



International Journal of Learning, Teaching and Educational Research
Vol. 24, No. 12, pp. 322-343, December 2025
<https://doi.org/10.26803/ijlter.24.12.14>
Received Aug 17, 2025; Revised Oct 29, 2025; Accepted Nov 3, 2025

From Pen to Performance: Unpacking the Impact and Practicality of PTW Rubrics on Malaysian L2 Learners' Self-Assessment

Nurul Fariena binti Asli*  and Mohd Effendi @ Ewan bin Mohd Matore 
Faculty of Education,
Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor Darul Ehsan, Malaysia

Abstract. One of the characteristics of an effective writing self-assessment tool for second language learners (L2) is its ability to provide meaningful feedback and foster autonomous learning. However, in assessing self-assessment for writing, a dearth of studies has addressed how the primary trait type of rubric could support learner engagement meaningfully and provide task-specific feedback. Utilizing Bachman's (2005) assessment use argument (AUA) validation framework, which focused on test usefulness qualities, this study investigated the impact and practicality of primary trait writing (PTW) rubrics as tools to evaluate students' writing ability where the perceptions from both student self-assessments and teacher assessments were gathered. A qualitative approach using open-ended interviews was employed to gather data following the classroom implementation of the PTW rubrics with six Malaysian secondary school students. The verbatim transcripts were analyzed thematically using ATLAS.ti version 22. The findings revealed eight themes related to the impact and three themes related to the practicality of PTW rubrics. The themes showed that the PTW rubrics positively influenced student learning and were practical for self-assessment activities. This study introduces another strategy for evaluating L2 students' writing skills as PTW rubrics are considered as a new tool for teachers to train students for self-assessment. Thus, students who can be independent learners and hold accountability in their own learning progress can be produced. This paper also adds to the existing body of knowledge by introducing a CEFR-aligned, self-assessment-oriented rubric that bridges theory and classroom practice for Malaysian L2 learners.

Keywords: Primary trait rubrics (PTW); Rater-mediated assessment; Self-assessment; Teacher assessment; Malaysia

*Corresponding author: Nurul Fariena binti Asli; p97677@siswa.ukm.edu.my

1. Introduction

Writing is one of the crucial skills that reflect students' competency in second language (L2) learning, where it serves not only as a means of communication but as a tool for cognitive and academic development. In educational settings, the ability to write effectively is an essential skill that needs to be mastered as it is assessed to determine language proficiency, primarily where English is taught as a second or foreign language (Dosi & Douka, 2021; Sun & Lan, 2024).

Consequently, writing assessment has gained prominence as one of the approaches to mastering the skill. Particularly, self-assessment has received growing attention due to its potential in promoting learner autonomy (Ebrahimi et al., 2021), metacognitive awareness (Fachrunnisa et al., 2024), and meaningful engagement during the writing process (Chambers & Harkins Monaco, 2023). While engaging in self-assessment activity, learners are able to reflect on their writing, identify areas for improvement, and take ownership of their progress (Anh et al., 2023).

Meanwhile, from a measurement perspective, self-assessment is considered a form of performance assessment where students take on the role of the rater. While this encourages student rater autonomy, it introduces variability in scoring due to the students' limited assessment experience and potential biases (Foley, 2019) compared to teacher assessment. In Malaysia, the demand to promote self-assessment activity as a form of formative assessment has been laid out in The English Language Roadmap by the Ministry of Education (MOE).

However, no standardized scoring method was prescribed for classroom use, leading to inconsistencies in assessment practices and teacher resistance (Franchis & Maslawati, 2023). Despite being emphasized in the policy, the implementation of self-assessment in writing activities in the L2 classroom settings remains limited and hindered by the lack of effective and appropriate scoring tools that offer clear, task-specific feedback and enhance student understanding. Thus, a critical need to develop and validate a practical scoring guide that can enhance consistent self-assessment practices while improving teacher acceptance and learner engagement in L2 writing.

In this context, scoring rubrics are seen as tools or methods that can support the validity and reliability of students' self-assessment, guiding and anchoring their judgments in assessing their essays (Vasileiadou & Karadimitriou, 2021). One significant issue lies in the type and use of scoring rubrics that effectively support self-assessment activities. While holistic and analytic rubrics have been extensively researched and used, they fail to guide learners through specific task demands as required for formative assessment contexts.

Meanwhile, the primary trait rubric has received limited attention in educational research even though it is widely recognized as particularly suitable for L2 learners (Weigle, 2010). This gap is due to the complexity of developing a primary trait rubric (Weigle, 2010; Veloo et al., 2018; Asli et al., 2024). The lack of studies using primary trait writing (PTW) rubrics to support learners' engagement and

formative self-assessment represents a critical gap that this study needs to address.

Given the specific needs of L2 learners, exploring the impact and practicality of the primary trait rubric for writing assessment is essential to address this gap and validate this rubric as an appropriate assessment tool for self-assessment in a secondary school setting. The impact and practicality are two out of five qualities that can be used to examine the usefulness of an assessment tool, as proposed by Bachman's (2005) assessment use argument (AUA) validation framework. Although these qualities were designed for assessing test use, they can also be applied to evaluate rubrics (Lallmamode, 2012). These underexplored validation areas help ensure that self-assessment rubrics function effectively in real classroom environments and are not theoretically explained only.

Addressing these gaps, the present paper examines the impact and practicality of PTW rubrics as tools for secondary school students to conduct self-assessment activities for L2 classrooms in Malaysia. Drawing on the AUA framework, this research explores the perceptions of both students and teachers regarding the use of PTW rubrics during classroom assessment activities. Therefore, this study seeks to address two research questions:

RQ1: What are the impacts of using PTW rubrics in self-assessment activities based on the students' and teachers' perceptions?

RQ2: To what extent are PTW rubrics practical for self-assessment activities based on the students' and teachers' perceptions?

2. Literature Review

2.1 Writing assessments and self-assessment in L2 education

Integrating self-assessment into writing assessments is essential for fostering self-directed learning, which has gained importance in the Malaysian education system and aligns with the Common European Framework of Reference for Languages (CEFR) (Adnan & Sayadi, 2021). Self-assessment fosters student learning ownership, thereby augmenting their reflective capacity, metacognitive awareness, and accountability for their advancement (Andrade, 2019).

