


Arab World English Journal (AWEJ) Volume 15. Number 1, March 2024  
DOI: <https://dx.doi.org/10.24093/awej/vol15no1.19>

Pp.298-313

## Malaysian Students' Perceptions towards Using Peer Feedback to Cultivate Evaluative Judgement of Argumentative Writing

Xiao Xie 

Faculty of Modern Languages and Communication, Universiti Putra Malaysia, Selangor, Malaysia.

Vahid Nimehchisalem 

Faculty of Modern Languages and Communication, Universiti Putra Malaysia, Selangor, Malaysia.

Corresponding Author: [vahid@upm.edu.my](mailto:vahid@upm.edu.my)

Mei Fung Yong 

Faculty of Modern Languages and Communication, Universiti Putra Malaysia, Selangor, Malaysia.

Ngee Thai Yap 

Faculty of Modern Languages and Communication, Universiti Putra Malaysia, Selangor, Malaysia.

Received:11/22/2023

Accepted:03/1/2024

Published: 03/20/2024

### Abstract

Although peer feedback is a prevailing practice to promote evaluative judgement, its influence on the development of this higher-order cognitive ability has not yet been adequately explored. Specifically, there is a dearth of research that examines the benefit of providing and receiving peer feedback in developing students' understanding of assessment standards of writing. The purpose of this study is to explore students' perceptions of how different feedback roles influenced their three types of evaluative judgement of ESL argumentative writing, namely hard, soft and dynamic evaluative judgement. During five weeks, 24 undergraduate students enrolled in an English argumentative writing course at a Malaysian public university were randomly assigned to three distinct peer feedback roles, namely feedback provider, feedback receiver, or feedback outsider, to participate in the peer feedback activities. Thematic analysis of pre- and post-intervention surveys indicated that different feedback roles varied in facilitating the development of evaluative judgement. Despite the limitation of domain-specific knowledge, strategically integrating peer feedback into writing course design afforded students opportunities to cultivate the three types of evaluative judgement. This study translates the theoretical framework of evaluative judgement into identifiable goals within the course of English argumentative writing and sheds light on the cognitive mechanisms inherent in different feedback roles, which enables educators and researchers to better dissect peer feedback curriculum design and student-centred assessment activity.

**Keywords:** Argumentative writing, evaluative judgement, feedback roles, Malaysian students' feedback, peer feedback, writing domain

**Cite as:** Xie, X., Nimehchisalem, V., Yong, M.F., & Yap, N.T. (2024). Malaysian Students' Perceptions towards Using Peer Feedback to Cultivate Evaluative Judgement of Argumentative Writing. *Arab World English Journal*, 15(1): .298-313. DOI: <https://dx.doi.org/10.24093/awej/vol15no1.19>

## Introduction

Evaluative judgement is defined as “the capacity to make decisions about the quality of work of oneself and others” (Tai et al., 2018, p. 467). Its significance extends beyond the scope of individual courses to the realm of lifelong learning (Boud & Soler, 2016), because frequently disregarded in university assessments is the incorporation of direct, authentic experiences for students to evaluate both others and themselves (Tai et al., 2016). This inclusion assists students in comprehending the intricate process of forming judgement, as highlighted by Nicol and Macfarlane-Dick (2006), making them more effective learners and meeting the demands of the postgraduate workplace (Boud, 2000).

The three distinct types of evaluative judgement (Nelson, 2018), adopting a historical perspective, await further validation through empirical evidence. The first is termed “hard evaluative judgement”, involving an objective assessment of correctness in paradigmatic features. The second, “soft evaluative judgement”, is rooted in the evaluation of value and quality, with the emphasis not placed on right or wrong but on appraising the significance of elements. The third, “dynamic evaluative judgement”, pertains to how content is handled and the organisation of thoughts. Given that evaluative judgement is not universally applicable but depends on specific standards, individuals acquire expertise in particular disciplines or subject areas. In doing so, they can effectively assess the quality of work within that field.

Peer feedback is considered to be important for the development of evaluative judgement (Nicol et al., 2014), and both providing feedback and receiving feedback are crucial for learning (Tai et al., 2018). Nevertheless, in the context of ESL/EFL writing, extensive research over the past three decades has primarily focused on reviewing the impact of various peer feedback methods on overall writing improvement (Gao et al., 2019) and investigating students' and educators' attitudes and experiences with peer feedback (Chang, 2016). As noted by Vasu et al. (2016), considering the educational context of Malaysia, where teachers are often expected to be accountable for the learning of their students, “peer feedback, although highly valued, turned out to be the least preferred in this context despite the current pedagogical trend that focuses more on students' active participation in improving their peers' writing than being dependent on teachers' feedback” (p. 164). It is also confirmed that this teacher-led-unidirectional feedback style places learners in a passive role, often resulting in a lack of understanding of peer feedback comments or difficulty in applying strategies to improve their learning (Boud & Molloy, 2013; Hattie & Timperley, 2007). Thus, peer feedback practices are still little understood in terms of the pedagogical implications of fostering and promoting learners' engagement affectively, behaviourally, and cognitively (Cheng & Zhang, 2024; Gan et al., 2023). Furthermore, it is important to take an in-depth look at the benefits and challenges associated with providing and receiving peer feedback as a means of encouraging students' proactive engagement in peer feedback (Tai et al., 2018). Despite this, little research has been conducted to determine what learning benefits each of these feedback roles can provide in the context of ESL/EFL writing (Allen & Katayama, 2016; Gao et al., 2023; Huisman et al., 2018).

