Borneo Engineering & Advanced Multidisciplinary International Journal (BEAM)
Volume 5, Issue 1, May 2026, Pages 23-34

Empowering Voices: A Pilot Evaluation of An Opinion-Expression Activity for Academic English Development Among Polytechnic Students

Charles Cheng Wuen King[1]*, Shobhanambigha Sivaguru[1], Ting Wei Shiow[2]

[1]Politeknik Balik Pulau, Jalan Pinang Nirai, 11000 Balik Pulau, Pulau Pinang, Malaysia
[2]Kolej Komuniti Tawau, Perdana Square Commercial Centre, Batu 4, Jalan Apas, 91000, Tawau, Sabah, Malaysia

*Corresponding author: charles@pbu.edu.my
Please provide an **official organisation email** of the corresponding author

Full Paper

Article history
Received
31 January 2026
Received in revised form
24 March 2026
Accepted
13 April 2026
Published online
1 May 2026

Abstract

Academic English skills are essential for polytechnic IT diploma graduates transitioning to undergraduate study, yet current vocational English curricula often lack adequate support for this transition. This one-group, pretest–posttest pilot study evaluated the feasibility and preliminary efficacy of a 13-week Task-Based Language Teaching (TBLT) opinion-expression intervention. Thirty first-semester IT diploma students at a Malaysian polytechnic completed pre- and post-intervention academic speaking and writing assessments. Performances were evaluated blindly by a trained rater using CEFR-aligned B2 First rubrics. The intervention required students to post weekly asynchronous written responses on Padlet and deliver one in-class oral presentation. Academic speaking scores (max = 25 points) increased significantly, from $M = 9.95$ to 12.92, $t(29) = 5.50$, $p < .001$, $d_z = 1.00$. Academic writing scores (max = 20 points) rose from $M = 5.37$ to 6.37, but this overall gain was not statistically significant, $t(29) = 1.64$, $p = .111$, $d_z = 0.30$. Intra-rater reliability was good for speaking (ICC = 0.84) and moderate-to-good for writing (ICC = 0.72). Baseline-adjusted analyses showed that writing gains varied significantly by prior SPM English grades, $F(2, 26) = 5.04$, $p = .014$; higher pretest scores also predicted smaller gains, $p = .023$. Furthermore, the quality of Padlet responses positively predicted writing improvement, $b = 1.11$, $p = .039$, whereas no significant predictors of speaking gain were identified. These exploratory findings reveal a "modality split," suggesting the intervention broadly improved speaking proficiency but yielded constrained, proficiency-dependent gains in writing. Despite limitations such as the lack of a control group and reliance on single-rater scoring, the results establish feasibility for a full-scale randomized controlled trial and highlight the need for modality-specific scaffolding in TVET academic English curricula.

Keywords: - Academic English, Task-Based Language Teaching (TBLT), opinion expression, polytechnic students, curriculum development

1. Introduction

As the lingua franca for international scholarship, scientific research, and professional enterprise (Muñoz et al., 2025), the mastery of academic literacy in English is critical. Furthermore, it is critical for any aspiring undergraduates to acquire the ability to comprehend, articulate, and negotiate sophisticated academic concepts in English as more than 80% of the world's indexed scientific journals and research outputs are published exclusively in English (Thompson, 2024). Consequently, this non-negotiable demand (Razkane et al., 2025) requires that these aspirants undergo significant shift from fluency in Basic Interpersonal Communication Skills (BICS), which is characteristic of English curricula found in foundational or vocational study (Teemant & Pinnegar, 2019), to the more rigorous and more formalized Cognitive

Academic Language Proficiency (CALP), which typifies formal undergraduate study.

In spite of this, there is currently a systemic and pedagogical mismatch between the need for mastering academic English skills and the English curriculum that forms the program structure of the Diploma of Information Technology (DIT) in Malaysian polytechnics. Specifically, the sole English language course that DIT students are required to take and pass, namely English for Digital Technology (EDT), mandates only the mastery of intensely operational, task-oriented, and narrowly-vocational outcomes, namely describing products and services related to digital technology, outlining basic operational processes and procedures, and handling workplace enquiries and customer's complaints (Saad, 2024). While these routine-based communicative skills are useful for immediate entry into technical employment within the IT sector, these skills lack the necessary exposure and pedagogical scaffolding for mastering the context-reduced, cognitively demanding, and rhetorically complex language skills that are critical for succeeding in university study. For these reasons, these DIT students are ill-equipped to prepare for the mandatory Malaysian University English Test (MUET), and if they do manage to get into university, they might struggle with the dense academic discourses, independent research work, and critical synthesis that are expected in university study (Tan & Raman, 2007).

To address this issue, we formulated this study to investigate the feasibility of introducing and integrating an innovative pedagogical intervention to help these DIT students to develop and expand their academic literacy skills in English. The core of this intervention revolves around productive skills, namely academic writing and speaking, and the reason for this is grounded by the fact that productive language skills are more challenging to master psycholinguistically and typically require active mobilization as opposed to passive reception. Operationally, this intervention utilized a task-based language teaching (TBLT) based opinion-expression activity to shift the locus of learning from interpersonal communication of daily and familiar issues to active and intellectual engagement that is contextualized within academic discourses that are typical of undergraduate-level communication. Through this intervention, students are taught and encouraged to utilize the Stand-Justify-Example (SJE) pathway to articulate their own perspectives and reasoning on peer-generated questions via written and spoken formats. Indirectly, students engage in high-level cognitive processes such as formulating distinct opinions, providing logical justifications, and citing specific, real-world examples to substantiate their claims in a second/foreign language. Furthermore, this dual-modality approach deliberately mirrors the nature of communicative aspects of future university study, which typically relies on written compositions and spontaneous spoken academic discourse to exhibit a learner's mastery of subject knowledge, analytical competence, and critical evaluation skills.