In autonomous learning, self-assessment enables students to evaluate their strengths and weaknesses (Aldosari & Alsager, 2023) and fostering skills that promote lifelong learning (Butler, 2023). The CEFR's focus on learner autonomy facilitates this method by offering a systematic framework for language competency, enabling students to establish learning objectives and assess their language skills through "can-do" descriptors (Vasiljevic, 2023). These advantages highlight the importance of self-assessment in facilitating language development, particularly for second language (L2) learners, and confirm its congruence with educational autonomy and CEFR principles.

Even though self-assessment has many advantages, many challenges and constraints exist in its practical implementation, especially in the assessment of

writing. Writing assessments are subjective. It includes evaluating several linguistic elements, including content, structure, vocabulary, syntax, coherence (Knoch, 2021), organization, language use, and mechanics (Hajan et al., 2024). Teachers usually encounter disagreements when conducting writing assessments, and L2 learners struggle to assess their writing without explicit instruction, leading to unsatisfactory self-assessment practices. In addition, with the absence of valid and reliable scoring rubrics, students' interpretation of performance criteria is inconsistent, resulting in invalid and unreliable self-assessments (Harsch et al., 2024). Developing a specific rubric that emphasizes particular traits as criteria may help L2 learners focus on essential elements of writing and ensure that they are aligned with CEFR standards.

Identifying self-assessment rubrics that are specific and sufficient to address the unique needs of L2 learners primarily within the CEFR framework is challenging. Consequently, the accurate self-assessment activity is neglected, constraining the students' ability to assess their strengths and weaknesses and adjust their learning accordingly (Zhang & Zhang, 2022). These issues explain the need for validated CEFR-aligned self-assessment rubrics primarily designed for L2 writing assessment.

The developed rubrics for L2 learners must provide explicit level-specific descriptors where effective self-assessment with valid and reliable evaluations can be conducted (Weigle, 2010). Creating a new self-assessment rubric entails challenges related to design, validation, and psychometric characteristics. Developing a rubric that explicitly delineates the specificity suggested by the CEFR would involve iterative testing and validation to confirm that the rubric's criteria align with both learner needs and CEFR performance descriptors (Roche & Booth, 2021).

2.2 Assessment Use Argument (AUA)

Bachman and Palmer (1996) developed six criteria for assessing test usefulness, which are adapted from Messick's complex definition to the language evaluation field. The following are the six criteria: (a) construct validity refers to the accuracy and suitability of inferences made on a test taker's aptitudes from their test results, (b) reliability explains the degree of consistency in test performance as shown by the test results, (c) authenticity refers to the practical relevance and realism of the assessment criteria and tasks, (d) impact is the positive consequences the test may have on individuals and institutions within the immediate test setting and beyond and, (e) practicality which refers to the extent to which resources support the design and development of the test, as well as its ongoing use. This paper focuses on two out of the six criteria, which are the impact and practicality.

The concept of "impact" in Bachman's writing refers to the effects a test may have on people, rules, or procedures in the classroom, school, educational system, or society. The influence on test takers is evaluated using three different techniques. First, their perception of the Target Language Use (TLU) domain is influenced by their test-taking experience. Second, they are immediately impacted by the feedback they get about how well they performed on the test. Based on their exam

results, the test takers' choices may directly affect them (Bachman & Palmer, 1996). Third, the washback effects from teachers since teachers' instructions are also affected negatively and positively based on the test or classroom assessment (Rathnayake, 2025). Therefore, this study evaluated teacher and students' opinions about the impact of using PTW rubrics based on the effects mentioned by Bachman and Palmer (1996).

Other than that, according to Bachman (2005) in AUA, an assessment's practicality should be evaluated in light of how it affects the assessment's overall quality. Based on the opinions of teacher and students, this study gathered evidence of the PTW rubrics' practicality as well by taking into account variables including available funding, administration, scoring time, and rater availability, practicality guarantees that the assessment is possible to be conducted (Bachman, 2005).

Bachman's (2005) AUA framework serves as the guiding foundation for this study. While the framework encompasses six test usefulness qualities, this research focuses on practicality and impact, as these are most relevant to evaluating the feasibility and educational value of the PTW rubrics in Malaysian secondary classrooms. The practicality component guided the study design by emphasizing whether the rubrics can be effectively implemented within classroom constraints, such as time, clarity, and ease of use for both teachers and students.

The impact component informs the data collection and analysis processes, focusing on how the rubrics influence students' learning experiences, motivation, and engagement in writing. Evidence for these components was derived from semi-structured interviews. The interview data were analyzed thematically, allowing the study to demonstrate that the PTW rubrics are practical for classroom application and positively impact learners' writing awareness and autonomy.

3. Methodology

3.1 Research Design

This study employed a qualitative research methodology to investigate the impact and practicality of PTW rubrics as tools to evaluate students' writing ability based on students' and teachers' perceptions. Figure 1 presents the study flow for this paper. It started with a 60-minute English lesson, in which students were given time to read the writing task and write their essays. A rater training session followed, where the researcher had a discussion with the teacher and student raters. This was to ensure their understanding of the PTW rubrics and how to apply the descriptors accurately. Students rated a sample script with the researcher to develop a shared understanding before they rated their essays individually.

While weak students received clarification on the process, no assistance was given on scoring, and discussions among peers were not allowed to ensure the integrity of the self-assessment. The teacher also rated the photocopies of students' essays

after that. Once the rating session was over, selected students were called for the interview sessions with the researcher.

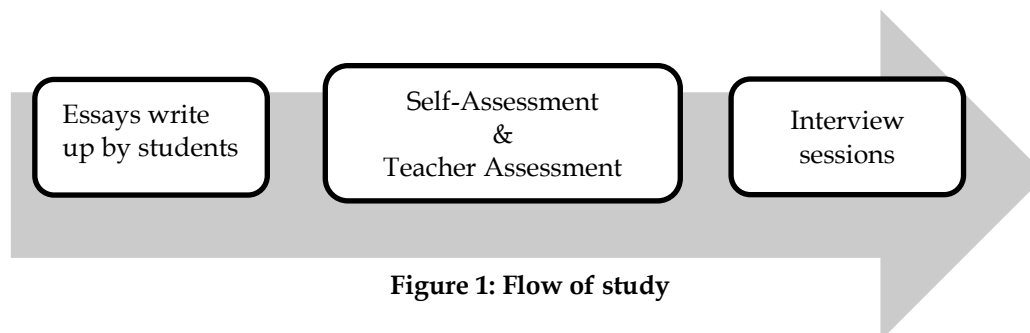


Figure 1: Flow of study

3.2 Context and Participants

The overall participants in this study were 149 Form 4 students from seven public secondary schools located in one of Malaysia's states, Pahang. The researcher only opted for national secondary schools, assuming they had undergone the same training and had the same exposure to the new syllabus.

