Faculty of Letters Journal of Social Sciences

Revisiting the Phenomenon Of Writing Assessment Literacy: The Delineation Of Its Mediating Dynamics

Sevda Balaman[1,a,*]

[1] Deparment of Translator and Interpretor, Faculty of Letters, Sivas Cumhuriyet University, Sivas, Türkiye
*Corresponding author

Research Article

ABSTRACT

The formation of a sound assessment system is critical for enhancing language learning as a vital constituent of the process and its relevance to the writing instruction is not the exception, as well. In this vein, a number of elements manifesting in the writing evaluation framework have been noted, one of which is undoubtedly teachers. They navigate the assessment procedure as an important stakeholder of the mechanism by assuming various roles, including task designing, test administrating, evaluating, grading, and feedback-providing (Weigle, 2007: 195). Correspondingly, the writing assessment literacy of teachers, maintained by the three-structure schemata, composed of their core knowledge, conceptual beliefs, and current implementations (Crusan, Plakans, & Gebril, 2016: 43), has a salient role in directing this complex system. However, the existing research, which addresses teachers' literacy in the writing assessment specificity, is insufficient in number. To this end, this review study is primarily intended for mapping out the underlying components of teachers' literacy for writing evaluation, in light of the fundamental principles in theory. Additionally, this study aims to succinctly summarize the main findings of the previous studies conducted in this sense by demonstrating an overall picture of the relevant literature with a purpose of contributing to the writing pedagogy and giving directions to the future research.

Keywords: assessment, formative assessment, literacy, summative assessment, writing

Yazma Değerlendirme Okuryazarlığı Olgusunun Yeniden Gözden Geçirilmesi: Aracılık Eden Dinamiklerinin Betimlenmesi

ÖZ

Sağlam bir değerlendirme sisteminin oluşturulması, sürecin hayati bir bileşeni olarak, dil öğrenimini geliştirmek için kritik öneme sahiptir ve bu durumun yazma öğretimiyle ilgisi de bir istisna değildir. Bu bağlamda, yazma değerlendirmesi çerçevesinde ortaya çıkan çok sayıda unsura dikkat çekilmiştir; bunlardan biri şüphesiz öğretmenlerdir. Öğretmenler, aktivite tasarımı, sınav uygulanması, değerlendirilmesi, not verilmesi ve geri bildirim oluşturulması da dahil olmak üzere çeşitli roller üstlenerek mekanizmanın önemli bir paydaşı olarak değerlendirme sürecini yönetmektedir (Weigle, 2007: 195). Buna bağlı olarak, öğretmenlerin bilgileri, kavramsal inançları ve mevcut uygulamalarından oluşan üç-parçalı bir yapı ile sürdürülen, yazma değerlendirme okuryazarlığı (Crusan, Plakans, & Gebril, 2016: 43), bu karmaşık sistemi yönlendirmede etkin bir role sahiptir. Ancak, yazma değerlendirmesi özelinde, öğretmenlerin okuryazarlığını ele alan mevcut araştırmalar sayıca yetersizdir. Bu sebeple, bu inceleme çalışması öncelikle teorideki temel ilkeler ışığında öğretmenlerin yazma değerlendirme okuryazarlığının temelini oluşturan ögeleri belirlemeyi amaçlamaktadır. Ayrıca bu çalışma, yazma eğitimine katkıda bulunmak ve gelecek araştırmalara yön vermek amacıyla, ilgili alan-yazının genel bir resmini ortaya koyarak bu anlamda yapılan mevcut araştırmaların temel bulgularını kısa ve öz bir şekilde özetlemeyi amaçlamaktadır.

Anahtar Kelimeler: değerlendirme, biçimlendirici değerlendirme, okuryazarlık, özetleyici değerlendirme, yazma

[a] sbalaman@cumhuriyet.edu.tr   https://orcid.org/0009-0004-2390-7323

Introduction

The concept of assessment has been acknowledged as an integral component of instruction over the past two decades (Shepard, 2000), with its functionalities in giving proper information about the effectiveness of teaching practices, prescribing the weakness and the strength of learners, and aiding the learning process as a teaching tool, alongside measuring the performance in the respective skill (Lee, 2017). In line with this, there has been a repetitive call from the worldwide spectrum for reconceptualizing the phenomenon of classroom assessment, built on the bases of the assessment for learning (AfL) paradigm (Black & William, 2009; Lee, 2007, 2011; Scarino, 2013), which refers to a mechanism where "teaching, learning, and assessment form a symbiotic relationship" that ends up with a marked modification and improvement of learning and teaching skills (Lee, 2017: 36).