Although evaluative judgement is becoming increasingly relevant in higher education assessment literature, there is still a lack of research showing how peer feedback impacts students' understanding of assessment standards in ESL argumentative writing. For peer feedback to be effective in Malaysian higher education, where most feedback discussions are typically teacher-led-unidirectional and seldom result in follow-up actions (Carless & Boud, 2018; Vasu et al., 2016), the current study aims to add empirical evidence regarding benefits and challenges of

providing feedback and receiving feedback to the student-centred assessment research base. Therefore, the research objective is to examine students' perceptions regarding different feedback roles' impact on their development of evaluative judgement of ESL argumentative writing. Specifically, this study is guided by the following research questions:

RQ1: How do different peer feedback roles promote students' development of hard evaluative judgement in argumentative writing?

RQ2: How do different peer feedback roles promote students' development of soft evaluative judgement in argumentative writing?

RQ3: How do different peer feedback roles promote students' development of dynamic evaluative judgement in argumentative writing?

## Literature Review

In a previous theoretical effort, Tai et al. (2016) put forward the proposition that evaluative judgement comprises two interrelated and harmonious components. The first component involves grasping the essence of what constitutes quality, while the second component involves applying this understanding to evaluate various works, whether one's own or others. Strategies for cultivating evaluative judgement, and adopting a social constructivist viewpoint (Vygotsky, 1978), include the use of rubrics, exemplars, self-assessment, peer assessment, and feedback (Tai et al., 2018).

While this perspective may seem innovative, its roots can be traced back to Sadler's (1989) exploration of formative assessment in shaping students' evaluative knowledge and evaluative expertise. Emphasising how students benefit from discerning the quality of their own and peers' work, there is a growing body of research explicitly placing themselves within the innovative framework of evaluative judgement. More researchers have provided detailed explanations of alternative theoretical perspectives concerning the conceptualisation of evaluative judgement. These perspectives encompass social material viewpoints (Ajjawi & Bearman, 2018), historical standpoints (Nelson, 2018), dual-process methodologies (Joughin, 2018), and epistemological outlooks (Goodyear & Markauskaite, 2018). Luo and Chan (2023) asserted that evaluative judgement is not a fixed feature possessed or lacking in students but rather a multidimensional structure encompassing knowledge, attitudes, abilities, behaviours, and identity-related aspects. This not only advocates for a more process-oriented approach to cultivating students' evaluative judgement skills in the curriculum but also emphasises the need for an integrated course design to address the necessity of emphasising various aspects of evaluative judgement and their synergistic relationships.

In a recent empirical study, Chong (2021) examined the use of exemplars in the preparation for the IELTS test, based on Nelson's (2018) classification of evaluative judgement. The study examined teacher-student dialogue patterns and university students' perceptions about using exemplars to improve their understanding of assessment criteria for academic writing tasks. The study indicated that teachers/researchers in IELTS testing employ various interaction strategies to cultivate students' abilities in the hard, soft, and dynamic dimensions of evaluative judgement. Chong (2021) pointed out that hard evaluative judgement "concerns areas involving grammatical (e.g., the use of correct verb tenses) and mechanical accuracy (e.g., the use of correct punctuation)"; soft evaluative judgement "is demonstrated through students' use of a variety of vocabulary and sentence patterns"; and dynamic evaluative judgement is about "how one communicates with the audience of the work through presentation of logical and structured ideas"

(p. 5). The results of the study indicated that concerning hard evaluative judgement, i.e., focusing on the correctness of vocabulary and grammar features in writing, students seemed to pay more attention to the correct use of sentence-level features rather than word-level features. As students needed additional types of expertise, such as grammatical knowledge, to fully comprehend this requirement related to language correctness, using exemplars alone had a less effective impact on students' development of hard evaluative judgement. In the development of students' soft evaluative judgement, more students focused on the importance of using various vocabularies than the importance of using various sentence structures, differing from students' hard evaluative judgement. Among these three types of evaluative judgement, students demonstrated robust declarative knowledge in the dynamic evaluative judgement of IELTS writing tasks. They were proficient in articulating various aspects of content and organisational requirements for English academic writing tasks.

Besides the aforementioned use of exemplars (Chong, 2021), in the context of second-language writing, learners often engage in providing suggestions to their peers and receiving comments on their writing performance in pairs or groups through written or oral means (Yu & Lee, 2016). As emphasised by Hyland and Hyland (2006), the central goal of this feedback practice is to cultivate students' autonomy as authors. Effective evaluative judgement requires dynamic interaction among individuals, their peers, and established performance benchmarks to clarify the characteristics of learners' outstanding performance, compare their current performance to these standards, and determine what they can do to improve (Hattie & Timperley, 2007).

As a prevailing practice, peer feedback possesses four key features that render it a suitable educational approach for developing students' evaluative judgement (Nicol, 2014). Firstly, by directly engaging students in multiple evaluative judgement behaviours through reviewing peers' work, peer feedback allows students to transfer thoughts generated through this comparative process to their work. Secondly, peer feedback provides a platform for interacting with standards and norms, helping students not only develop and interpret concepts of quality but also formulate their standards and norms. Thirdly, for feedback providers, explaining the feedback contributes to enhancing and building students' own knowledge and understanding, suggesting that generating feedback may be more beneficial to students' knowledge generation than receiving feedback. This is because, in the review process, feedback providers elevate their reader awareness (Chang, 2015; Tsui & Ng, 2000), enhance understanding of global issues in writing (Min, 2005), and promote reflection on personal writing (Nicol, 2014). Finally, for feedback receivers, accepting different viewpoints fosters a culture of constructive criticism, enabling participants to cultivate a growth mindset and accept opportunities for improvement. The key is how students interact with and respond to the feedback received (Hattie & Timperley, 2007; Kim, 2009; Nicol & Macfarlan-Dick, 2006).