Seven public schools were chosen to ensure that the participants were evaluated based on the performance standard (TP), which required students from different proficiency levels to participate so that a balanced mixture of all abilities was involved in this study. TP was used in Malaysia to determine students' proficiency level in each skill, and these were based on the CEFR level. TP 1 and TP 2 were for students with low proficiency, TP 3 and TP 4 were for the average, and TP 5 and TP 6 were for the highly proficient students.

Each student produced one essay, which was assessed using 4 PTW rubrics comprising four traits (criteria). Rater training was conducted using the 4 PTW rubrics and a sample essay to make sure students used the rubrics effectively. During the self-assessment implementation, students used the rubrics to rate their essays, while the teacher concurrently assessed the same ones using the rubrics. The teacher's scores and students' scores were observed, and they were similar. It showed consistency between the rubric interpretation.

An open-ended, semi-structured interview session with **six** selected students was conducted to yield the students' perceptions after the self-assessment activity. Meanwhile, one teacher rater was interviewed after the rating session finished. From each category, two students, a girl and a boy, were chosen, either on a volunteer basis or selected by their teacher. The breakdown of students and a teacher as interview participants is shown in Table 1 below.

Table 1: Breakdown of participants in the interview session

Category	Number of students and code
TP 6 (High proficiency)	1 Female (TP6 GIRL) 1 Male (TP6 BOY)
TP 4 (Average proficiency)	1 Female (TP4 GIRL) 1 Male (TP4 BOY)
TP 2 (Low proficiency)	1 Female (TP2 GIRL) 1 Male (TP2 BOY)
Teacher	1 Female (TCHR)

3.3 Instrument and Data Collection Procedure

Two instruments were used in this study: (a) four PTW rubrics and (b) teacher's perception interview and students' perception interview, adapted from Lallmamode (2012), which consisted of structured and semi-structured questions to elicit feedback regarding the quality and impact of the PTW rubrics. The PTW contained four rubrics with different criteria (content, cohesive devices, format, and sentence fluency). Experts validated the interview questions before using them. The interview protocol consisted of three parts: (a) an introduction that explained the purpose of the interview, the duration, confidentiality, and consent, (b) questions about the impact and practicality of the rubrics, and (c) closing remarks by thanking the participants.

3.4 Data Analysis

The Atlas Ti program (Version 22) was used for the qualitative data analysis, which involved independently analyzing the teacher and student interview datasets. The primary goal was to respond to the following study questions: (a) What are the impacts of using PTW rubrics in self-assessment activities based on the students' and teacher's perceptions? and (b) To what extent are PTW rubrics practical for self-assessment activities based on the students' and teacher's perceptions? Figure 2 illustrates the procedures involved in analyzing qualitative data. All recordings were transferred to the computer following the completion of the interviews. The audio files were extracted into scripts using Google Transcriber to obtain the transcription.

The researcher then began coding the transcriptions in any conceivable way. Coding was the process of organizing textual or visual data into digestible informational categories, searching for databases that support the code, and then assigning a name to the code (Creswell & Poth, 2018). Furthermore, it was recommended that diagrams be made to illustrate the connections between codes or novel ideas. Meanwhile, the visual aids facilitated the identification of code overlap. For instance, the concurrence tool can look for potential overlaps between codes and a network diagram of codes in Atlas.Ti (a qualitative software program) to visualize code interactions (Creswell & Poth, 2018).

Once coded and thematized, the extracted data were examined to confirm that they match the assigned code. Two inter-raters or inter-coders were chosen, and comprehensive instructions were given on Cohen's Kappa intercoder reliability

technique. Using Cohen's Kappa intercoder calculation, these data were then used to determine the Kappa value (Cohen, 1960). The mean Kappa value of 0.79 indicated a significant agreement (substantial agreement) between the raters, demonstrating the reliability of the coded data in this study.

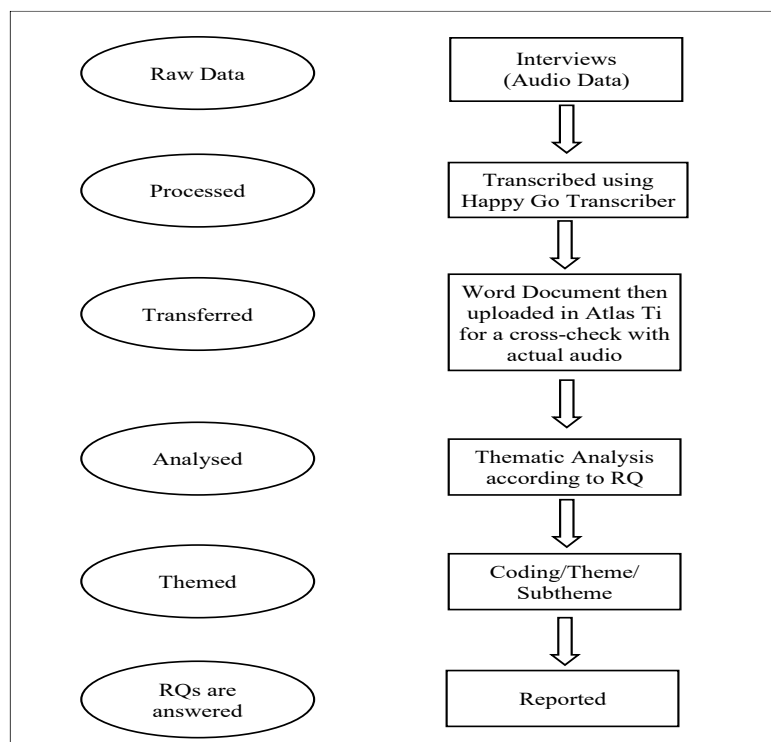


Figure 2: The steps in qualitative data analysis

4. Results and Findings

4.1 The impacts of the PTW rubrics

An essential factor to consider in developing and validating scoring rubric is the inclusion of stakeholders' perspectives. The findings illustrate the comprehensive effects experienced by all participants while utilizing the PTW rubrics. Table 2 below delineates the categorization of all codes corresponding to the subtheme. The data were consolidated into various subthemes and themes to address the study issue about impact. Eight themes had emerged as the primary impacts students and instructors experienced following the PTW rubrics' utilization. (Please disregard the grammatical faults seen in the snippets, as second language learners produce them.)

