7), 3628-3651.
- Tobin, G. A., & Begley, C. M. (2004). Methodological rigour within a qualitative framework. *Journal of Advanced Nursing*, 48(4), 388-396.
- Tzanidou, Ch. (2024). *Teacher Evaluation*. Master's Thesis, University of West Attica.
- Tzitzika, L., & Stavakou, P. A. (2024). Evaluation of Educational Work in Greece-A Qualitative Research Study. *Journal of Advances in Education and Philosophy*, 8(11), 670-680.
- Warren, C. A., & Karner, T. X. (2005). *Discovering qualitative methods: Field research, interviews, and analysis*. Los Angeles: Roxbury.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. New York: The New Teacher Project. <https://files.eric.ed.gov/fulltext/ED515656.pdf>.