Indeed, this holds true for English-as-a-foreign-language (EFL) or second language (L2) writing milieus (Crusan, 2010; Lee, 2017; Weigle, 2007), given that this skill, reported as an arduous task (Hyland, 2003; Nguyen & Truong, 2021), inevitably necessitates a robust assessment system, whose primary purpose is for enabling learners to gain proper competence for yielding high-quality outputs with long-term monitoring and guidance in composing (Lee, 2011, 2017). The assessment of writing is a multifaceted construct, characterized by a number of foci (Sohrabi, Ghanbari, & Abbasi, 2022; Wang, Lee, & Park, 2020), ranging from the subjective evaluation of what constitutes a well-written text (Crusan, 2010, 2013) to rater-based and task-related issues (Huang, 2009). Inarguably, teachers are the primary navigators of this complicated process with their roles in designing, administering, grading, and communicating the results of the classroom tests in a reliable and valid way (Weigle, 2007: 195). In this sense, their beliefs, perspectives, attitudes, conceptual ideas, and knowledge that make up their assessment literacy, which controls certain parameters such as the criteria, the method, the timing, and the scorer of the measurement, have a significant role in predicting the efficiency of the assessment in this skill (Crusan, 2013). In this vein, a new term, i.e., "*second language writing assessment literacy*", has been coined by Crusan, Plakans, and Gebril (2016: 44), to showcase the uniqueness of the Writing Assessment Literacy (WAL) as a specialized domain, fueled by the interconnected dynamics of teachers' core knowledge, conceptions, and implementations (Crusan et al., 2016: 43; Lam, 2019: 81).

Insofar, much research has been dedicated to the exploration of how to assess writing from various angles (e.g., Andrade, Du, & Mycek, 2010; Esfandiari, & Myford, 2023; Ghanbari & Barati, 2020; Joo, Seong, Suh, Jung, & Purpura, 2023; Li, 2023). However, far less has been revealed about teachers' literacy regarding the writing assessment system, albeit the salience of and emphasis on their positions in conditioning this process (Crusan et al., 2016). Therefore, this review study is aimed to thoroughly elaborate different elements at play in the assessment procedure in L2/EFL composition classes, based on its theoretical grounding and to present a synopsis of the previous studies conducted within the scope of teachers' WAL.

Assessment of writing

Writing is a complicated skill, layered by different sub-processes that a writer has to regulate such as "thinking, planning, organizing, and linking as well as several levels of language manipulation" (Hamp-Lyons, 2003: 163). Similarly, its assessment poses challenges that urge teachers to design an effective writing task, score the competence reliably, and give descriptive feedback to students (Nguyen & Truong, 2021), by meeting varying purposes such as placing learners in appropriate classes according to their proficiency, diagnosing their weak or strong sides, informing about their achievement, giving information about how competent they are in a particular writing activity, and evaluating their general competence in composing (Hyland, 2003: 214).

Historically, the writing assessment is reliant on product-oriented approaches, framed by the premise that students' final products be evaluated by prioritizing the accuracy element. Since the ultimate goal is to yield error-free performances (Nunan, 1999), scoring the outputs is exclusively based on surface-level structures of the text, including syntax, lexis or cohesive devices (Hyland, 2003). Such a framework stands for the summative assessment model, whose main function is to assign an end-of-course or end-of-year grade for the performance which sums up the student's progress in the corresponding skill (Hyland, 2003; Scarino, 2013; Shieh & Cefai, 2017: 30). Summative assessment broadly refers to the assessment of learning (AoL) framework (Lee, 2017: 1), which is settled on the tenets of behaviorist theory of learning and looks for objective evidence of achievement with an emphasis on the production of numeric scores (Shepard, 2000). In the writing specificity, the AoL perceives the composing path as a linear process, which aims at forming an error-free text (Grabe & Kaplan, 1996), designed in a single-drafting methodology (Lee, 2017). However, upon recognizing the criticality of the "learning-to-write" as well as the "writing-to-learn" strategies for gaining competence in composing (Manchón, 2011: 3), the traditional product-based assessment model has been challenged by the process-oriented assessment approach (Mamad & Vígh, 2023). The latter type primarily values the composing journey, rather than how the output looks like (Applebee, 1986), in which learners go through a recursive and non-linear procedure (Hyland, 2003), by looping back and forth among the steps of "brainstorming, planning, drafting, revising, editing, and publishing" (Lee, 2017: 14). In this sense, portfolio assessment is noted as a good example of this type of assessment, where students compile a series of written outputs in a range of genres, drafted in multiple cases over a course of time (Lam, 2013).

The process-based consideration in writing pursues the formative assessment (Grabe & Kaplan, 1996), which serves to judge performances with an ultimate purpose of leading to substantial gains and sustaining learning through the ongoing delivery of feedback (Black & William, 2009). The formative assessment paradigm largely represents the AfL schemata, which has recently garnered a deal of attention with its attributes in feeding back students' responses for directing their subsequent renditions (Houston & Tompson, 2017: 8). While the AoL paradigm has been designed to regulate the aims of "accountability, or of ranking, or of certifying competence" (Black, Harrison, Lee, Marshall, & Wiliam, 2004: 10), the AfL module is primarily intended for promoting learning and fine-tuning the instruction (Wang et al., 2020). This assessment type, i.e., AfL, observes the principles of self-regulation (Lee, 2017) by encapsulating its driving facets, which surround students' rigorous endeavors in defining learning and growth goals, tracking their success, and deciding how to fill in any gaps (Andrade, Huff, & Brooke, 2012: 8). In this student-centered methodology, the concept of feedback acts as a catalyst for the assessment (Lee, 2017), directly resonating with self-regulated learning (Lam, 2018). In the AfL modelling, the source of the feedback might derive from a teacher, a peer or the self, though the teacher feedback dominates the AoL process (Gardner, 2006). Whoever the source is, the central plank of this phenomenon lies at its quality, which translates into the formation of "clear, descriptive, and diagnostic feedback" offered to the learners against a set of points of reference formulated in the "feed up" stage (Lee, 2017: 15). In line with this, the employment of analytic grading grids is highly recommended for giving constructive feedback, over the holistic scoring schemes based on the rating routine with an overall impression of multiple traits (González, 2018), since the former measurement tools can thoroughly remark various dimensions in a performance in light of certain standardized criteria, hence increasing the possibility of the uptake of feedback (Weigle, 2002).