Table 2: Themes and subthemes for impact

Themes		Subthemes
1	Guidance from the rubric	The rubric acts as a teacher
2	Through the lens of a teacher	Emotion
		Understand how the teacher marks
		Know how to rate
3	Better understanding of the assessment	Understand better
		Give more info to score
4	Gain better results	Increase marks
		Perform better in exam
		Know how to score
5	Improve writing	Know how to write correctly
6	Confident to do self-assessment	Confident to use the rubric
7	Self-Awareness	Know own capability
8	Successful implementation of self-assessment	Use the rubric independently

Interviews revealed that students experienced a beneficial effect on their writing after utilizing the PTW rubrics. Several participants indicated that the rubrics functioned as a substitute instructor, since using the PTW rubrics provided guidance akin to a real teacher. They were also astonished that the rubric mirrored the teacher's grading precisely. Consequently, one of the benefits students derived from the PTW rubrics was an outline for enhancing their writing "*I think it's really impressive because the rubric are really like a teachers marking*". (TP4 GIRL).

Additionally, another effect observed by students was that the rubric enabled them to perceive evaluations from a teacher's perspective. Students empathized with the teacher's perspective while evaluating their essays emotionally "*I can understand what the teacher's feel because when I use the rubric, it is easier for me to search for what is right and what is wrong*".(TP2 GIRL). Consequently, it demonstrated that students were imbued with an emotional awareness, enabling them to empathize with the teacher's sentiments when evaluating their essays. Consequently, the study demonstrated that the rubrics facilitated students' empathy towards the teacher's emotions throughout the assessment of their writings. Additionally, a participant noted that he had a superior comprehension

of the assessment work, as the rubrics provided guidance on scoring criteria for the writing exercise, which he could use for his essay.

Furthermore, one participant said the rubrics afforded her a clear understanding of how to construct an essay that would attain a perfect grade, as she noted, "... I know how to write a full mark essay..." (TP4 GIRL) and for the student, after using the PTW rubrics, she was aware of the straightforward method to obtain high scores for her essay by uttering, "the way to get marks are really easy for me after seeing the rubrics" (TP4 GIRL). Furthermore, a student asserted that they observed an enhancement in their essay ratings due to applying the rubrics, claiming that "it can definitely help me increase my marks in writing and assist me in my exam to score better marks." (TP6 BOY). Simultaneously, a participant conveyed assurance in her capacity to excel in her assessments following the application of the PTW rubrics by saying, "I can perform better in my exams." (TP6 GIRL).

Another theme that emerged was improving writing, where the use of PTW rubrics had an impact and helped to improve students' writing. Two students mentioned that "...know how to do the essay" (TP2 BOY) and "What I learned is that I can actually improve my essay better just by using the rubrics. I can also learn new such as cohesive devices. Also, better use of my sentences and more." (TP4 BOY). Alongside the students' input, the teacher affirmed that the PTW rubrics were advantageous in directing pupils towards precise writing and improving their writing abilities. "The rubrics included all the critical components of the essay necessary for students to evaluate and grade. It can help the students improve their writing, where the students have to think about what they need to use, what type of cohesive devices that they need to use, how many cohesive devices, or how many sentences that you need to write in this one essay. And also, do they answer all of the content or the questions given and do they write in one paragraph, which is the format of an email. (TCHR)

Moreover, it was found that students experienced a boost in confidence when engaging in self-assessment activities after utilizing the PTW rubrics. The low-proficiency girl had mentioned this as she stated, "Yes, yes, I feel confident to use the rubric on my own." (TP2 GIRL), and the average proficiency boy said, "and I believe I can do it with confidence." (TP4 BOY). The teacher also supported this in the extract:

"I think the students will have confidence to use the rubric on their own because they would think that the rubric will help them to write a better essay. So they will have confidence, and they will use it." (TCHR)

Another impact derived from this study is the increase in students' self-awareness regarding their proficiency and capability in essay writing. During the conversation, a boy asserted his awareness of his abilities by stating, "I can even understand more of my capabilities in my essay." (TP4 BOY). The boy articulately conveyed his self-awareness of his ability. The teacher indicated that students exhibited self-awareness of their mistakes when using the rubrics.

The study concluded that the interview indicated a favorable result through the effective execution of self-assessment activities. This was apparent from the participants' capacity to autonomously employ rubric, as illustrated by these excerpts. "Yes, I definitely can use the rubric on my own..." (TP6 BOY), "Yes, I after

learning today I can know how I know how to use the rubric without a teacher's advice." (TP4 GIRL), and the teacher also responded *"Yes, they can use the rubric on their own, but only after being explained in the training session, and then they can use it on their own."* (TCHR). Considering all the excerpts provided, it is indisputable that both students and teachers had a decidedly good opinion of PTW rubric's impact on them. The statements from all participants were unequivocal in expressing their satisfaction with the utilization of PTW rubrics as an assessment tool for assessing the essays.

4.2 The practicality of PTW rubrics

The practicality of the PTW rubrics yielded three themes drawn from the subthemes, as illustrated in Table 3. Data saturation was attained when all participants discussed the same element. (Kindly disregard the grammatical faults seen in the excerpts, as they are produced by second language learners)

Table 3: Themes and subthemes for practicality

Themes	Subthemes
Suitable for all proficiency levels	Can be used by all levels of students' proficiency
	All levels can use
Time-efficiency	Less time to use
Rater-friendly rubric	Easy to use
	Simple rubric
	Understand all the words in the rubrics

The interview findings suggested that PTW rubrics were beneficial for students in self-assessment activities because they were suitable for individuals with diverse skill levels. This study recruited students with diverse skill levels to guarantee that the rubrics could serve as a learning tool for all Form 4 students in Malaysia. The participants were asked, *"Do you think all students can use the rubrics?"*. The excerpt, *"I think it can be useful, especially for the lower guys. They are sometimes struggling, especially in English, so implementing something like this can actually help them to improve their essay."* (TP4 BOY). The interview served as clear evidence, as all participants unanimously agreed that they believed that all students could utilize the rubrics.