As shown in previous research, learning mechanisms involved in providing peer feedback include initiating problem detection, encouraging problem diagnosis, and modifying strategies (Patchan & Schunn, 2015). On the contrary, the learning mechanisms associated with receiving peer feedback encompass acquiring information about current performance and anticipated performance, as well as developing strategies to bridge this gap (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006). The traditional understanding of peer feedback has mainly focused on the advantages gained by students receiving peer feedback. However, a significant shift in perspective is emerging, suggesting an awareness that students taking on the role of feedback providers are now the primary beneficiaries of this practice (Nicol et al., 2014; Tai et al., 2016;

Xie et al., 2022). Students providing feedback engage in a cognitive process that involves advanced thinking. This process encompasses complex psychological operations, such as applying assessment standards, identifying potential issues in the reviewed work, and formulating viable solutions to address these issues. However, reviewing the results of feedback literature and empirical cases, there is limited research focused on comparing the learning performance and cognitive effects of different roles in the peer feedback process (Allen & Katayama, 2016; Gao et al., 2023; Huisman et al., 2018). As far as the authors of this study are aware, there has been no research comparing the effects of different peer feedback roles on students' evaluative judgement of ESL argumentative writing.

## Method

### *Participants*

The investigation transpired throughout the second semester of the academic year 2022/2023, spanning from week two to week six, encompassing five weeks. Conducting a convenience sampling method, 24 local students were recruited from a public university in Malaysia for an undergraduate writing course, mainly females aged between 19 and 21. They had achieved at least Band Three on the Malaysian University English Test (MUET), demonstrating a reasonably fluent and fairly appropriate use of the English language, despite numerous grammatical errors. MUET is predominantly utilised for university admissions in Malaysia, with nine bands ranging from Band Five as the highest to Band One as the lowest. Their bands mainly correspond to levels B1 and B2 of the Common European Framework of Reference (CEFR). Hence, it can be deduced that these participants possessed an intermediate level of proficiency in English, enabling them to generate clear, detailed texts on a broad spectrum of subjects. They were also capable of articulating viewpoints on topical issues, presenting the advantages and disadvantages of various options.

These participants were randomly assigned to three different groups, each given specific roles during the feedback process. These roles included feedback providers, feedback receivers, or feedback outsiders. The feedback providers were responsible for offering insights, suggestions, and constructive feedback to help the receivers improve their writing. Simultaneously, feedback receivers had the opportunity to reflect on the feedback provided by peers. This reflective process was crucial for participants to understand their own writing performance and appreciate different perspectives on their work. The feedback outsider would not actively engage in providing comments on their peers' writing essays nor receive feedback on their own work. Each participant was informed of their assigned role through email. This strategic role allocation aimed to achieve a dual purpose: safeguarding the privacy of each participant and fostering an open and supportive environment within the learning framework. Ethical approval was granted by the Ethics Committee for Research Involving Human Subjects (JKEUPM) at Universiti Putra Malaysia (reference number: JKEUPM-2023-411).

### *Research Instruments*

Researchers utilised the evaluative criteria developed by Aryadoust (2012) as the chosen tool and presented various examples to the participants. The criteria effectively delineated five domains of argumentative writing and provided guidelines for subsequent peer feedback activities, consistent with Weir's (1990) interpretation of the five major attributes of writing patterns: Relevance and Adequacy of Content (RAC), Compositional Organisation (CO), Coherence (C),



Vocabulary (V), and Grammar (G). This study employed the categorisation outlined by Chong (2021), wherein the assessment of the precision of Grammar (G) and Vocabulary (V) is classified as hard evaluative judgement; the assessment of the range and scope of Grammar (G) and Vocabulary (V) falls within soft evaluative judgement; and the evaluation of the Relevance and Adequacy of Content (RAC), Compositional Organisation (CO), and Coherence (C) is categorised under dynamic evaluative judgement.

In response to the introduction of the evaluative criteria, the participants were required to complete three writing tasks and to submit a written response addressing each writing prompt in 300-500 words: (1) The role of music in bringing together people of different cultures and age groups; (2) The potential obsolescence of print newspapers and books due to the availability of online reading; (3) Social and practical challenges of living in a foreign-language-speaking country. As part of the research tools, two writing exemplars representing different proficiency levels were created for each prompt. Based on these exemplars, researchers conducted classroom discussions and analysis of the five domains of the evaluative criteria.

### ***Research Procedures***

During this five-week training programme, two distinct sessions, each lasting 1.5 hours, were conducted every week. This structure ensured that each participant received a total of 15 hours of training in the classroom. This study embraced a Process-Genre Approach (Huang & Zhang, 2022; Rahimi & Zhang, 2022), incorporating a strategically arranged set of activities to enhance participants' experiences with peer feedback. The Process-Genre Approach aims to cultivate student writers' understanding of linguistic features, rhetorical structures, and writing skills in prewriting, drafting, revising, and editing, empowering them to craft texts that serve specific communicative purposes within a particular genre.