In addition to the detailed feedback-provision function, the use of analytic rubrics can also lessen the subjectivity, inherent in the performance-based assessment (González, 2018; Nguyen & Tsuong, 2021), as they cover certain guidelines to track. Especially inexperienced assessors, i.e., peers or the self, can have difficulty in coordinating certain parameters such as the criteria to rate and their respective descriptors, or weightings (Weigle, 2002). And such assessment tools can be a frame of reference in this sense (Hyland, 2003). Considering that "the score is ultimately what will be used in making decisions and inferences about writers" (Weigle, 2002: 108), these rubrics can serve these purposes by providing reliable, transparent, and consistent scorings between and in-subjects renditions (Jonsson & Svingby, 2007). Additionally, by taking advantage of these rating schemes, self and peer-evaluation strategies, conceptualized within the underpinnings of the AfL framework as alternative measurement types (Wang et al., 2020), might enable learners to become more autonomous, reflect on their own learning, and have a deep understanding of what is expected from them. By spreading their efforts and engagement through such activities, students' life-long learning can also be promoted (Sambell & McDowel, 1998; Wanner & Palmer, 2018).

In short, the assessment of writing does not operate in a vacuum, but manifests itself as a multi-dimensional task, where an array of factors has an anticipating role in generating a productive and reliable measurement. In this vein, teachers involve in the assessment process as an important stakeholder, by assuming the roles of designers, evaluators, and facilitators of the process (Weigle, 2007). The effectiveness of assessment methodology is mostly dependent on their reflections, knowledge, beliefs, cognitions, and practices (Lee, 2017), all of which serve the grounding of the teacher's assessment literacy in writing (Crusan et al., 2016: 43; Lam, 2019: 81).

Writing Assessment Literacy (WAL)

Assessment literacy is a phenomenon, mediated by teachers' interpretation of what assessment stands for with its defining merits and how it should be properly applied (González, 2018). According to Scarino, the knowledge base of literacy "encompasses not only diverse assessment paradigms, theories, purposes, and practices related to elicitation, judgment and validation in diverse contexts, but also learning theories and practices and evolving theories of language and culture" (2013: 314). This framework is composed of "the fundamental principles of the 'what', 'why', and 'how,' without which teachers cannot engage with assessment at a deeper level" (Xu & Brown, 2016: 159). More notably, its uptake is filtered by teachers' conceptions, which denote their belief system, consisting of intuitive perceptions of the nature and aim of the assessment, mostly framed by cognitive and affective strands. Eventually, this knowledge base, guided by the teachers' conceptional scheme (Xu & Brown, 2016), is interpreted and exercised in a variety of ways, such as concretizing the construct, designing tasks to elicit performances from students, specifying how students respond to these activities, formulating evaluation criteria and standards, and translating evidence as part of verification procedures (Scarino, 2013: 323).

In the writing domain, Lam (2019: 81) presents a three-structure WAL system which is built on the framework of teacher assessment literacy, formulated by Xu and Brown (2016: 155). The first component, "*the knowledge base*", is comprised of "assessment knowledge of designing, implementing, grading, and providing feedback for improving student learning" (Lam, 2019: 79). "*Teacher conceptions*", the second dimension of the system, are deemed "an internal guiding framework of how teachers perceive the purposes and uses of assessment relating to their beliefs" (Lam, 2019: 80). The

last part of WAL covers *"assessment practices"* (Lam, 2019: 80), interconnected with the previous two dimensions mentioned in this section. Such practices refer to how and which evaluation implementations are performed in the writing specificity (Crusan et al., 2016; Lam, 2019: 80).