The teacher offered a comprehensive commentary on the applicability of rubric for students with diverse skill levels. She articulated her concern, particularly for students with limited proficiency, by asserting that, *"Use the rubric? My students, the one that you have done the research, they are at the lowest level of proficiency, three and four. That's the lowest in my school. And yes, it has been proven that they can use the rubric. All proficiency students, TP 1 and 2, they can use rubric, but do they write the*

essay? They don't really write. If they don't write the essay, then the rubric will not be used for it all. But if they do write the essay, one or two sentences, then they at least can mark using at least one of the rubric components." (TCHR). Despite the teacher's concern for the low-ability students over the usage of PTW rubrics, her remark was optimistic, and she anticipated that those students would be capable of utilizing the rubrics to evaluate their essays.

Other than that, the students were asked, "Do you think it requires a lot of time to use the rubrics?" to determine how much time they need to spend using the PTW rubrics. All participants concurred that the rubrics were time-efficient, necessitating minimal student time commitment. The answers to the question are, "I think no." (TP4 GIRL) and "No, since the rubric is very easy and I think it can be done in short time." (TP6 BOY). The teacher also supported this, "Yes. Not that a lot of time. They need to be given time because they are not examiners. They have never marked their friends, their essay, or their own essay. They will take a few times, a few minutes to be aware of that, to familiarize themselves with the rubric. Maybe some of them have heard the word rubric first time when they were at that training session. They will have a few minutes in order to familiarize, but then they will know how to use it." (THCR). The teacher noted that students needed minimal time to apply the rubric. She elucidated that students might first require time to adapt to the rubrics. However, once they achieve proficiency, they only need a short time to apply them. The participants' comments revealed that PTW rubrics were time-efficient, rendering them a practical tool for self-assessment in the classroom.

Another issue that emerged from the study question on students' and teachers' perception of the practicality of rubrics was the concept of a rater-friendly rubric. The interview data revealed several subthemes that could be categorized under the primary theme of rater-rubric compatibility. The responses were categorized into three distinct tables. The subthemes were easy to use, the rubric was simple and straightforward, and one could understand all the words. The rubrics were claimed to be an assessment tool that was easy to use by the students, "...given in an easier way." (TP6 BOY), "...easier method without thinking too much." (TP6 GIRL), and it was agreed by the teacher as well, "Yes, it can be understood. I don't think this is a hard word like C1 or C2 vocabulary. All of these use B1 and B2 vocabulary so the students can understand if they read the rubric. So, they would understand. Proficient students will understand without being explained further on." (TCHR).

The rubrics were considered effective not just for their user-friendly design but also for their simplicity and clarity, which enhance student utilization. The notion of rubrics was clear, since the descriptors were uncomplicated. Additionally, one participant expressed that the rubric layout was simple for him. He said, "Layout is simple for me". (TP4 BOY). The rubrics were deemed practical as all the words were comprehensible to the students, as mentioned, "Yeah, mostly I can understand the words," (TP6 BOY) and "understand" (TP2 GIRL). The interview excerpts indicate that the three subthemes related to the rater-friendly themes demonstrate the practicality of PTW rubrics for L2 learners in self-assessment activities within the classroom.

5. Discussion

5.1 The impacts of the PTW rubrics based on AUA

In Bachman's work, the term "impact" refers to the consequences a test may exert on people, policies, or practices within the classroom, school, educational system, or society (Bachman, 2005). The results from the interviews about students' and teachers' perceptions of the effects experienced after utilizing the PTW rubrics are entirely favorable. The most significant influence is observed on the students themselves and the classroom procedures, where self-assessment can be implemented. Three ways exist to evaluate the impact of rubric utilization.

Consequently, the familiar structure of the PTW rubric provides a sense of reassurance and continuity, allowing students to engage more confidently with the self-assessment process. Students demonstrate the ability to empathize with their teacher's perspective while evaluating their essays, reflecting a culturally embedded respect for teacher authority. The data indicates that students had a positive experience throughout the evaluation process, and these favorable experiences shape their perceptions of rubric-based assessment, as reflected in their interview responses. To be more specific, the influence of the rubrics on students was evaluated based on 3 different techniques.

First, their test-taking experience influences their view of the TLU domain (Bachman & Palmer, 1996). The interviews reveal that the direction offered by the PTW rubrics instills in students a sense of having their teacher present, since it mirrors the teacher's evaluation criteria. Students possess the capacity to empathize with their teacher's emotions while evaluating their essays. The data indicates that students have positive experience in the evaluation process. These favorable experiences have shaped their perspectives and opinions, as seen by their input during the interview.

Additionally, the students emphasized during the interview that they possess sufficient confidence to conduct self-assessments utilizing the PTW rubrics. This has shaped their perspectives on self-evaluation exercise, which can be executed in the classroom, as they have positive experiences with the assessment procedure in this study. All students, regardless of skill levels, said throughout the interviews that they can independently utilize the rubrics, including the teacher.

Secondly, the type of feedback they receive regarding their test performance directly affects them (Bachman & Palmer, 1996). This self-assessment activity utilizing PTW rubrics provides rapid feedback from the rubric descriptors, and students' choices may directly influence their outcomes depending on the scores obtained post-test (Bachman & Palmer, 1996). This study reveals that upon completing their essay assessments and receiving their scores, students report an enhanced comprehension of the assessment task and the expectations of the question, leading them to recognize the necessary actions to improve their essay scores. Moreover, with direct feedback highlighting their strengths and faults, students can achieve higher marks or improve their results. Consequently, it is demonstrated that students benefit from utilizing the PTW rubrics.