The training commenced with an introductory session aimed at acquainting participants with the evaluative criteria (Aryadoust, 2012). Subsequently, each participant independently completed and submitted the writing tasks. In the following session, the principal researcher conducted a classroom discussion and analysis of the five domains of the evaluative criteria for each of the three writing tasks. This analysis was based on two exemplars of writing, representing different proficiency levels. During the peer feedback phase, each peer feedback provider put forward revision suggestions for the essays of four feedback receivers for each task. To maintain the quality of feedback, the principal researcher explicitly instructed the providers to clarify identified issues, provide reasons for modification, and offer suggestions for improvement. Upon receiving the peer feedback outcomes, the principal researcher discreetly emailed the feedback comments to the intended receivers, maintaining anonymity. This process aimed to assist the feedback receivers in engaging in deep introspection regarding their writing performance concerning the specific writing tasks, while simultaneously safeguarding their privacy. Ultimately, this pedagogical design aimed to promote an environment conducive to collaborative learning, encouraging constructive feedback exchange without concerns about personal sensitivity.

### ***Data Collection***

A pre-intervention survey was carried out at the commencement of the training to gather participants' initial views and expectations regarding peer feedback, as well as their comprehension of the five domains of writing. Similarly, a post-intervention survey was undertaken at the culmination of the five-week writing training to collect participants' reflections

and insights following their involvement in peer feedback activities. This involved exploring the challenges and difficulties encountered by participants in assessing the five domains of argumentative writing and revealing the impact of different feedback roles on the development of their evaluative judgement.

In the first and fifth weeks of training, the principal researcher administered pre- and post-intervention surveys to all participants using the Google Forms online platform. By utilising the online survey, participants could express their thoughts and opinions in a more comfortable and confidential environment, free from external influences. The Google Forms online platform provided participants with a user-friendly interface, allowing them to complete the survey at their convenience. Participants could reflect on their experiences, articulate their thoughts comprehensively, and provide valuable feedback in a systematic and structured manner, thereby gaining a more holistic understanding of their perspectives and experiences.

### ***Data Analysis***

Regarding data analysis, this study employed a thematic analysis approach, harmoniously combining inductive and deductive coding methods, aiming to comprehensively understand identified themes (Fereday & Muir-Cochrane, 2006). By adopting both inductive and deductive coding techniques, the analysis process involved exploring newly emerging themes and applying them to existing theoretical frameworks or concepts, ensuring a thorough examination of the data. The researchers selected Nelson's (2018) evaluative judgement framework and an established evaluative judgement template as predefined basic dimensions for survey design and subsequent analysis (Chong, 2021).

Qualitative data for this study were analysed using NVivo (1.5.1) software. To enhance the credibility and reliability of the reported qualitative data, the initial phase of thematic analysis was conducted by the principal researcher who diligently perused the dataset, recording initial impressions. Once preliminary ideas were discerned, the principal researcher proceeded to code interesting features across the entire dataset, participating in discussions with other researchers to refine the coding and organise the data associated with each code. Three themes were substantiated by vivid and robust extracts from the data, grounded in both semantic and latent content. These themes encompassed the usefulness of peer feedback on evaluative judgement, reflection on writing knowledge construction, and evaluative judgement on different writing domains. Additionally, given the adoption of Nelson's framework (2018) in this study, the subsequent results are categorised according to the classification of hard, soft, and dynamic evaluative judgement.

### **Results**

Research Question One pertains to the hard evaluative judgement of the accuracy of Vocabulary (V) and Grammar (G) domains in writing. In the pre-intervention survey, concerning the Grammar (G) domain, eight participants underscored the significance of clear and accurate grammar. Furthermore, six participants highlighted the necessity of using appropriate tenses based on the selected topic (Participant No. 8). This demonstrates their awareness of the importance of accurately employing tenses in writing to maintain consistency and coherence. Additionally, two participants stressed the completeness of sentence elements and the diversity of syntactic structures (Participant No. 16). This indicates their understanding of the importance of constructing well-formed syntactic structures to enhance the overall quality of argumentative essays.

"I believe grammar features can adopt past, present, and future tenses based on the selected topic." (Participant No. 8, Feedback Provider)

"Ensure sentences are complete and comprehensive, with no fragments that leave the reader searching for the intended meaning." (Participant No. 16, Feedback Receiver)

In the post-intervention survey, in the evaluation of the Grammar (G) domain, some feedback providers mentioned that their own grammar limitations had impeded them from making accurate judgement (Participant No. 1). This quote reflects the participants' struggle in identifying grammatical errors due to a limited understanding of grammar rules. While feedback receivers generally acknowledged the benefits of receiving feedback, Participant No. 16 noted that receiving incorrect feedback disrupted their evaluative judgement. This excerpt underscores the crucial role of teachers in actively supervising and guiding feedback exchange to ensure learners receive accurate and helpful input. Teachers can provide explanations, examples, and additional resources to reinforce learners' understanding of the feedback they receive.

"The grammar section was the most challenging for me because my grammar is poor, and I can't identify the author's grammar errors." (Participant No. 1, Feedback Provider)

"My grammar mistakes were also wrongly corrected. Some feedback providers gave me incorrect corrections, making me doubt what is right and what is wrong." (Participant No. 16, Feedback Receiver)

In the pre-intervention survey, regarding hard evaluative judgement, few students paid attention to accuracy in the Vocabulary (V) domain and seldom mentioned the focus on the three vocabulary-level language features highlighted in the evaluation criteria: collocations, parts of speech, and spelling, demonstrating deficiencies in their lexical knowledge. In the post-intervention survey, in the Vocabulary (V) domain, the most frequently mentioned challenge among participants was related to their vocabulary limitations, highlighting the dilemma of accurately assessing vocabulary use due to personal knowledge limitations and encountering unfamiliar word choices (Participant No. 3). This statement indicates that participants found providing feedback on vocabulary choices challenging because they were uncertain about suitable alternatives for incorrect words.