Along those same lines, a number of researchers have conducted research which represents these triadic components of the WAL construct. Such studies will be briefly presented with their major findings in this part. Crusan et al. (2016), for example, explored the dynamics of WAL by designing a pioneering inventory, addressing teachers' assessment literacy in the writing domain through a three-structure model, composed of teachers' knowledge, their philosophies, and practices in the writing assessment. A total of 702 teachers working at various tertiary institutions from 41 countries were recruited for the study. A 54-item self-report survey was designed by the researchers, whose items were formulated in the Likert scale, alongside open-ended and multiple-choice question formats (Crusan et al., 2016: 46). The results displayed that teachers had sufficient knowledge in the design of writing tasks for assessment. They also reported having the knowledge of alternative assessment models and integrated writing activities. However, a majority of the participants were unsure of how to design a rubric for measuring the writing competence (Crusan et al., 2016: 49). As for their beliefs, the general attitude reflected on by a substantial majority had a positive tendency. A majority confirmed that a range of writing tasks, i.e., out-of-class writing assignments, portfolios, and timed in-class writing tasks, were helpful in the assessment. But surprisingly, two-thirds of the participants reported that the measurement of writing tasks inevitably yields inaccurate results. Additionally, in the written comments made by the considerable number of the teachers, writing assessment was deemed as challenging and frustrating (Crusan et al., 2016: 49-50). With respect to their practices, a vast majority reported benefiting a multiple-draft methodology, self-assessment activities, and rubrics in the writing evaluation process (Crusan et al., 2016: 50-51).

The modified and adapted version of the same instrument designed by Crusan et al. (2016) was also administered by Weng (2023), with an intention of investigating the level of teachers' WAL with its sub-structures, and secondarily displaying the possible relations between these dimensions of the construct and the participants' demographic profiles. Data were gathered from 219 tertiary level EFL teachers in the Chinese setting. The findings indicated that although they had fairly enough knowledge in the writing instruction, a small portion of the teachers put that they had taken formal training, with a direct reference to the writing evaluation. The findings also displayed that the rating of the participants' assessment knowledge collapsed into the "somewhat familiar" cluster (Weng, 2023: 64). Specifically, high frequency scores were measured in the

familiarity of the key implementations such as the design of effective writing tasks, scoring via rubrics, and integrated writing tasks. However, comparably less number of the teachers had knowledge for the application of portfolios (56%) and alternative methods in the evaluation of writing tasks (48%). Regarding their beliefs, many respondents held positive attitudes towards writing assessment. Yet, nearly half of the participants reported that the scoring of writing tasks tends to yield subjective measurements (51%). Additionally, 53% of the participants stated that the criterion of content should be prioritized over grammar. Interestingly, the use of essay writing was favored by only 43% of the teachers, despite being a frequently-adopted technique for the evaluation in the Chinese context. In the practices cluster, most teachers stated utilizing the scoring rubrics (75%) and self-assessment tasks (60.3%) (Weng, 2023: 65). Further, the findings revealed certain correlations between WAL and demographics. Regression analyses were also performed, confirming that participants' knowledge and beliefs could predict their assessment practices (Weng, 2023).

Sohrabi et al. (2022) also replicated the employment of the instrument formed by Crusan et al. (2016), by exploring teachers' perceptions of WAL in the Iranian EFL context, with a sample of 118 participants. The quantitative data showed that the teachers were adequately knowledgeable about the key implementations including portfolios, integrated writing activities, and rubric creation and employment. Considering their beliefs, overall, the teachers found various writing applications helpful for evaluation. However, they were undecided about whether the scoring of the writing texts has a subjective nature. While 42% of the respondents endorsed on this issue, about 41% of the participants showed disagreement. A similar score was also calculated in judging the effectiveness of self-assessment tasks. Although 39% of the teachers were ambivalent, 32% of the participants indicated disapproval and only 26% agreed on the efficiency of this task type (Sohrabi et al., 2022: 11). All in all, although a more positive trend for the writing assessment was reported by the participants as part of their knowledge and belief systems, their perceived WAL level in these two constructs mismatched their assessment practices in composing to some extent. That is, unlike the substantial endorsement rating calculated in the knowledge and beliefs components, the results were not that high in the practices sub-scale. A downward trend was measured in the points such as rater-training, rubric creation, the use of portfolios, integrated writing tasks, and the integration of computer technology (Sohrabi et al., 2022).

Similarly, Thirakunkovit (2019) also designed a study in which an online survey formed by Crusan et al. (2016) was administered to 52 Thai university instructors with follow-up interviews with 21 participants. The findings indicated that while the majority of the instructors took teacher training in writing evaluation, nearly 30% of the participants reported never having received any formal

education in this sense (Thirakunkovit, 2019: 1140). Moreover, a considerable number of the instructors indicated a need for receiving more training in the reliable rubric design and the implementation of alternative assessment techniques. Regarding their beliefs, the majority endorsed on the idea that the evaluation process is time-consuming. Moreover, nearly half of the participants believed that the scoring of writing tasks could yield inaccurate results. A substantial majority also reported that the evaluation of compositions inevitably ends up with subjective scorings. Another interesting finding is that a considerable rate of the participants was uncertain about the efficiency of the self-assessment procedures. Nearly one-third of the instructors reported being undecided about the need for the portfolio application (Thirakunkovit, 2019: 1142). Regarding the day-to-day practices in class, only 56% of the participants reported using a multiple-draft methodology and only 52% of the participants stated benefiting rubrics or a set of criteria in each assignment (Thirakunkovit, 2019: 1140). In the interview data, it is noted that interviewees did not much frequently employ self-assessment methods, peer evaluation, or portfolio use (Thirakunkovit, 2019).