Lastly, as for teachers, negative or positive washback effects are seen as impacting their teaching (Rathnayake, 2025). The teacher's impact on her teaching, whether bad or positive, is not explicitly indicated. Nonetheless, she mentions her limited knowledge of cohesive devices and expressed intent to explore the topic further, as she had only instructed her students to use basic exploration. It can be inferred that she will incorporate cohesive devices in her instruction, as they are utilized in the PTW rubrics. The teacher concurs that her students can utilize the rubrics to discover areas for improvement and asserts that implementing PTW rubrics in the classroom will enhance their essay writing skills.

Consequently, she states her intention to implement the rubric in her classroom. The teacher's explanation exemplifies positive washback, as the assessment process (writing and self-evaluation) fosters effective teaching and learning techniques, enhancing student performance. Positive washback pertains to the advantageous impacts of language assessment on the teaching and learning process (Bokiev & Abdul Samad, 2021). Thus, positive impacts are evident from using PTW rubrics in the classroom self-assessment activity.

5.2 The practicality of the PTW rubrics based on AUA

The practicality of rubrics in educational assessment is a topic of ongoing debate, primarily concerning their development from the developers' standpoint, as Vercellotti (2021) emphasized. The prevalent concerns regarding the practicality of rubrics pertain to their efficacy in administering examinations, considering factors such as financial resources, time limitations for administration and scoring, the availability of evaluators (Bachman, 2005), the alignment of rubric design with learning objectives, and the necessity for rubrics to delineate clear criteria and standards for assessment (Vercellotti, 2021). It has been recognized that there is a lack of studies investigating the practicality of rubrics from the students' perspective. This study aims to investigate the practicality of the PTW rubrics, as Bachman (2005) recommended in AUA, by examining their impact on the overall quality of student assessments.

This study's findings enhance the notion of AUA practicality, as Bachman (2005) emphasizes that AUA practicality is related to the test development process. However, within the context of this study, as it pertains to a self-assessment tool utilized in classroom activities, students' perceptions, as the main stakeholders, are evaluated to ascertain the practicality of these rubrics. Additionally, the study encompasses the teacher's viewpoint on this subject. The evidence of practicality was utilized to substantiate the rationale for this research topic. All students who participated in this study's interviews claim that the PTW rubrics are helpful and that they intentionally use them to enhance their work and grades. The findings reveal that PTW rubrics are practical as self-assessment tools in terms of their suitability to be used by all students regardless of their proficiency level, time efficiency, and as rater-friendly rubrics.

Rubrics are practical for all students when they offer a definitive set of criteria for evaluating work or performance. A key characteristic of the primary trait rubric, which concentrates solely on specific attributes (Davis, 2018; Veloo et al., 2018), is

its accessibility for students across all proficiency levels. Consequently, the participants concur that the rubrics can be fully utilized, including by low-proficiency students.

Additionally, the interviews reveal that students and teachers affirm the short grading time when utilizing the PTW rubrics to evaluate their essays. The rating method required minimal time, as evidenced during the rubric implementation phase, where students completed it in approximately 5-10 minutes. Despite the brevity of the grading period, it does not undermine the advantage of delivering immediate feedback to students. PTW rubrics are user-friendly, since participants, including the instructor, report that the rubrics are straightforward to use, feature a clear layout, and contain comprehensible terminology.

5.3 Previous rubric-based self-assessment in L2 writing and the Malaysian educational context.

Research on how rubric-based self-assessment is used in L2 learners' writing has outlined potential advantages and problems (refer to Section 2.1). Self-assessment promotes learner autonomy but is a challenge in practice as writing assessment is intrinsically subjective, involving the judgment of a range of linguistic features like content, organization, lexis, syntax, and cohesion (Knoch, 2021).

Teachers frequently face inconsistent judgments, and unreliable or invalid scores can be awarded to L2 writers who cannot accurately self-evaluate without explicit instruction, leaving them liable for false-positive or negative evaluations (Harsch et al., 2024). Furthermore, without validated and scale-aligned rubrics, self-assessment tools cannot capture the performance criteria that are task-specific or attend to a range of proficiency levels of L2 learners (Zhang & Zhang, 2022; Roche & Booth, 2021).

While most general rubrics contain broad descriptors and fail to make task-specific delineations, PTW rubrics pinpoint task-specific traits that align with the communicative intent and the writing task's linguistic characteristics. Such task specification enables L2 learners to concentrate on the core aspects of writing for valid and reliable self-assessment in alignment with CEFR levels. This study illustrates the impact and practicality of PTW rubrics, which meaningfully facilitate learner engagement and self-evaluation, demonstrating their potential to improve self-assessment in L2 writing classrooms in Malaysia.

The discussions on PTW rubrics' positive impact and practicality reflect the Malaysian educational context's features, which focus more on teacher-centered practices and an exam-oriented system. Learners are typically dependent on teacher evaluation and rarely engage in assessing their own work. The PTW rubrics are pivotal in bridging the gap by providing structured yet comprehensible criteria that empower students to take greater responsibility for their learning. Consequently, it aligns with Weigle's (2002) assertion that L2 learners require highly specialized criteria to evaluate their writing effectively. The evidence substantiates the rationale for the study question, demonstrating that PTW rubrics are practical for evaluating L2 learners' writing in self-assessment activities among Malaysian L2 learners in formative classroom

contexts. This is also vital to promote student-centered teaching and learning and shared responsibility among students and teachers (Kiang & Yunus, 2021), in line with its policy advocacy in Malaysia.

6. Limitation and Implication

There are two limitations of this paper. The first limitation concerns the PTW rubrics, which are restricted to specific writing performance tasks. This is due to the fact that primary trait scoring is highly precise, whereas the intended use in this study is for composing straightforward essays, such as emails and informal letters. Even though this enables researchers to concentrate exclusively on the effectiveness of the criteria in capturing the desired writing traits primarily for L2 learners, teachers cannot use the developed rubrics for all types of essays.

Another limitation is the participants in this study. Only six out of 149 students were chosen, and they were randomly picked based on their willingness and permission from their teachers after the self-assessment session was conducted. Two of the participants were students from low proficiency categories, and there might be an issue with the validity of the rating. The researcher tried to address that with rater training, hoping that it would help the low achievers understand the use of rubrics when rating their essays.