"Determining the correct vocabulary choice is a challenge for me because I am unsure of words that can replace, even if I can identify the selected incorrect word." (Participant No. 3, Feedback Provider)

Research Question Two is related to the development of participants' soft evaluative judgement regarding the range of Vocabulary (V) and Grammar (G) domains. None of the participants mentioned syntax diversity in their pre- and post-intervention surveys, indicating their lack of awareness of grammatical range, which needs to be addressed by the writing lecturer.

In the pre-intervention survey of the Vocabulary (V) domain, participants had different views on vocabulary diversity use. Almost half of the participants (10 out of 24) believed that more advanced and complex vocabulary should be used to show a higher degree of lexical knowledge (Participant No. 13), while six participants mentioned the importance of using simple yet precise vocabulary to effectively convey meaning. Another consideration raised by some participants is to consider the expected audience to determine the level of vocabulary difficulty (Participant No. 2).

"Maybe advanced vocabulary skills, like using rare words with the same meaning." (Participant No. 13, Feedback Receiver)

"It depends on the audience; your vocabulary range must suit your audience." (Participant No. 2, Feedback Provider)



Interestingly, in the post-intervention survey, one of the feedback providers, Participant No. 1, elaborated on her significant development in understanding the Vocabulary (V) domain through the peer feedback process. This indicates that she became more aware of the importance of selecting vocabulary that effectively conveys ideas rather than pursuing novelty.

“In the first week, I thought, in the vocabulary domain, the author should use uncommon or unusual vocabulary. But later I shifted my focus from using strange vocabulary to using appropriate, understandable, and not repetitive vocabulary.” (Participant No. 1, Feedback Provider)

Research Question Three is related to participants' dynamic evaluative judgement, concerning three domains of the evaluative criteria, namely, Relevance and Adequacy of Content (RAC), Compositional Organisation (CO), and Coherence (C). The results are presented in these three domains accordingly below.

Firstly, in the pre-intervention survey, participants demonstrated the highest level of clarity in their awareness of the Relevance and Adequacy of Content (RAC) domain. The majority of participants (21 out of 24) could articulate specific requirements and expectations related to this domain, emphasising the importance of in-depth discussion of the topic with strong reasoning and arguments (Participant No. 8).

“In my view, a good argumentative essay needs an in-depth discussion of the topic with solid reasoning and argument support.” (Participant No. 8, Feedback Provider)

Nevertheless, in the post-intervention survey, concerning the Relevance and Adequacy of Content (RAC) domain, some feedback providers acknowledged encountering diverse challenges when making evaluative judgement in this regard. The author's subjectivity introduced difficulties in reaching consistent and fair evaluation (Participant No. 7). Ambiguities in arguments and supporting evidence posed challenges for participants in evaluating the relevance of content. The presence of conflicting viewpoints in the essays further complicated the determination of the overall coherence and persuasiveness of the argument (Participant No. 3).

“The perception of relevance and accuracy can vary from person to person, making it difficult to provide objective feedback.” (Participant No. 7, Feedback Provider)

“The opinion is given on one side, but the supporting sentences are leaning towards the opposite. Apart from that, some of the essays are also straying from the points or questions given.” (Participant No. 3, Feedback Provider)

In the post-intervention survey, eight feedback receivers generally conveyed that receiving constructive feedback from their peers improved their evaluative judgement with respect to the Relevance and Adequacy of Content (RAC) domain. This process aided them in pinpointing deficiencies in their writing. For instance, upon reflecting on the feedback, they identified occasions where they veered away from the topic and failed to adequately address the task, despite their initial perception that they had performed well (Participant No. 11).

“After receiving feedback, I realised I went off-topic, not really addressing the task. I thought I did well. I'll be more careful about the relevance and adequacy of my content.” (Participant No. 11, Feedback Receiver)

Secondly, within the Compositional Organisation (CO) domain, during the pre-intervention survey, half of the participants demonstrated a fundamental yet lucid understanding of the requisites in this domain. They underscored the significance of a well-structured English argumentative essay, encompassing elements such as an introduction, body paragraphs, and a conclusion (Participant No. 12).

“In my view, it must follow the standard format of an introduction, three separate argument points, and a conclusion.” (Participant No. 12, Feedback Receiver)

In the post-intervention survey, it emerged that students exhibited the highest level of competence in assessing the appropriateness of paragraphs within written pieces. The Compositional Organisation (CO) domain proved to be the most straightforward for the majority of participants to evaluate (Participant No. 4).

“You don’t even need to read them fully, just check if the CO is balanced. I mainly focus on each paragraph, key points, introduction, and conclusion.” (Participant No. 4, Feedback Provider)

Thirdly, 16 participants had difficulty explaining the assessment requirements about the Coherence (C) domain, showing a lack of domain-specific knowledge, including the use of cohesive devices, pronouns, and synonyms. Among them, nine participants’ responses were marked by confusion, as they mistakenly intermingled elements with those from the Relevance and Adequacy of Content (RAC) domain (Participant No. 2). Conversely, eight participants acknowledged the use of cohesive devices (Participant No. 13), indicating their rudimentary understanding of the importance of maintaining coherence in an essay.

“It connects the contents as it requires a lot of details and elaborations.” (Participant No. 2, Feedback Provider)

“You can use appropriate connectors and sentence linking words to connect the articles.” (Participant No. 13, Feedback Receiver)

In the post-intervention survey, nine participants openly acknowledged experiencing confusion when assessing the Coherence (C) domain. Despite participating in peer feedback activities, a sense of uncertainty persisted in their ability to distinguish correct cohesive devices (Participant No. 2). Furthermore, grappling with the broader concept of coherence presented challenges for rendering sound judgement within this domain.