In a mixed-methods study, Lam (2019) also analyzed Hong Kong secondary school teachers' literacy in the writing assessment. The data were derived from a questionnaire, alongside telephone interviewing and observations. The results indicated that they had moderately assessment knowledge as part of their WAL. Although half of the teachers knew how to ensure fair assessment and classroom-oriented evaluation, some teachers were still not relatively knowledgeable about the theoretical stance of validity and reliability issues and test construction in the same way (Lam, 2019: 83). Regarding their conceptions, they had a positive stance for the employment of the alternative assessment techniques, the use of criterion-reference rubrics, and the multiple-drafting approach. Concerning their practices, even though a substantial majority reported using in-class writing tasks, relatively smaller number of the participants stated utilizing group projects (36.9%) and portfolios (12.3%). In the scoring issue, less than the half reported using a rubric with multiple traits when assigning a grade (Lam, 2019: 84).

In short, this overview of the pertinent early research addressing L2/EFL teachers' assessment procedures displays the emerging themes in the WAL mechanism, grounded on a theoretical foundation. This review also highlights the points where teachers feel competent and efficacious in guiding the assessment process as well as the situations where they need more professional training (Table 1).

Çizelge 1: Önceki araştırmada bildirilen, özel dikkat gerektiren bulguların özeti

Table 1. Summary of the findings that require specific attention, reported by the previous research

Researchers	The points where teachers feel less knowledgeable	Teachers' negative conceptions of certain assessment points	Implementations where teachers less frequently get involved
Crusan et al. (2016)	• rubric design	• finding challenging/ frustrating • inaccurate scorings	
Weng (2023)	• portfolio design • alternative methods	• subjective measurements	
Sohrabi et al. (2022)		• subjective measurements • self-assessment tasks	• rater-training • rubric creation • portfolio application • integrated writing tasks • computer technology
Thirakunkovit (2019)	• formal training in writing • rubric design • alternative assessment methods	• time-consuming • inaccurate/ subjective scorings • self-assessment methods • portfolio employment	• multiple-draft methodology • rubric use • self-assessment methods • peer-evaluation • portfolio assignment
Lam (2019)	• validity • reliability		• portfolios • group projects • rubric

Discussion and implications

From the theoretical philosophies and the findings of the early research summarized above, it is evident that teachers' WAL is a multi-structured concept, operationalized by three major components, i.e., knowledge, conceptions/beliefs, and practices (Crusan et al., 2016). These three structures denote dynamic, context-specific, and inter-related elements, proving a predictive role for affecting one another, situated on a cyclical continuum (Lam, 2019; Xu & Brown, 2016). In this vein, there is a pressing need for language assessors to accumulate comprehensive knowledge, built on a theoretical stance covering "a core body of formal, systematic, and codified principles concerning good assessment practice" (Xu & Brown, 2016: 155), which

results in the formation of the assessment conceptions schemata entailing their perspectives for the utility of a range of evaluation methodologies. Accordingly, the reliable and valid practices of writing assessment might be largely offset by the core knowledge basis of how to enact specific tasks and the internalization of the significance of those implementations in the belief system (Weng, 2023). And in the reverse directionality, the deployment of a specific assessment task might shape the knowledge framework and the beliefs/conceptions entity, in such a way that will eventually result in the approval or disapproval of that method for the subsequent uses (Xu & Brown, 2016). This triadic interplay among these three foci in the assessment paradigm underscores the prominence of extensive training based on both theoretical and practical assessment applications as part of teacher professionalism to gain in-depth understanding of the core principles in the field, gain cognition relating thereto, and employ them properly in class (Lam, 2019).

Notably, the aforementioned studies (e.g., Sohrabi et al., 2022; Thirakunkovit, 2019) have also proven that some teachers do not have much expertise and experience in the systemization of the alternative assessment strategies. One such field is the portfolio assignment that merits further guidelines in order for its incorporation in class. As it is known that the portfolio application stresses self-reflection, learner autonomy, and the rigorous engagement in the composing path rather than valuing the final output (Hamp-Lyons & Condon, 2000). However, based on the data elicited in some early studies (e.g., Lam, 2019; Sohrabi et al., 2022; Weng, 2023), it appears that some teachers need in-depth knowledge for its role in facilitating formative assessment and more expertise in how to functionalize this technique effectively. To this end, teachers might be guided for properly navigating the portfolio-assessment approach, pertaining to the issues such as specifying the objectives and their links to curricular aims, laying out the content according to the needs, selecting the relevant tasks to be utilized and assessing the student progress and the efficiency of this application (Delett, Barnhardt, & Kevorkian, 2001).