Meanwhile, the implications of the study can be seen in the field of second language teaching in Malaysia, the new curriculum, which is based on CEFR, leads to a struggle for English teachers as lessons focus more on student-centered. This study provides new insights into classroom-based assessment practices that teachers can use concerning self-assessment practices. In addition, it also introduces another strategy for evaluating second language students' writing skill performance assessment. The availability of PTW rubrics will be a guiding tool for teachers to train students for self-assessment, thus producing students who can be independent learners.

Training sessions for the interpretation and application of the PTW rubrics are needed for more effective implementation. It can be integrated into the writing lessons for the students, and in this study, it is proven that the rater training session has helped students use the PTW rubrics. Even though the English proficiency among students varies, teachers can adapt the PTW rubrics as they are developed for all levels of students. However, challenges may arise in scaling this approach as it relates to teacher time constraints and limited assessment literacy among students. Thus, professional development workshops and peer assessment activities can help sustain the effectiveness of PTW rubrics in all educational contexts.

7. Conclusion

In conclusion, PTW rubrics positively impacted student learning and are practical for incorporation into classroom evaluation methodologies. This work contributes to formative classroom writing assessment by demonstrating that PTW rubrics enhance clearer instructional guidance and structured assessment procedures for teachers. Since the rubrics serve as a practical tool, teachers are able to align their feedback with explicit performance criteria, which ensure consistency and

transparency in evaluating students' writing. Meanwhile, the specificity of the PTW rubrics supports teachers in scaffolding instruction, such as rater-training. Meanwhile, the rubrics foster learner autonomy by encouraging students to practice self-assessment, reflect on their writing performance, and take responsibility for their learning progress.

Since this study has highlighted the practicality and positive impact of the PTW rubrics on Malaysian L2 learners' self-assessment activities, future research can employ a mixed-methods research design. This will provide more comprehensive data to enhance further understanding of the effectiveness of PTW rubrics. Quantitative data will provide empirical evidence, and if this kind of data is integrated with qualitative data on learners' writing performance, a more nuanced perspective can be gained. This is in terms of how PTW rubrics influence learners' engagement, motivation, and self-assessment development.

8. Disclosure statement

No potential conflict of interest was reported by the author

9. Acknowledgments

The authors wish to acknowledge the use of Grammarly in writing this paper. This tool was used to help improve the language and grammar in the paper. In addition, the use of Consensus is also a great help in finding the related studies for this paper. The paper remains an accurate representation of the author's work and intellectual contributions.

10. References

- Adnan, N., & Sayadi, S. (2021). ESL students' readiness for self-directed learning in improving English writing skills. *Arab World English Journal*, 12(4), 503-520. <https://doi.org/10.24093/awej/vol12no4.33>
- Aldosari, M., & Alsager, H. (2023). A step toward autonomy in education: Probing into the effects of practicing self-assessment, resilience, and creativity in task supported language learning. *BMC Psychology*, 11. <https://doi.org/10.1186/s40359-023-01478-8>
- Andrade, H. (2019). A critical review of research on student self-assessment. *Frontiers in Education*. <https://doi.org/10.3389/feduc.2019.00087>
- Anh, T. T. N., Nguyen, N. T., & Nguyen, A. T. T. (2023). The impact of online Self-Assessment on learning outcomes and Self-Assessment skills among Grade 11 students in Vietnam. *International Journal of Learning Teaching and Educational Research*, 22(4), 21-35. <https://doi.org/10.26803/ijlter.22.4.2>
- Asli, N. F., Mohd Matore, M. E. E., & Mohd Yunus, M. (2024). Construct validity of primary trait writing rubrics based on assessment use argument (AUA) validation framework. *Heliyon*, 10(22), 1-19. <https://doi.org/10.1016/j.heliyon.2024.e40053>
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2(1), 1-34. https://doi.org/10.1207/s15434311laq0201_1
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford University Press.
- Bokiev, U., & Abd Samad, A. (2021). Washback of an English language assessment system in a Malaysian university foundation programme. *The Qualitative Report*, 26(2), 555-587. <https://doi.org/10.46743/2160-3715/2021.4349>
- Butler, Y. G. (2023). Self-assessment in second language learning. *Language Teaching*, 57, 42-56. <https://doi.org/10.1017/S0261444822000489>

- Chambers, A. W., & Harkins Monaco, E. A. (2023). Increasing student engagement with self-assessment using student-created rubrics. *Journal of the Scholarship of Teaching and Learning*, 23(2), 96–99. <https://doi.org/10.14434/josotl.v23i2.33715>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Creswell, J., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage Publications.
- Davis, L. (2018). Analytic, holistic, and primary trait marking scales. In J. I. Liontas (Eds.), *The TESOL encyclopedia of English language teaching* (pp. 1–6). John Wiley & Sons. <https://doi.org/10.1002/9781118784235.eelt0365>
- Dosi, I., & Douka, G. (2021). Effects of language proficiency and contextual factors on second language learners' written narratives: A corpus-based study. *International Journal of Research Studies in Education*, 10(5), 1–18. <https://doi.org/10.5861/ijrse.2021.5076>
- Ebrahimi, M., Izadpanah, S., & Namaziandost, E. (2021). The impact of writing self-assessment and peer assessment on Iranian EFL learners' autonomy and metacognitive awareness. *Education Research International*, 1–12. <https://doi.org/10.1155/2021/9307474>
- Fachrunnisa, N., Salija, K., & Sahril, S. (2024). Students' experiences with self-assessment in improving English writing. In *Proceedings of the Hybrid National Conference (HNC 2024)* (Vol. 1, No. 1). E-Proceeding FKIP Universitas Bosowa. <https://doi.org/10.31850/ncp.v1i1.1212>
- Foley, J. A. (2019). Issues on the initial impact of CEFR in Thailand and the region. *Indonesian Journal of Applied Linguistics*, 9, 359–370. <https://doi.org/10.17509/ijal.v9i2.20233>
- Franchis, C., & Mohamad, M. (2023). Implementation of ESL assessment in Malaysian schools: A systematic literature review paper. *International Journal of Academic Research in Progressive Education and Development*, 12(2), 2311–2326. <http://dx.doi.org/10.6007/IJARPED/v12-i2/17537>
- Hajan, B. H., Castillo-Hajan, B., Artes, R., Jr, Jackaria, P. M., Mastul, A. H., Ayangco-Derramas, C., & Sali, F. Z. V. (2024). Analysis of the writing quality of senior High school students' Research Introductions. *International Journal of Learning Teaching and Educational Research*, 23(5), 219–238. <https://doi.org/10.26803/ijlter.23.5.12>
- Harsch, C., Koval, V., Kanistra, P., & Delgado-Osorio, X. (2024). Validating an integrated reading-into-writing scale with trained university students. *Assessing Writing*, 62, 100894. <https://doi.org/10.1016/j.asw.2024.100894>
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527–535. <https://doi.org/10.1037/0033-2909.112.3.527>
- Kiang, N. H., & Yunus, M. M. (2021). What do Malaysian ESL Teachers Think About Flipped Classroom? *International Journal of Learning Teaching and Educational Research*, 20(3), 117–131. <https://doi.org/10.26803/ijlter.20.3.8>
- Knoch, U. (2021). Assessing writing. In G. Fulcher & L. Harding (Eds.), *The Routledge handbook of language testing* (pp. 297–308). Routledge. <https://doi.org/10.4324/9781003220756-19>
- Lallmamode, S. P. (2012). *The development of an argument-based validation of an analytic rubric for L2 writing electronic portfolios* (Doctoral dissertation). International Islamic University Malaysia.
- Rathnayake, W. (2025). From Tests to Teaching: Exploring the Washback Effect in Language Pedagogy. *Journal of Desk Research Review and Analysis*. <https://doi.org/10.4038/jdrra.v2i2.47>