Given the novelty and high cognitive load of this intervention, we found that it is scientifically necessary to conduct an initial evaluation of its pedagogical efficacy, psycholinguistic impact, and logistical viability through a systematic pilot study before determining if a larger, randomized controlled trial (RCT) is justifiable soon. Accordingly, this study is guided by the following research objectives:

1. To explore the extent of change in students' academic writing and speaking skills following participation in a 13-week long opinion-expression activity.
2. To examine whether the changes observed in students' academic English performances vary according to their prior English proficiency.
3. To investigate whether the quality of student's participation, as reflected in their written responses, is associated with academic writing and speaking gains.

2. Literature Review

In his work on BICS and CALP, Cummins (2000) noted that L2 learners typically require a minimum of five to seven years of instruction to approach peer-appropriate academic proficiency; consequently, the psycholinguistic burden of acquiring CALP in an L2 is an immense one as learners are required to execute complex, higher-order cognitive operations such as synthesizing disparate information, evaluating complex theories and abstract concepts, and structuring logical and cohesive arguments through an unfamiliar language. Furthermore, the mastery of L2 is inherently difficult as academic communications are often context-reduced, thus forcing learners to rely almost entirely on highly precise linguistic cues, complex syntaxes, and a vast repertoire of low-frequency vocabulary to convey and extract meanings without the aid of situation scaffolding, which is often found in BICS. As such, the consequences of an educational system that fails to bridge the developmental gap between BICS and CALP are severely and systematically damaging; when students possess high level of BICS but are deficient in CALP, this situation may lead to educators and policymakers to mistakenly assume that learners are proficient in an L2 (Cummins, 2008). Such misdiagnosis often leads to premature entry of students into very cognitively demanding academic environments without adequate support, thus resulting in academic and psychological issues.

For Malaysian polytechnic's DIT graduates who only possess sufficient mastery of BICS, beginning university study without the required mastery of academic literacy skills may result in difficulties in processing academic texts or participating in scholarly discourse due to ceiling effect (Rodríguez, 2010). However, an expression-opinion activity designed using the principles of TBLT may offer them the means of building and expanding their existing academic literacy skills to overcome the said ceiling effect, even within a curricular situation that does not explicitly mandate the mastery of these academic skills. Per Moore's definition of language task within the context of second

language acquisition (SLA), an opinion-expression task that is designed according to the four cardinal TBLT principles – (i) namely prioritization of semantic communication to solve problems and communication gaps, (ii) contextualization within authentic and real-world communicative situations, (iii) deliberate engagement of higher-order cognitive functions, and (iv) explicit completion of non-linguistic outcomes – will force students to rely on their entire interlanguage repertoire and notice their own linguistic gaps implicitly. When students notice how their lack of linguistic ability impedes the process of achieving specific communicative outcomes, a targeted "need to know" situation then arises organically. As a result, students shift from being passive recipients of language rules to active linguistic problem-solvers, thus mirroring the framework suggested by Kolb's models of experiential learning.

In addition, the TBLT-grounded opinion-expression activity may provide opportunities for what Long (1996) called negotiation of meaning and implicit feedback. In a spoken-based opinion-expression activity where a less capable communicator is tasked with expressing and defending an opinion, the communicator rarely achieves these communicative outcomes successfully on the first attempt as they may use lexical items incorrectly, misapply complex grammar structures, or struggle with phonological articulation, thus resulting in failure of comprehension on the part of the more capable listeners. This failure, which then triggers negotiation of meaning, forces the interlocutors to deploy interactional adjustments such as comprehension checks (e.g. "*do you understand what I said?*"), confirmation checks (e.g. "*Right?*"), clarification requests (e.g. "*you mean this, not that, right?*"), targeted repetitions et cetera, which according to Long are specific grammatical and conversational modifications that can increase comprehensibility of input, and subsequently acquisition of target language structures. Furthermore, as the less capable communicator receives implicit, real-time feedback that their original outputs are flawed or inadequate, they are prompted to modify these outputs by formulating and testing new linguistic hypotheses until comprehension (between the interlocutors) are partially or successfully achieved. When executed properly, an interactive opinion-expression activity can act as a powerful scaffold and psychological motivator for second language development irrespective of a learner's prior language proficiency.

As noted previously, a well-designed opinion-expression activity provides not only opportunities for negotiation of meaning and implicit, real-time feedback, but may also cause learners to notice their own communicative gaps and attempt to acquire the linguistic structures that they need to bridge the said gap. Schmidt, building on Swain's Output Hypothesis, which posits that language learning can only occur when learners move from passive and semantic processing to active and syntactic processing (Swain, 1985), further suggests that learners need to consciously direct their attention to the language structures that they need to facilitate uptake and acquisition (Schmidt, 1990). A written-based opinion-expression activity is thus the best way to facilitate uptake via noticing; compared to spoken artefacts, written artefacts such as digital boards are visually permanent and thus provide learners with extended processing time to comprehend their peers' written responses, isolate unfamiliar vocabulary and new linguistic structures, and register them as intake in a conscious manner. Later on, Swain expanded her original Output Hypothesis to extend two more functions to noticing; noticing does not only lead learners to recognize the limitation of their own interlanguage, but also encourages them to test new grammatical hypotheses and lexical items to try to convey their intended meaning, as well as serving as a tool for reflection, where learners use language to discuss, analyze, and internalize linguistic rules via collaborative tasks (Swain, 1995).