The deployment of self-assessment procedures is also identified as another prominent instance where teachers should gain more awareness and competence (Sohrabi et al., 2022; Thirakunkovit, 2019). Self-editing strategies are critical to maximize the affordances for self-regulating one's own learning (Wang et al., 2020), through purposefully monitoring and observing attempts, exercised by cognitive and metacognitive orientations (Diab, 2011). Additionally, self-based revisions are known for leading to learners' agentic engagement, in such a way that helps to discern weak and strong patterns in the writing outputs (Esfandiari & Myford, 2013), with a judgment of to what extent the stated goals or the criteria have been met (Andrade et al., 2010). Moreover, given the nature of writing competence promoted by life-long efforts, the importance of self-feedback is undeniable as a

continuous assessment tactic for sustaining the writing development in the long-run (Jiliang & Kun, 2007). Therefore, it seems urgent for assessors to provide adequate supervision to learners, through which these innovative ways will be ultimately added to their learning strategies repertoire (Sohrabi et al., 2022).

Another emerging theme in the previous studies (e.g., Sohrabi et al., 2022; Weng, 2023) is that some teachers endorsed on the potential of assessing writing in generating subjective or inaccurate scorings. Indeed, the utility of analytical rubrics can minimize this drawback to a great extent by eliminating arbitrary measurements (González, 2018; Thirakunkovit, 2019). In this vein, it seems that teachers need more guidance in the construction of scoring tools, accompanied by their practice in the classroom setting (González, 2018), which might eventually ensure the reliability, consistency, and the validity in the assessment (Hyland, 2003). In this sense, professional direction through training programs can help gain in-depth understanding of the proper design and use of rating instruments in the writing field (Dempsey, PytlikZillig, & Bruning, 2009).

Conclusion

In sum, this review study can contribute to the writing pedagogy by scrutinizing the internal structure of the WAL, which demonstrates an interplay among three important dimensions of assessment, i.e., knowledge, the conceptual framework, and practices, and how these three components manifest in the evaluation of composed texts (Crusan et al., 2016; Lam, 2019). Also, the portrait of teachers' perceptions of their WAL levels, in light of the previous research results, can give an idea about where they feel competent and where they need more guidance in coordinating the writing assessment (e.g., Crusan et al., 2016; Lam, 2019; Sohrabi et al., 2022). Depending on this overview of theoretical principles and empirical findings drawn from the previous research, some new studies can also be attempted to revitalize this phenomenon by addressing other angles. To begin with, most of the earlier research is dominated by the employment of self-reported questionnaires as the primary data-collection tool (e.g., Crusan et al., 2016); however, such instruments can be updated in a way that captures other aspects, not emphasized in-depth. For example, the notion of feedback is not thoroughly addressed in the previous inventories. Therefore, considering that researchers (e.g., Lee, 2017; Weigle, 2007) have vouched for the criticality of systemic attention devoted to the relationship between feedback literacy and the WAL, the concept of feedback should find its own place in the construction and validation of a new scale for measuring WAL. Moreover, even though self-report instruments can yield important findings, consultation to teachers in a qualitative research design, e.g., the application of interview protocols, can shed light on the aspects, otherwise will presumably remain unexamined. In addition, new research based on gathering opinions derived from the other stakeholders of

the assessment process, e.g., students, can also report promising results. Last, the WAL paradigm should be addressed, conforming to the contemporary writing instruction, shaped by the advent of technology (Al-Bahlani & Ecke, 2023). To this end, new inventories can be formulated in a coverage of modern writing assessment ways, facilitated by digital tools.

Extended Summary

Given the role of the assessment in ensuring the effective writing instruction through the fulfillment of an array of respective functionalities such as evaluating local and general composing competence, specifying strengths or weakness of the learners, or informing the writing pedagogy about the efficiency of the instruction (Lee, 2017), it is no surprising to find that this concept has been steadily emphasized as a central parameter shaping the writing education (Lee, 2007; Shepard, 2000). In this sense, two assessment theories have received a growing recognition in the composing genre: The formative vs summative assessment approaches (Hyland, 2003). While the former, also known as the assessment for learning (AfL) schemata, is based on facilitating instruction by means of continuous feedback over the course of time with an ultimate purpose of guiding the subsequent performances (Black et al., 2004), the latter, denoting the assessment of learning (AoL) orientation, is about granting an ultimate grade which outlines one's proficiency in the respective skill at the end of a course (Lee, 2011). In contrast to its counterpart, the AfL methodology is highly stressed for enabling learners to develop their self-regulatory competencies, noted as the catalyst for taking initiatives to direct their learning and to manage the process with the rigorous endeavors. In this sense, the provision of feedback, either by the teacher, peer or the self, is deemed as the pillar of self-regulation by helping learners have a clear idea about their progress and eventually provide a basis for the subsequent renditions (Lee, 2017). For such aims, especially the employment of analytic rubrics is suggested to systematize the assessment methodology by highlighting different aspects in a performance against a set of criteria. Accordingly, students might be deeply informed about the strong and problematic points in their renditions (Weigle, 2002). Overall, it is evident that writing evaluation is a complicated process, driven by a number of factors (Sohrabi et al., 2022) and teachers assume the primary role of navigating this complex system. Especially, their literacy of the writing assessment is noted as a driving force for coordinating this procedure (Crusan et al., 2016). However, there exists a lack of research, exploring the assessment literacy of teachers in the writing genre specificity. For this reason, this review study is intended for presenting theoretical considerations, with a direct reference to the writing evaluation framework and summarizing the main findings of the existing research (e.g., Crusan et al., 2016; Lam, 2019; Sohrabi et al., 2022; Thirakunkovit, 2019; Weng, 2023) in order for displaying a general portrait of writing teachers' assessment literacy levels.