- Roche, T., & Booth, S. (2021). A collaborative approach to assuring standards: Using the CEFR to benchmark university pathway programs' English language outcomes. *Language Teaching Research Quarterly*, 26, 18–38. <https://doi.org/10.32038/ltrq.2021.26.02>
- Sun, Y., & Lan, G. (2024). Translingual practices in Chinese as a second language writing: Variations across language proficiency levels. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/13621688241299906>
- Vasileiadou, D., & Karadimitriou, K. (2021). Examining the impact of self-assessment with the use of rubrics on primary school students' performance. *International Journal of Educational Research Open*, 2, 100031. <https://doi.org/10.1016/j.ijedro.2021.100031>
- Vasiljevic, Z. (2023). The impact of the CEFR on foreign language education in Japan. *Journal of International Scientific Publications: Educational Alternatives*, 21, 270–288. <https://doi.org/10.62991/EA1996152708>
- Veloo, A., Aziz, N. H. A., & Yaacob, A. (2018). The most suitable scoring method to assess essay writing in ESL classrooms. *Advances in Language and Literary Studies*, 9(4), 19–25. <https://doi.org/10.7575/aial.v.9n.4p.19>
- Vercellotti, M. L. (2021). Beyond the rubric: Classroom assessment tools and assessment practice. *The Electronic Journal for English as a Second Language*, 25(3), 1–16.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
- Weigle, S. C. (2010). Scoring procedures for writing assessment. In J. C. Alderson & L. F. Bachman (Eds.), *Assessing writing* (pp. 116–141). Cambridge University Press.
- Weir, C. (2005). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan.
- Zhang, X. S., & Zhang, L. J. (2022). Sustaining learners' writing development: Effects of using Self-Assessment on their foreign language writing performance and rating accuracy. *Sustainability*, 14(22), 14686. <https://doi.org/10.3390/su142214686>

Appendix 1

Interview Questions (Students)

1. What is your opinion on self-assessment rubrics?
2. Do you ever use any rubric?
3. Which rubric is quite challenging when you want to mark your essay?
4. Do you think that the rater training session helps you to understand better?
5. Do you think the rubrics can help you to improve your English?
6. Can you understand all the words used in the rubric?
7. Do you think you can use the rubrics on your own?
8. Do you think the rubric has helped you understand the assessment requirements or what is required of you?
9. Do you think you will have any problem using the rubric on your own?
10. Do you feel confident using the rubrics?
11. Do you think it helps you to understand how your teachers mark your essay?
12. Do you think you can practice writing your essay using rubrics?
13. Are the rubrics practical for students to use?
14. Do you think all students with all proficiency levels can use the rubric?
15. Can the rubric be used as material in the classroom?
16. Do you have any suggestions for improvement?

Appendix 2

Interview Questions (Teacher)

1. What is your opinion on the primary trait of writing self-assessment rubrics?
2. Do you ever use a primary trade rubric to mark your student's essays?
3. Which trait do you think is difficult to mark?
4. Which criteria did you find less suitable?
5. Which criteria did you find suitable?
6. Do you think that the training session helps students to use the rubric?
7. How do you think rubric can help students to improve their writing?
8. Do you think the student can understand the vocabulary used in the rubrics?
9. Do you think students can use the rubric on their own?
10. Do you think rubric can help students understand the assessment requirements?
11. Do you think the students will have any problem using the rubric?
12. Are the rubrics practical for use by the students?
13. Do you think it requires a lot of time to use the rubric?
14. Do you think students with all levels of proficiency can use the rubrics?
15. Would you use rubric as one of your materials in class?
16. Do you think that students will have the confidence to use the rubric on their own?
17. Do you think the rubrics help students to understand how teachers mark their essays?
18. Will the rubrics have a positive impact on your students?

Appendix 3

Rubric Samples

PRIMARY TRAIT WRITING RUBRICS

RUBRIC 1

Trait	Scores and Scoring Criteria				
	4	3	2	1	0
Content	I can respond to the questions with well-detailed elaboration (3 supporting details).	I can respond to the questions with detailed elaboration (2 supporting details).	I can respond to the questions with elaboration (1 supporting detail).	I can respond to the questions with no elaboration (no supporting detail).	I do not respond to the questions at all.

Score

RUBRIC 2

Format of an email
 1-Subject
 2-Salutation
 3-Body of the email
 4-Signature

Trait	Scores and Scoring Criteria				
	4	3	2	1	0
Format	I write the email using 4 elements of the format correctly.	I write the email using 3 elements of the format correctly.	I write the email using 2 elements of the format correctly.	I write the email using 1 element of the format correctly.	I write the email without using the format.

Score