All things considered, we can hypothesize, with some reasonable confidence, that the complex interplay of mechanisms of negotiation of meaning, implicit feedback, pushed output, and noticing would make our expression-opinion activity a good candidate for improving and expanding DIT students' academic literacy skills in both speaking and writing skills. However, empirical evidence from research on second language (L2) writing suggests that improvement in academic speaking may not necessarily translate into improvement in academic writing. Two hypotheses may account for this phenomenon; firstly, academic writing in L2 often requires massive working memory resources (Al-Fattah, 2018) and is cognitively more challenging than listening, speaking, or reading in L2 (Dixon & Nessel, 1983). L2 learners often rely on first language translation strategy to generate ideas and structure their arguments in their first languages (L1) before attempting the highly mentally-taxing, word-for-word translation process into L2. Such process then becomes a massive source of extraneous cognitive load, which consumes the entire working memory and leaving zero residual capacity for germane load – the actual mental effort that is needed to notice and fix grammatical errors, refine syntaxes, and learn new academic vocabulary (Parveen, 2025). Conversely, sustained opinion-expression activity may condition learners to begin to think in L2, shift away from L1-to-L2 strategy, and use the now freed-up cognitive capacity to focus on linguistic structures and synthesize meanings using target language. A lower-proficiency learner who can achieve this may benefit more from this shift than higher-proficiency learners, whose cognitive load is already relatively optimized.

Secondly, improvement in one modality (e.g. academic speaking) may not seamlessly transfer to another modality (e.g. academic writing) due to modality effect and Transfer-Appropriate Processing (TAP) effect. According to Sweller et al. (1998), human working memory is not a single and unified construct but is further divided into partially independent processing streams such as a visual/pictorial channel and an auditory/verbal. Within the context of learning, this compartmentalization implies that learning is highly modality-specific, and that stimulation

of visual modality (e.g. typing responses) may not optimally support or strengthen the distinct vocal apparatus or the real-time auditory processing required for spoken interaction. Furthermore, the TAP effect suggests that the degree of success in the retrieval of information from memory is directly proportional to the degree of overlap between the cognitive processes engaged during encoding (learning) phase and retrieval (testing) phase (Bramão & Johansson, 2018). This implies that learning is a highly contextual activity: if an instructional activity highly favors cognitive pathways that are specific to real-time speech production, fluency, verbal negotiation, and phonological articulation, then these skills would transfer exceptionally well to an academic speaking test that require similar cognitive pathways. However, these highly specific cognitive pathways do not seamlessly transfer to the divergent demands of academic writing task that often require meticulous planning, recursive editing, and adherence to rigid structural conventions, most of which require a different set of cognitive pathways altogether.

2.1 The Current Study

As discussed in the previous subsection, a well-designed pedagogical intervention that is rigorously grounded in seamless combination of well-established SLA and cognitive theories has the potential to assist students in acquiring academic literacy skills that are critical for successful university study. TBLT and interactionist-based approaches set the stage for meaningful and productive process of expression of opinion via SJE pathway – when asked to express one's opinion on an issue, learners will first express their **S**tand (e.g. Yes/No/Not Sure), followed by their **J**ustification for their stand, and one or two concrete and appropriate **E**xamples to support both their stand and justification. While engaged with this pathway, learners are forced to produce a comprehensible output to close a communicative gap (such as opinion gap, information gap, justification gap). Should they lack the necessary linguistic structures and vocabulary to do so, they may notice their own linguistic deficit, which theoretically should lead them to hypothesize and experiment with various linguistic rules to close the gap.

Equally important to the successful establishment of the effects of the intervention is rigorous and scientific assessment of the said effects via pre and post tests. At this stage, we might be inclined to think that a large-scale RCT is to establish definitive, causational and/or correlational links. However, such attempt often involves novel routines, digital platforms, and unverified assessment rubrics that may be methodologically hazardous and resource inefficient. To address these issues, a one-group pre-test/post-test pilot study may be more useful for establishing feasibility of a full-scale RCT, assessing potential safety issues and recruitment potential (Thabane et al., 2010), as well as reducing the potential for an underpowered RCT due to inaccurate sample size. Furthermore, a pilot study also acts as a theoretical stress test that allows us to observe whether theoretical predictions will hold true in real-world settings, thus permitting further refinement of instructional design before vast resources are expended on a control group. With this consideration in mind, we aim to examine the effects of the intervention by testing the following hypotheses quantitatively:

H1a: *Participants will demonstrate significant improvements in academic speaking performance after intervention.*

H1b: Participants will demonstrate significant improvements in academic writing performance after intervention.

H2a: Academic speaking gains will differ significantly based on students' prior English proficiency (SPM grade), after controlling for baseline scores.

H2b: Academic writing gains will differ significantly based on students' prior English proficiency (SPM grade), after controlling for baseline scores.

H3a: Higher quality Padlet participation is positively associated with greater academic speaking gains, after controlling baseline scores and proficiency.

H3b: Higher quality of Padlet participation is positively associated with greater academic writing gains, after controlling baseline scores and proficiency.

3. Methodology

3.1 Participants

For this pilot study, we recruited participants from a single class (Classroom Y) at a Malaysian polytechnic (Polytechnic X) located in the northern region of Peninsular Malaysia. Classroom Y was selected via convenience sampling because a member of the research team served as the primary course instructor of their EDT course, thus affording the research team with regular access to research site and participants. To minimise potential coercion associated with the instructor-student relationship, participants were informed verbally that their participation was voluntary, while informed consent procedures emphasised that non-participation would not affect course grades.

Classroom Y initially consisted of 34 full-time students who were enrolled in first semester study of Diploma in Information Technology (DIT) program. The initial cohort was split evenly by gender (17 male participants, 17 female participants) and consisted almost entirely of SPM leavers (n = 33). As our study required examining the moderating effect of participants' prior English language proficiency, we opted to operationalize this using their English language grade in SPM irrespective of their enrolment pathway. To ensure manageable factor levels for moderation effect analysis, we collapsed their original letter grades (A to F) into three distinct proficiency bands: A-range (n = 9), B-range (n = 8), and C/D-range (n = 17). By the end of the pilot study, two participants withdrew from the institution during the study period, which resulted in their incomplete pre–post speaking and writing data being excluded from post-intervention analyses. Two

additional participants were absent from the post-speaking assessment, and two were absent from the post-writing assessment. We then conducted paired pre–post analyses on 30 complete cases for speaking and 30 complete cases for writing.