In theory, it is conceptualized that assessment literacy in writing is composed of teachers' knowledge, beliefs, and current practices (Crusan et al., 2016: 43). Depending on this internal structure of the construct, this notion has been explored in the EFL or L2 composing process. In light of such existing studies (e.g., Crusan et al., 2016; Lam, 2019; Sohrabi et al., 2022; Thirakunkovit, 2019; Weng, 2023), certain themes have been identified for further attention. First, it has been proven that some teachers do not have adequately expertise and experience in the systemization of the alternative assessment ways (e.g., Sohrabi et al., 2022; Thirakunkovit, 2019). In this sense, further guidelines are needed by teachers for the embedment of the portfolio assignment in the classroom assessment. Self-assessment tasks also appear as the other innovative technique where teachers do not feel much experienced about (e.g., Sohrabi et al., 2022). Another emerging theme reported by some of the participants in the existing research (e.g., Sohrabi et al., 2022; Weng, 2023) is that writing assessment might generate subjective or inaccurate scorings. Indeed, as a solution to this problem, the use of analytical rubrics can offer more reliable, consistent and valid measurements as compared to holistic scoring schemes (González, 2018; Thirakunkovit, 2019). However, teachers seem to need more training in the construction and implementation of scoring tools (Lam, 2019; Thirakunkovit, 2019). In short, what is discussed in this paper can delineate the underlying mechanism of the assessment literacy in the writing specificity based on the theoretical principles and reveal the points where teachers feel competent and need more guidance as part of their professionalism (Crusan et al., 2016; Lam, 2019).

References

Al-Bahlani, S. M., & Ecke, P. (2023). Assessment competence and practices including digital assessment literacy of postsecondary English language teachers in Oman. *Cogent Education, 10(*2), 1-18, 2239535. https://doi.org/10.1080/2331186X.2023.2239535

Andrade, H., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy & Practice, 17*(2), 199-214, https://doi.org/10.1080/09695941003696172

Andrade, H., Huff, K., & Brooke, G. (2012). *Assessing learning: The students at the center series*. New England: The Nellie Mae Education Foundation.

Applebee, A. N. (1986). Problems in process approaches: Toward a reconceptualization of process instruction. In A.R. Petrosky & D. Bartholomae (Eds.), *The teaching of writing: Eighty-fifth Yearbook of the National Society for the Study of Education, Part II* (pp. 95-113). Chicago: National Society for the Study of Education.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 9–21.

Crusan, D. (2010). *Assessment in the second language writing classroom.* University of Michigan Press.

Crusan, D. (2013). Assessing writing. In A. J. Kunnan (Ed.), *The companion to language assessment. Volume 1: Abilities, contexts, learners* (pp. 201–215). Hoboken, NJ: Wiley-Blackwell. https://doi.org/10.1002/9781118411360.wbcla067

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing, 28*, 43-56. http://dx.doi.org/10.1016/j.asw.2016.03.001

Delett, J. S., Barnhardt, S., & Kevorkian, J. A. (2001). A framework for portfolio assessment in the foreign language classroom. *Foreign Language Annals, 34*(6), 559–568.

Dempsey, M. S., PytlikZillig, L. M., & Bruning, R. H. (2009). Helping preservice teachers learn to assess writing: Practice and feedback in a Web-based environment. *Assessing Writing, 14*(1), 38-61. https://doi.org/10.1016/j.asw.2008.12.003.

Diab, N. M. (2011). Assessing the relationship between different types of student feedback and the quality of revised writing. *Assessing Writing, 16*(4), 274–292.

Esfandiari, R., & Myford, C. M. (2013). Severity differences among self-assessors, peer- assessors, and teacher assessors rating EFL essays. *Assessing Writing, 18*(2), 111-131. https://doi.org/10.1016/j.asw.2012.12.002

Ghanbari, N., & Barati, H. (2020). Development and validation of a rating scale for Iranian EFL academic writing assessment: A mixed-methods study. *Language Testing in Asia, 10*(1), 1–21.

Gardner, J. (2006). *Assessment and learning*. London: Sage.

González, E. F. (2018). *The impact of assessment training on English as a foreign language university professors' classroom writing assessment: Reported practice and perceptions* (Unpublished PhD Dissertation). University of Southampton, the UK.

Grabe, W. & Kaplan, R. B. (1996). *Theory and practice of writing*. London: Longman.

Hamp-Lyons, L., & Condon, W. (2000). *Assessing the portfolio: Principles for practice, theory and research*. Cresskill, NJ: Hampton.

Hamp-Lyons, L. (2003). Writing teachers as assessors of writing. In Kroll, B. (ed.), *Exploring the Dynamics of Second Language Writing* (162-189). New York, USA: Cambridge University Press.