“Evaluating coherence is quite challenging because I’m not very clear on this aspect. I cannot differentiate between correct or incorrect cohesive elements.” (Participant No. 2, Feedback Provider)

In addition to the aforementioned findings about the research questions, it is imperative to examine the impact of non-engagement in peer feedback on feedback outsiders’ evaluative judgement. During the post-intervention survey, there was a sense of stagnation and a lack of clarity regarding their learning progress (Participant No. 17, Participant No. 21, and Participant No. 23). Without the valuable insights and suggestions from peers, these participants felt that they lacked guidance in their writing improvement journey. They expressed a desire to receive feedback to identify areas for growth and to gain a better understanding of their writing strengths and weaknesses.

“I couldn’t learn anything, and my evaluating skill is still stagnant.” (Participant No. 17, Feedback Outsider)

“It makes it harder for me to write a better essay without any feedback.” (Participant No. 21, Feedback Outsider)

“I don’t know if I’m actually improving or not without any feedback. I want to know what I need to improve and what I am lacking in the writing.” (Participant No. 23, Feedback Outsider)

## Discussion

On Research Question One of hard evaluative judgement, this study found that the utilisation of peer feedback has limited effects on the students' advancement in evaluative judgement concerning the accuracy of grammar, morphology, syntax, or structure. This is related to the challenge students face in making clear explanations based on vocabulary and grammar rules, given their lack of specific domain knowledge in making hard evaluative judgement related to language correctness. In alignment with Chong's (2021) assertion that relying solely on exemplars is insufficient for fostering students' evaluative judgement, the outcomes of this study affirm that peer feedback itself can only partially improve students' domain-specific knowledge, such as accurate use of vocabulary and syntactic features, since participants only provided very general descriptions or directly expressed difficulty in making judgement on these domains. Therefore, although peer feedback is useful in developing students' understanding of evaluative criteria, teachers must develop students' domain-specific knowledge through other instructional tasks.

To address Research Question Two, regarding soft evaluative judgement, which involves students judging a variety of vocabularies and sentence structures, the current study found that students' understanding regarding the grammatical and lexical ranges varied. On the one hand, concerning the development of soft evaluative judgement in the Vocabulary (V) domain, after the peer feedback intervention, some participants shifted from a mere focus on the wide range of vocabulary diversity to utilising appropriate and contextually accurate vocabulary. On the other hand, this cohort of Malaysian participants tended to overlook syntactic diversity, exposing their deficiency in grammatical knowledge. Because of limited knowledge of the Grammar (G) domain, students' awareness of grammatical evaluative criteria did not translate into accurate soft evaluative judgement, as found by Chong (2021). This necessitates educators to devise specific pedagogical designs to effectively scaffold students' peer feedback; otherwise, students often fail to either proactively engage in providing peer feedback or incorporate the comments into revisions because the feedback they received is too vague or lacks specific suggestions.

Research Question Three is related to dynamic evaluative judgement. The most interesting conclusion drawn in this study is different from Chong's (2021) conclusion regarding students' strong declarative knowledge in dynamic evaluative judgement of IELTS writing. Chong (2021) attributed this to the higher implicitness of the requirements in content and organisation under dynamic evaluative judgement in the IELTS writing evaluation criteria, where students generate thoughts based on their previous experiences and thinking habits. This study found that students demonstrated relatively accurate judgement in the development of thought, and logical organisation of paragraphs, but had unclear and weak judgement in using cohesive devices, pronouns and synonym references. In accordance with the findings of Aryadoust (2012) and Weir (1990), the results of this study indicate that, for dynamic evaluative judgement related to the Relevance and Adequacy of Content (RAC) domain, it measures sociolinguistic knowledge by assessing responses related to the interactive or task-setting topic; meanwhile, the Compositional Organisation (CO) and Coherence (C) domains reflect discourse knowledge. There is a need to further distinguish the knowledge composition required for dynamic evaluative judgement in argumentative writing and require teachers to provide targeted guidance.

Both providing feedback and receiving feedback enabled this cohort of university students to practice evaluative judgement in their argumentative writing learning, allowing them to understand what constitutes high-quality work and how to create such work (Allen & Katayama, 2016; Gao et al., 2023; Huisman et al., 2018). This study adopted a social constructivist perspective

(Vygotsky, 1978), emphasising the active and constructive nature of knowledge development about the diverse peer feedback roles in the development of evaluative judgement. Firstly, regarding providing feedback, this collaborative approach enhanced feedback providers' understanding of effective writing skills while fostering a sense of ownership and responsibility towards the learning process (Nicol et al., 2014). This aligns with other research findings indicating that providing peer feedback contributes to strengthening and constructing students' own knowledge and promoting deeper understanding (Chang, 2015; Min, 2005; Nicol, 2014; Tsui & Ng, 2000). It also should be noted that, since most students lacked experience in providing feedback, they found it challenging to read others' writings effectively and offer constructive opinions. It has been demonstrated that peer feedback training enhances feedback providers' ability to provide more precise and accurate feedback (Chang, 2015; Min, 2005). Secondly, regarding receiving feedback, students began to appreciate different perspectives, gaining insights into their writing abilities and areas for improvement (Hattie & Timperley, 2007; Nicol & Macfarlan-Dick, 2006). By accepting constructive criticism and incorporating valuable insights into revisions, participants cultivated a culture of continuous growth and development as writers. This process encouraged a growth mindset, fostering a willingness to adapt and improve based on constructive input. Thirdly, regarding the absence of peer feedback, the feedback outsiders' excerpts highlighted the importance of integrating peer feedback into the learning process, underscoring the necessity of creating opportunities for students to actively participate in providing and receiving feedback (Hyland & Hyland, 2006). By incorporating peer feedback sessions into writing instruction, educators can cultivate a collaborative and supportive learning environment (Carless & Boud, 2018; Xie et al., 2022). Students can benefit from the diverse perspectives and collective wisdom of their peers, enabling them to continually improve their writing skills and evaluative judgement (Nicol, 2014).