3.2 Instrumentation and Data Collection

Our SJE-based opinion-expression activity is designed around three primary tools, namely participant-generated questions, a Padlet board, and a custom holistic rubric to measure the quality of participant's written responses on the Padlet board. Prior to data collection, we invited participants to submit their opinion questions on familiar, non-sensitive topics to foster genuine interest and to meet Ministry of Education's CEFR B1 baseline proficiency (Ministry of Education Malaysia, 2018, p. 225). At the beginning of Week 2, we chose an opinion question at the beginning of an EDT class and invited participants to use the SJE pathway to formulate and post their written response to the question on Padlet. We then invite a student randomly to come forward to share their response orally. To ensure every participant has opportunity to share their response with the class orally, we repeated the same procedure at the end of the EDT class and during each EDT class. At the end of the intervention period, we evaluated participants' Padlet post using the 8-point custom holistic rubric.

To measure changes in participants' academic writing and speaking performance, we administered tests for each component before and after the 13-week intervention period, respectively. We aligned these assessments with Ministry of Education's CEFR B2 exit target for post-secondary education (Ministry of Education Malaysia, 2018, p. 26) by sourcing for relevant and suitable assessment tasks from Cambridge English's B2 First Examination assessment materials. We selected Cambridge English's assessment framework for two reasons: its B2-level cognitive demands and expected output matched the expected outcomes of our intervention, and one member of the research team was already professionally trained to assess Cambridge English's examinations.

In academic speaking tests, participants engaged in short individual turns, paired interactions, and extended discussions to provide spoken exemplars for rating. Through these tasks, participants were required to state and elucidate their opinions on a wide range of familiar topics. To prevent collusion and test-retest familiarity, we assigned speaking pairs randomly and used 20 distinct speaking prompt sets. To score participants' academic speaking performance, we utilized Cambridge English's official 25-point holistic rubric to rate each of these criteria equally: global achievement, grammar and vocabulary, pronunciation, interactive communication, and discourse management.

In academic writing tests, participants were required to respond to a writing prompt by composing a 180-to-200-word essay to present and justify their opinion. We then rated these written responses using Cambridge English's official 20-point holistic rubric by evaluating each of these criteria equally: content, communicative achievement, organization, and language. It is important to note that the present pilot study does not assign definitive CEFR levels to individual participants as this is not the study's primary objective, and that participants did not complete the full battery of B2 First Examination.

3.3 Data Analysis Procedures

Prior to hypothesis testing, we followed Tabachnick and Fidell's guidelines to ensure that only complete cases – participants with both pre- and post- intervention scores in both academic speaking and writing components – were retained for paired analyses using listwise deletion, thus ensuring that comparisons that are reflected within participants change over time. All statistical analyses were performed using R version 4.5.2 (Wickham et al., 2012) via RStudio version 4.5.2 (Allaire, 2012).

In addition, we also evaluated intra-rater reliability on 50% stratified random subsamples (for both pre- and post-speaking and writing components) using intraclass coefficient (ICC), specifically a two-way mixed-effects model for single measures, ICC(3,1). We then interpreted the ICC scores using Koo and Li's guidelines: <.50 = poor reliability, .50 - .75 = moderate reliability, .75 - .90 = good reliability, and > .90 = excellent reliability (Koo & Li, 2016). We also conducted fundamental parametric assumptions for all models including normality through Shapiro-Wilk test, homoscedasticity using Breusch-Pagan and Levene's tests, and homogeneity of regression slopes using a Type III Analysis of Variance (ANOVA) test. To ensure that our inferential findings remain robust against potential scoring variability, we planned sensitivity checks using composite intra-rater scores.

Moving onto hypothesis testing, we ran three distinct sets of analyses. For H1a and H1b, we used exploratory paired samples t-tests together with the non-parametric Wilcoxon signed-rank tests to evaluate overall pre- to post-intervention gains, and calculating Cohen's d_z for effect sizes. To interpret the effect sizes, we relied on Cohen's benchmark: .2 = small, .5 = medium, .8 = large (Cohen, 1988). Statistical significance was set at $\alpha = .05$. For H2a and H2b, we conducted exploratory Analysis of Covariance (ANCOVA) within a general linear model (GLM) framework to examine if prior English language proficiency moderated any gains while controlling for baseline scores. We established effect sizes using partial eta squared (ηp^2) and conducted post-hoc pairwise comparisons using Tukey's Honestly Significant Difference (HSD) adjustment. Finally, we used multiple linear regression in H3a and H3b to determine if the average quality of participants' Padlet responses predicted gains in their academic speaking and writing performances, while adjusting for baseline performance and prior English language proficiency.

3.4 Statistical Software and Reproducibility

To ensure the transparency and reproducibility of results, several specialized packages were utilized beyond the base stats library in R. Data organization and sampling (e.g., slice sample) were performed using the tidyverse suite, specifically dplyr (Wickham et al., 2023) and tidyr (Wickham et al., 2024). Intra-rater reliability analysis utilized the irr package (Gamer et al., 2019) for estimating coefficients of interrater reliability. For inferential modelling, the car package (Fox & Weisberg, 2019) was utilized for Levene's tests and the computation of Type II Sums of Squares, while emmeans (Lenth & Piaskowski, 2025) was used for calculating marginal means and Tukey-adjusted post hoc comparisons. All standardized effect sizes and their associated 95% confidence intervals were estimated using the effectsize package (Ben-Shachar et al., 2020). Finally, all graphical outputs were generated using ggplot2 (Wickham, 2016), and the final report was rendered via knitr (Xie, 2015).