Houston, D., & Thompson, J. N. (2017). Blending formative and summative assessment in a capstone subject: 'It's not your tools, it's how you use them'. *Journal of University Teaching & Learning Practice, 14*(3), 1-15. https://doi.org/10.53761/1.14.3.2

Huang, J. (2009). Factors affecting the assessment of ESL students' writing. *International Journal of Applied Educational Studies, 5*(1), 1-17.

Hyland, K. (2003). *Second language writing.* London: Cambridge University Press.

Jiliang, C., & Kun, W. (2007). An investigation of scorer reliability of self and peer assessment of EFL writing among Chinese college students. *CELEA Journal, 30*(1), 3–11.

Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144. https://doi.org/10.1016/j.edurev.2007.05.002

Joo, S. H., Seong, Y., Suh, J., Jung, J.-Y., Purpura, J. E. (2023). Assessing Korean writing ability through a scenario-based assessment approach. *Assessing Writing, 58*, 100766. https://doi.org/10.1016/j.asw.2023.100766

Lam, R. (2013). Two portfolio systems: EFL students' perceptions of writing ability, text improvement, and feedback. *Assessing Writing, 18*(2), 132–153.

Lam, R. (2018). Understanding assessment as learning in writing classrooms: The case of portfolio assessment. *Iranian Journal of Language Teaching Research, 6*(3), 19-36. https://doi.org/10.30466/ijltr.2018.120599

Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System, 81*, 78-89. https://doi.org/10.1016/j.system.2019.01.006

Lee, I. (2007). Assessment for learning: integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *The Canadian Modern Language Review, 64*(1), 199-213. https://doi.org/10.3138/cmlr.64.1.199

Lee, I. (2011). Bringing innovation to EFL writing through a focus on assessment for learning, *Innovation in Language Learning and Teaching, 5*(1), 19-33, DOI: 10.1080/17501229.2010.502232

Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Singapore: Springer.

Li, A. W. (2023). Using *Peerceptiv* to support AI-based online writing assessment across the disciplines. *Assessing Writing, 57*, 100746. https://doi.org/10.1016/j.asw.2023.100746

Mamad, A., & Vígh, T. (2023). Moroccan EFL public university instructors' perceptions and self- reported writing practices. *Heliyon, 9*(5), 1-12. https://doi.org/10.1016/j.heliyon.2023.e15626

Manchón, R. M. (2011). Situating the learning-to-write and writing-to-learn dimensions of L2 writing. In R. M. Manchón (Ed.), *Learning-to-write and Writing-to-learn in an Additional Language* (pp. 3–14). Philadelphia, PA: John Benjamins Publishing Company.

Nguyen, T. H. H., & Truong, A. T. (2021). EFL teachers' perceptions of classroom writing assessment at high schools in Central Vietnam. *Theory and Practice in Language Studies, 11*(10), 1187-1196. https://doi.org/10.17507/tpls.1110.06

Nunan, D. (1999). *Second language teaching & learning*. Boston, MA: Heinle & Heinle.

Sambell, K. & McDowell, L. (1998). The value of self and peer assessment to the developing lifelong learner. In Rust, C. (Ed.). *Improving Student Learning: Improving Students as Learners* (pp. 46-56), Oxford, UK: Oxford Centre for Staff and Learning Development.

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation

in assessment and in teacher learning. *Language Testing, 30*(3), 309–327.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Shieh, J., & Cefai, C. (2017). Assessment of learning and teaching in higher education: A case analysis of a university in the south of Europe. *Malta Review of Educational Research 11*(1), 29–47.

Sohrabi, Z., Ghanbari, N. & Abbasi, A. (2022). Iranian EFL teachers' perceptions of writing assessment literacy: a countrywide study. *Language Testing in Asia, 12*(23), 1-19. https://doi.org/10.1186/s40468-022-00172-7

Thirakunkovit, S. (2019). Exploring writing assessment literacy and classroom practices of Thai university instructors. *The Journal of Asia TEFL, 16*(4), 1135–1151.

Wang, L., Lee, I., & Park, M. (2020). Chinese university EFL teachers' beliefs and practices of classroom writing assessment. *Studies in Educational Evaluation, 66*(2020), 1-11. 100890. https://doi.org/10.1016/j.stueduc.2020.100890

Wanner, T., & Palmer, E. (2018). Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback, *Assessment & Evaluation in Higher Education, 43*(7), 1032-1047, https://doi.org/10.1080/02602938.2018.1427698

Weigle, S.C. (2002). *Assessing writing.* Cambridge, UK: Cambridge University Press.

Weigle, S. C. (2007). Teaching writing teachers about assessment. *Journal of Second Language Writing, 16*(3), 194-209.

Weng, F. (2023). EFL teachers' writing assessment literacy: Surveying teachers' knowledge, beliefs, and practises in China. *Porta Linguarum An International Journal of Foreign Language Teaching and Learning, 40 (2023), 57–74.* https://doi.org/10.30827/portalin.vi40.23812

Xu, Y., & Brown, G. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education, 58*, 149–162. https://doi.org/10.1016/j.tate.2016.05.010