## Conclusion

This study, set in the context of an argumentative writing course at a public university in Malaysia, provides qualitative evidence for the advantages and limitations of using peer feedback for the development of evaluative judgement among undergraduate students. The results indicate that peer feedback, overall, contributes to the development of students' hard, soft, and dynamic evaluative judgement, as they acquire a more sophisticated understanding of evaluative criteria. However, it is essential not to rely solely on peer feedback, as it may not comprehensively clarify and summarise students' domain-specific knowledge, including linguistic, discourse, and sociolinguistic aspects.

Some results of the thematic analysis are limited in generalisability due to the study's small scope and convenience sampling. Additionally, the study used perceptual data, which is subject to bias, novelty, or researcher effects. Future research could contemplate the adoption of quasi-experimental designs to investigate the effectiveness of peer feedback in enhancing students' evaluative judgement. The inclusion of data sources such as verbal protocols, eye-tracking data, or psychometric data could unveil the intricate cognitive processes. These supplementary data streams possess the potential to illuminate the very mechanisms that lead to the statistically significant variations discerned within the study. The integration of such data would not only provide a finer-grained understanding but could also potentially offer explanations for the patterns and outcomes observed. This would complement existing qualitative research, exploring variations

in learners' behaviours and perceptions across different feedback roles, significantly enhancing the robustness of the study's findings.

### **Funding**

This research is not funded.

### **Acknowledgments**

Not applicable.

### **Conflicts of Interest**

The authors declare no conflict of interest.

### **Authenticity**

This manuscript is an original work.

### **Artificial Intelligence Statement:**

AI and AI-assisted technologies were not used.

### **About the Authors**

**Xiao Xie** is a PhD candidate majoring in the English Language, at the Faculty of Modern Languages and Communication, UPM, Malaysia. He graduated from Sichuan University and Sun Yat-sen University, China. He takes a lively interest in peer feedback, argumentative writing and language assessment literacy. Email: [gs58879@student.upm.edu.my](mailto:gs58879@student.upm.edu.my)

ORCID: <https://orcid.org/0000-0003-3388-7464>

**Vahid Nimehchisalem** holds a PhD in TESL. He's an Associate Professor at the Faculty of Modern Languages and Communication, UPM. Learning-teaching materials evaluation and language assessment are his main areas of research interest. He's chief editor of the International Journal of Education and Literacy Studies and managing editor of the Journal of Language and Communication (JLC). ORCID: <https://orcid.org/0000-0002-5454-1895>

**Mei Fung Yong** holds a PhD in Linguistics and Second Language Teaching from Massey University, New Zealand. She served as an Associate Professor at the Faculty of Modern Languages and Communication, UPM, before her recent retirement. Her areas of research interests include collaborative project-based writing, game-based learning, self-assessment, and second language learning. E-mail: [yong@upm.edu.my](mailto:yong@upm.edu.my) ORCID: <https://orcid.org/0000-0002-5363-1864>

**Yap Ngee Thai** has a PhD in Linguistics. She is an Associate Professor at the Faculty of Modern Languages and Communication, UPM. Her work primarily involves second language acquisition and learning, particularly on speech perception and speech production. She has also worked on language development and processing among bilingual and plurilingual children and adults, and the neuroscience of language. E-mail: [ntyap@upm.edu.my](mailto:ntyap@upm.edu.my) ORCID: <https://orcid.org/0000-0001-8672-7128>

### **References**

- Ajjawi, R., & Bearman, M. (2018). Problematising standards: Representation or performance? In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: Assessment for knowing and producing quality work*, (pp. 41–50). London: Routledge.