4. Result and Discussion

4.1 Descriptive Statistics of Pre- and Post-Intervention Academic Speaking and Writing Test Performance

Table 1: Descriptive statistics of academic speaking and writing performances by prior proficiency band

Modality & Proficiency Band	*n*	Pre-Intervention Mean	Post-Intervention Mean	Gain Mean	Gain SD
Academic speaking (overall)	30	9.95	12.92	2.97	2.95
A-range	8	13.94	18.00	3.69	1.62
B-range	8	10.06	12.88	2.81	3.15
C/D-range	14	7.03	10.04	2.64	3.50
Academic writing (overall)	30	5.37	6.37	1.00	3.33
A-range	8	9.33	12.00	2.75	3.99
B-range	8	5.38	5.00	-0.38	4.14
C/D-range	14	3.06	3.93	0.79	1.97

4.2 Reliability and Assumption Checks

Next, we evaluated the intra-rater reliability of our stratified subsamples. The intra-rater reliability for academic speaking tests was good across all 30 recordings (ICC = .84, 95% CI [.65, .92], $F(29, 18.8) = 13.2$, $p < .001$). On the other hand, the intra-rater reliability for academic writing tests was moderate overall (ICC = .72, 95% CI [.42, .87], $F(29, 14) = 7.83$, $p < .001$), though baseline pre-intervention scores exhibited poorer internal reliability than pre-intervention scores.

We also performed statistical evaluation on specific parametric assumptions for each hypothesis. For paired-samples t-tests (H1a and H1b) and multiple regressions (H3a and H3b), results from Shapiro-Wilk and Breusch-Pagan tests indicated that the data met normality and homoscedasticity requirements ($p > .05$), respectively. For ANCOVA models in H2a and H2b, results from Levene's tests confirmed homogeneity of variances across all distinct proficiency bands ($p > .05$). However, a Type III ANOVA revealed a significant interaction between pre-intervention academic writing scores and SPM proficiency bands, thus violating the homogeneity of regression slopes for H2b ($F(2, 24) = 3.45$, $p = .048$). For this reason, we were required to interpret the subsequent ANCOVA of academic writing scores with caution.

4.3 Effects of Intervention on Academic Speaking and Writing Performances (H1a & H1b)

Results from H1a indicated that the 13-week SJE-based opinion-expression activity has resulted in a massive, positive shift in academic speaking performance. Participants significantly improved their academic speaking scores from a pre-intervention mean of 9.95 points to a post-intervention mean of 12.92 points. This 2.97-point mean gain was highly significant with large effect size, $t(29) = 5.50$, $p < .001$, $d_z = 1.00$); we can thus conclude that H1a is fully supported.

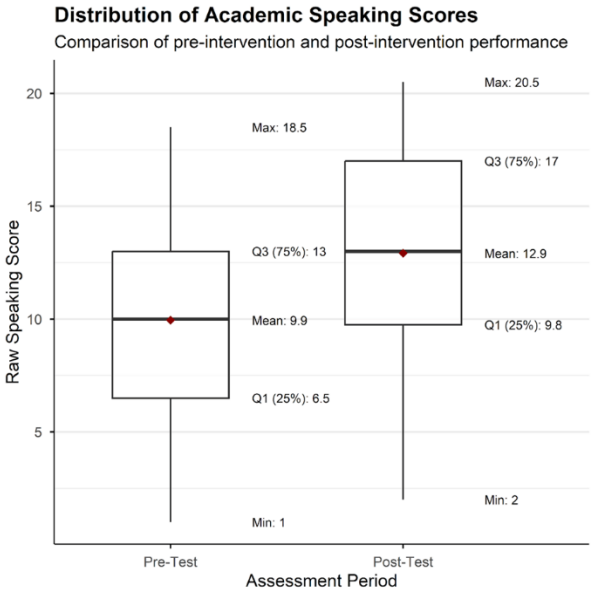

Fig. 1: Comparison of pre-intervention and post-intervention academic speaking test scores

In contrast, the pre- and post-intervention writing scores painted a completely different picture. Mean academic writing performance shifted slightly upward from pre-intervention score of 5.37 points to post-intervention score of 6.37; however, this 1.00-point gain was not statistically significant, $t(29) = 1.64$, $p = .111$, $d_z = 0.30$. We thus conclude that H1b was not supported; the SJE-based opinion-expression activity does produce a robust effect on academic speaking performance, but its short-term association with academic writing was minimal for the overall sample.

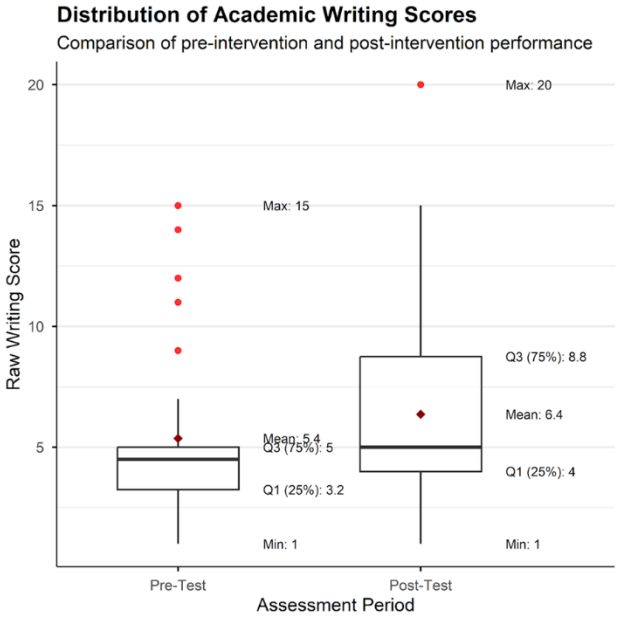

Fig. 2: Comparison of pre-intervention and post-intervention academic writing test scores

4.4 Moderating Effect of Prior English Language Proficiency (H2a & H2b)

For H2a and H2b, we ran exploratory ANCOVAs to assess if prior English language grade in SPM moderated outcomes in H1a and H1b. For academic speaking performance, prior English language proficiency had no significant main effect, $F(2, 26) = 1.36$, $p = .274$ ($\eta p^2 = .09$). We can thus conclude that H2a was not supported, and that academic speaking development trajectories were uniformly positive irrespective of participants' baseline proficiency.

On the other hand, we found that academic writing gain varied significantly by prior English language proficiency, $F(2, 26) = 5.04$, $p = .014$ ($\eta p^2 = .28$). Post-hoc pairwise comparisons with Tukey's HSD adjustment revealed that participants in the highest proficiency band (A-range) exhibited significantly larger adjusted mean gains than the B-range (mean difference = 5.28, $p = .014$) and C-D range (mean difference = 5.36, $p = .026$) groups. For these reasons, we conclude that H2b was supported, thus highlighting a distinct advantage for A-range group participants in development of academic writing skills.

4.5 Robustness Checks Against Scoring Variability

As a safeguard against measurement error due to intra-rater variability, we conducted subsequent sensitivity analysis checks on H1a, H1b, H2a, and H2b. First, we generated a hybrid dataset by replacing the original scores of 15 re-rated participants (selected randomly while adhering to group stratification) with their composite mean scores across the two rating periods (pre- and post-intervention). We then re-estimated all primary models for the four hypotheses.

The sensitivity checks confirmed the statistical conclusion of our primary analysis. The massive overall gain for academic speaking (H1a) remained highly significant ($t(29) = 5.20$, $p < .001$), while the overall gain for academic writing remained non-significant ($t(29) = 1.42$, $p = .168$). Similarly, the ANCOVA main effects regarding prior English language proficiency remained unchanged: non-significant for academic speaking performance ($p = .131$), and significant for academic writing performance ($p = .026$). Since there were no significant shifts in thresholds, we conclude that our primary findings for H1a, H1b, H2a, and H2b remained robust against potential rater drift.

4.6 Quality of written Padlet responses as predictor of gains (H3a & H3b)

In our final analysis, we examine whether the quality of asynchronous Padlet responses predicted improvements in academic speaking and writing. We aggregated five sampled sessions to compute a mean Padlet quality score for each participant ($M = 4.32$, $SD = 1.13$).

For H3a, we found that the quality of written Padlet responses did not significantly predict gains in academic speaking scores, $b = -0.24$, $SE = 0.56$, $t(25) = -0.43$, $p = .673$, 95% CI [-1.39, 0.91]. The overall model ($F(4,25) = 0.76$, $p = .560$) also did not significantly predict speaking gains, explaining approximately 11% of the variance ($R^2 = .11$). We thus conclude that H3a was not supported.

On the contrary, the regression model for academic writing performance was statistically significant ($F(4,25) = 4.16$, $p = .010$), explaining roughly 40% of the variance ($R^2 = .40$) in academic writing gains. After adjusting for baseline performance and prior English language proficiency, the quality of written Padlet responses emerged as a significant positive predictor of academic writing improvement, $b = 1.11$, $SE = 0.51$, $t(25) = 2.18$, $p = .039$, 95% CI [0.06, 2.16]. This means that for every 1-point increase in written Padlet response, gains in academic writing increased by about 1.11 points. We thus conclude that H3b was supported.

4.7 Summary of hypothesis testing

Table 2 provides a comprehensive overview of the statistical decisions for all primary and exploratory hypotheses.

This exploratory pilot study aimed to investigate how well a new, task-based language teaching (TBLT) activity, focused on expressing opinions, worked in developing academic speaking and writing skills in English for Malaysian polytechnic students. The study specifically looked at a 13-week opinion-expression activity that used the Stand-Justify-Example (SJE) method, along with an online platform (Padlet). The goal was to help students move from Basic Interpersonal Communicative Skills (BICS) to Cognitive Academic Language Proficiency (CALP).

Table 2: Summary of hypothesis testing

Hypothesis	Relationship	Statistical test	Statistical result	*p*-value	Decision
H1a	Pre- to post-intervention change (academic speaking)	Paired t-test	$t = 5.50$	< .001	**Supported**
H1b	Pre- to post-intervention change (academic writing)	Paired t-test	$t = 1.64$	.111	Not supported
H2a	Effect of prior proficiency on academic speaking gains	ANCOVA	$F = 1.36$	.274	Not supported
H2b	Effect of prior proficiency on academic writing gains	ANCOVA	$F = 5.04$	.014	**Supported**
H3a	Padlet quality as a predictor of academic speaking gains	Multiple regression	$b = -.24$	.673	Not supported
H3b	Padlet quality as a predictor of academic writing gains	Multiple regression	$b = 1.11$	.039	**Supported**

Note: All statistical decisions remained robust across sensitivity analyses using hybrid intra-rater composite scores.

Empirical findings from the paired pre- and post-intervention assessments present a highly complex, nuanced, and theoretically rich dataset; statistical results reveal a pronounced "modality split," wherein the intervention catalysed massive, universal, and statistically significant gains in academic speaking, yet resulted in constrained, proficiency-dependent, and modality-specific gains in academic writing. As this divergence in language acquisition trajectories demands rigorous theoretical contextualization, the following discussion will attempt to interpret these findings through the interconnected lenses of Output Hypothesis, Interaction Hypothesis, Cognitive Load Theory (CLT), Expertise Reversal Effect, and Transfer-Appropriate Processing (TAP).