- Allen, D., & Katayama, A. (2016). Relative second language proficiency and the giving and receiving of written peer feedback. *System*, 56, 96-106. <https://doi.org/10.1016/j.system.2015.12.002>.
- Aryadoust, V. (2012). Evaluating the psychometric quality of an ESL placement test of writing: A many-facets Rasch study. *Linguistics Journal*, 6(1), 8-33.
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>.
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698-712. <https://doi.org/10.1080/02602938.2012.691462>.
- Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400-413. <https://doi.org/10.1080/02602938.2015.1018133>.
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>.
- Chang, C. Y. H. (2015). Teacher modeling on EFL reviewers' audience-aware feedback and affectivity in L2 peer review. *Assessing Writing*, 25, 2-21. <https://doi.org/10.1016/j.asw.2015.04.001>.
- Chang, C. Y. H. (2016). Two decades of research in L2 peer review. *Journal of Writing Research*, 8(1), 81-117. <https://doi.org/10.17239/jowr-2016.08.01.03>.
- Cheng, X., & Zhang, L. J. (2024). Engaging secondary school students with peer feedback in L2 writing classrooms: A mixed-methods study. *Studies in Educational Evaluation*, 81, 101337. <https://doi.org/10.1016/j.stueduc.2024.101337>.
- Chong, S. W. (2021). University students' perceptions towards using exemplars dialogically to develop evaluative judgement: The case of a high-stakes language test. *Asian-Pacific Journal of Second and Foreign Language Education*, 6(1), 1-27. <https://doi.org/10.1186/s40862-021-00115-4>.
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5(1), 80-92. <https://doi.org/10.1177/160940690600500107>.
- Gan, Z., He, J., Zhang, L. J., & Schumacker, R. (2023). *Examining the relationships between feedback practices and learning motivation. Measurement: Interdisciplinary Research and Perspectives*, 21(1), 38-50. <https://doi.org/10.1080/15366367.2022.2061236>.
- Gao, Y., An, Q., & Schunn, C. D. (2023). The bilateral benefits of providing and receiving peer feedback in academic writing across varying L2 proficiency. *Studies in Educational Evaluation*, 77, 101252. <https://doi.org/10.1016/j.stueduc.2023.101252>.
- Gao, Y., Schunn, C. D. D., & Yu, Q. (2019). The alignment of written peer feedback with draft problems and its impact on revision in peer assessment. *Assessment & Evaluation in Higher Education*, 44(2), 294-308. <https://doi.org/10.1080/02602938.2018.1499075>.
- Goodyear, P., & Markauskaite, L. (2018). Epistemic resourcefulness and the development of evaluative judgement. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: Assessment for knowing and producing quality work*, (pp. 28-38). London: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>.

- Huang, Y., & Zhang, L. J. (2022). Facilitating L2 writers' metacognitive strategy use in argumentative writing using a process-genre approach. *Frontiers in psychology*, 13, 1036831. <https://doi.org/10.3389/fpsyg.2022.1036831>.
- Huisman, B., Saab, N., Van Driel, J., & Van Den Broek, P. (2018). Peer feedback on academic writing: undergraduate students' peer feedback role, peer feedback perceptions and essay performance. *Assessment & Evaluation in Higher Education*, 43(6), 955-968. <https://doi.org/10.1080/02602938.2018.1424318>.
- Hyland, K., & Hyland, F. (2006). Contexts and issues in feedback on L2 writing. *Feedback in Second Language Writing: Contexts and Issues*, (pp. 1-19). Cambridge: Cambridge University Press.
- Joughin, G. (2018). Limits to evaluative judgement. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: Assessment for knowing and producing quality work*, (pp. 60-69). London: Routledge.
- Kim, M. (2009). The impact of an elaborated assessee's role in peer assessment. *Assessment & Evaluation in Higher Education*, 34(1), 105-114. <https://doi.org/10.1080/02602930801955960>.
- Luo, J., & Chan, C. K. Y. (2023). Conceptualising evaluative judgement in the context of holistic competency development: Results of a Delphi study. *Assessment & Evaluation in Higher Education*, 48(4), 513-528. <https://doi.org/10.1080/02602938.2022.2088690>.
- Min, H. T. (2005). Training students to become successful peer reviewers. *System*, 33(2), 293-308. <https://doi.org/10.1016/j.system.2004.11.003>.
- Nelson, R. (2018). Barriers to the cultivation of evaluative judgement: A critical and historical perspective. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: Assessment for knowing and producing quality work*, (pp. 51-59). London: Routledge.
- Nicol, D. J. (2014). Guiding principles for peer review: unlocking learners' evaluative skills. In C. Kreber, C. Anderson, N. Entwistle, & J. McArthur (Eds.), *Advances and innovations in university assessment and feedback* (pp. 197-224). Edinburgh: Edinburgh University Press.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. <https://doi.org/10.1080/03075070600572090>.
- Nicol, D. J., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102-122. <https://doi.org/10.1080/02602938.2013.795518>.
- Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback: How students respond to peers' texts of varying quality. *Instructional Science*, 43, 591-614. <https://doi.org/10.1007/s11251-015-9353-x>.
- Rahimi, M., & Zhang, L. J. (2022). Effects of an engaging process-genre approach on student engagement and writing achievements. *Reading & Writing Quarterly*, 38(5), 487-503. <https://doi.org/10.1080/10573569.2021.1982431>.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144. <https://doi.org/10.1007/BF00117714>.
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher*

- Education*, 76, 467-481. <https://doi.org/10.1007/s10734-017-0220-3>.
- Tai, J., Canny, B. J., Haines, T. P., & Molloy, E. K. (2016). The role of peer-assisted learning in building evaluative judgement: Opportunities in clinical medical education. *Advances in Health Sciences Education*, 21, 659-676. <https://doi.org/10.1007/s10459-015-9659-0>.
- Tsui, A. B., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147-170. [https://doi.org/10.1016/S1060-3743\(00\)00022-9](https://doi.org/10.1016/S1060-3743(00)00022-9).
- Vasu, K., Ling, C. H., & Nimehchisalem, V. (2016). Malaysian tertiary level ESL students' perceptions toward teacher feedback, peer feedback and self-assessment in their writing. *International Journal of Applied Linguistics and English Literature*, 5(5), 158-170. <http://dx.doi.org/10.7575/aiac.ijalel.v.5n.5p.158>.
- Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes*. Cambridge: Harvard University Press.
- Weir, C. (1990). *Communicative language testing*. London: Prentice-Hall.
- Xie, X., Nimehchisalem, V., & Rafik-Galea, S. (2022). Paradigm shifts in peer feedback within learning-oriented language assessment. *ASEAN Journal of Applied Languages*, 1(1), 20-34.
- Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching*, 49(4), 461-493. <https://doi.org/10.1017/S0261444816000161>.