The assessment of H1a demonstrated a statistically significant and practically substantial enhancement in academic speaking performance after the 13-week intervention. Moreover, the extent of this improvement was determined by a large effect size. Consequently, the structured, cyclical expectation of publicly expressing and justifying academic viewpoints through the SJE pathway may have instigated a reorganization of learners' oral proficiency. Crucially, the subsequent ANCOVA for H2a demonstrated that these speaking gains did not differ significantly based on students' prior English proficiency, as measured by their Sijil Pelajaran Malaysia (SPM) grade. From an A-range student to a C/D-range student, the developmental trajectory in the oral modality was remarkably uniform and positive. This universal improvement in speaking may be attributed to the synergy of psycholinguistic mechanisms of Swain's Output Hypothesis and Long's Interaction Hypothesis. Within traditional diploma curriculum, students were accustomed to highly contextualized BICS environments where top-down communicative shortcuts sufficed. Conversely, when students were required to employ the SJE pathway in an oral presentation, they were unable to depend on these simplified strategies. The intervention triggered the output's "noticing function"; as learners attempted to express complex justifications and practical examples, they encountered immediate linguistic limitations – conscious realizations of the differences between their intended communications and their actual syntactic abilities. Moreover, the consistency of these oral improvements across all proficiency tiers (H2a) highlights the significant equalizing effect of interactional scaffolding. During real-time SJE presentations, when a lower-proficiency learner encountered difficulty in expressing justification, the immediate feedback from the instructor or peers (through clarification requests, comprehension checks, or recasts) potentially facilitated localized, dynamic scaffolding. Such mechanism has been observed and reported by Aqila (2025) who noted that task-based oral frameworks provide learners with opportunity for creativity and engagement, thus resulting in significant enhancements in communication skills regardless of baseline starting points. In polytechnic, this interactive environment effectively served as a great equalizer; C/D-range learners may have benefited from the communicative pressure and real-time peer scaffolding just as much as their A-range counterparts. The oral modality, therefore, functioned as a highly fluid, highly negotiated space where psycholinguistic barriers were collectively dismantled, leading to the massive effect size observed in the cohort.

On the other hand, the analysis for H1b revealed that the intervention did not produce a statistically significant overall improvement in academic writing performance for the aggregate sample. While the median score shifted slightly in a positive direction, the mean gain was accompanied by a rather small effect. The group's developmental trajectory in writing seemed to be resistant to short-term, generalized improvement; such divergence between oral and written outcomes forms the crux of the observed "modality split." This phenomenon cannot be explained by interactionist theories alone; it necessitates a rigorous examination of human memory architecture through Cognitive Load Theory (CLT). Academic writing in a second language (L2) is inherently characterized by exceptionally high element interactivity; as Sweller (2010) notes, as "*the act of coordinating these multiple, interacting elements is unrelated to learning the information being presented, imposing a heavy working memory load*". As observed through this study, when polytechnic students were required to compose a coherent academic argument during the post-intervention writing test, they had to simultaneously manage morphosyntactic accuracy, retrieve low-frequency academic vocabulary, structure logical discourse adhering to the SJE pathway, and maintain rhetorical conventions – all without the aid of paralinguistic cues or immediate peer negotiation. For the average polytechnic student in this cohort, the intrinsic cognitive load of the academic writing test was massive. Furthermore, novice writers often generate ideas and structure arguments entirely in their L1 before attempting a mentally taxing, word-for-word translation process into

English (Al-Fattah, 2018). This strategy thus acts as a massive source of extraneous cognitive load, entirely saturating the limited capacity of working memory. When the demands of translation and the intricate interactivity of the writing process overwhelm working memory, no capacity remains for germane load, which encompasses the cognitive effort needed to identify grammatical errors, improve complex sentence structures, or transfer novel academic vocabulary from short-term to long-term memory. Yao and Fan's (2025) recent research on cognitive load in L2 writing corroborated these findings, demonstrating that the central executive function of working memory is essential for allocating attentional resources to simultaneously manage content, structure, and audience considerations, thereby rendering writing a particularly challenging cognitive task for L2 learners. As a result, the 13-week intervention proved inadequate for most of the overall group in mitigating the significant cognitive constraints inherent in L2 academic writing.

Although the overall sample did not exhibit notable improvements in writing, the exploratory ANCOVA employed to assess H2b revealed a significant subgroup effect. The analysis showed a statistically significant main effect of pre-existing English language proficiency on academic writing gains. Subsequent pairwise comparisons indicated that participants within the A-range (high proficiency) attained significantly greater adjusted mean gains relative to both the B-range and C/D-range groups. This observation implies that prior linguistic proficiency could serve as a potent moderator in the development of L2 academic writing skills. This "rich-get-richer" dynamic in written L2 acquisition can be theoretically explained by the "Expertise Reversal Effect", an advanced corollary of Cognitive Load Theory advanced by Kalyuga et al. (2003) who suggested that "[i]nstructional techniques that are highly effective with inexperienced learners can lose their effectiveness and even have negative consequences when used with more experienced learners" (p. 23). In the context of this study, the high element interactivity of the independent academic writing task may have overwhelmed the working memory of the B-range and C/D-range learners, which potentially resulted in cognitive paralysis and static pre- and post-intervention scores. However, the A-range learners may have already automatized many lower-level linguistic operations such as basic grammatical encoding, spelling, and high-frequency vocabulary retrieval through their prior years of successful English study. Because these foundational linguistic schemas were already consolidated and stored in their long-term memory, the A-range students did not experience catastrophic working memory overload when faced with the writing assessment. Instead, their cognitive capacity was liberated; they could dedicate their available working memory entirely to germane cognitive load by processing macro-level discourse structures of the SJE pathway, refining their academic tone, and experimenting with the complex syntaxes required for CALP. The intervention could have successfully pushed the A-range students into higher academic registers because they possessed the baseline cognitive architecture required to survive the transition. On the contrary, C/D-range students were cognitively stranded – the lack of explicit guidance and low-element-interactivity scaffolding during writing assessments may have prevented them from translating their oral SJE competence into written modality.

In the final part of this study, we investigated how Padlet, an asynchronous digital platform, worked in the opinion-expression activity. The results for H3a showed that the quality of written Padlet response did not significantly predict their improvements in academic speaking skills. Conversely, the outcomes derived from the evaluation of H3b demonstrated that the quality of written Padlet responses was a statistically significant, positive predictor of academic writing improvements. This highly specific, directional transfer of skills - where text-based practice predicts text-based testing, but not oral testing – may be attributed to the manifestation of Transfer-Appropriate Processing (TAP). When students formulated their SJE responses asynchronously on Padlet, they were practicing language in a visual, text-based modality – typing on Padlet specifically stimulates the visual working memory processor and orthographic retrieval networks. Because the academic writing test demanded these exact same visual and orthographic cognitive pathways, the high-quality practice on Padlet may have transferred seamlessly to post-test writing test. Similar effect was also evident in the study conducted by Sugiarni et al. (2025) and their study showed that using Padlet in a process writing approach significantly improved students' writing skills. The reason for this was that Padlet provided structured, collaborative spaces, which increased visual engagement during the drafting phase. In contrast, using a digital board does not require the vocal apparatus, nor does it develop the real-time auditory processing and phonological loops needed for spontaneous speech. For these reasons, it is only logical to surmise that text-based asynchronous practice does not predict variance in the oral speaking test due to the constraints of Transfer-Appropriate Processing and the Modality Effect. The success of the Padlet intervention in predicting writing gains can thus be intricately attributed to Schmidt's Noticing Hypothesis: spoken words are inherently transient as they vanish the exact millisecond they are uttered, thus leaving lower-proficiency learners with virtually no time to consciously register novel lexical items or complex syntaxes. On the other hand, text artefacts generated and archived on a Padlet board possesses visual permanence – this permanence affords learners the extended processing time necessary to direct their focal attention toward linguistic forms in a conscious manner. High-quality participation on Padlet indicates that learners may have utilized this extended time actively to reflect meta-linguistically, thus resulting in successful conversion of environmental "input" into acquired "intake." This conscious, text-based metalinguistic reflection may have then directly scaffolded their subsequent performance in formal academic writing assessments.

5. Implications and Future Recommendations

As evidenced by findings from the hypothesis testing, the SJE-based opinion-expression activity has the potential to assist not only DIT but also other ESL/EFL learners to develop their academic literacy skills. The ability of the intervention to induct a massive effect size in the development of academic speaking suggests that structured TBLT routines may disrupt BICS-level stagnation. Furthermore, TBLT-based strategies may also be augmented technologically to assist SLA; Veas Aguirre and Pacheco Vasquez (2025) noted that technologically-mediated TBLT can potentially enhance learners' speaking fluency and interactional skills via authentic, meaningful tasks.

With these in mind, we offer several practical recommendations for English language educators working within TVET sector who aspire to bridge the gap between basic workplace English and academic English discourse. Firstly, we recommend that educators should aim to implement asynchronous scaffolding for lower-proficiency students so that they can reduce cognitive overload during real-time oral production. By providing these learners with digital boards (such as Padlet) or even notebooks/notepads, learners are allowed sufficient time to draft and revise their responses, use dictionaries and other assistive tools, and negotiate meaning without suffering the debilitating psychological pressure of a real-time audience. Shared responses on digital boards may also serve as persistent repository of exemplars of good language use, helping students to consciously notice and analyse high-quality vocabulary and linguistic structures before they are expected to model these exemplars in their own responses.

Secondly, given the strong evidence of modality effect and lack of cross-skill transfer, we believe that teachers should ensure that practice environments are aligned with appropriate modalities. Teachers should take care not to assume that proficiency in oral monologues or interactions would translate automatically into written competence. Instead, teachers should benefit from employing a "guidance fading" strategy – classroom instruction should first begin with SJE-based oral tasks to build semantic fluency, confidence, and basic conceptual schemas as L2 writing has been shown to be cognitively more challenging than L2 speaking. Only when these semantic concepts are firmly established in learners' memory should the instruction transition into written modality; at this stage, learners would benefit greatly from heavy, structured guidance (e.g. worked examples of academic paragraphs, explicit sentence starters, structured cloze exercises et cetera) to lower the extraneous load of L2 writing.

Thirdly, as evidenced by the data, a "one-size-fits-all" approach to designing and delivering academic writing instruction is fundamentally problematic and should be replaced with differentiated instructions. A-range students thrive under open-ended cognitive demands of generating independent, complex written responses as they have already automatized the skills required for these demands. However, similar conditions often result in cognitive paralysis among C/D-range students, which necessitates strictly differentiated instructions as they often require explicit, step-by-step grammatical instructions, and highly constrained writing tasks to build their foundational schemas and bypass the cognitive drain of L1-to-L2 translation task. Only when these skills are proceduralized properly should these students be pushed towards producing extended, independent academic discourse.

Finally, while this pilot study has provided valuable preliminary data to support the strategies above, it is also imperative that we acknowledge several methodological limitations. First, the 13-week timeframe was relatively brief to observe deep linguistic maturation. Furthermore, the small sample size ($N = 30$) from a single class and a single institution does not permit broad generalizability. Most importantly, relying on a one-group pre-test/post-test design means that we cannot isolate the effect of Padlet use from standard developmental maturation. As such, we would recommend that future research transition into a full RCT study – in which an active control group is established to receive traditional grammar-translation instruction or presentation-practice-production (PPP) style instruction – so that the exact effect from the SJE-based intervention can be isolated properly. In addition, future RCT studies may also benefit from delayed and repeated post-intervention measurements to measure long-term retention.

6. Conclusion

The shift from foundational vocational training to demanding undergraduate programs necessitates a significant linguistic transformation. This involves moving from the context-specific, cognitively accessible BICS to the abstract, formalized, and cognitively challenging CALP. Presently, for students enrolled in the DIT program at Malaysian polytechnics, this transition is hindered by a systemic pedagogical misalignment. The current curriculum emphasizes highly operational, routine-based English, which effectively limits their linguistic development and subsequently disadvantages them when confronted with the intricate discourse required by universities and gateway assessments such as the MUET.

This exploratory pilot study investigated the feasibility of disrupting this stagnation by introducing a TBLT-based intervention utilizing the SJE pathway across dual modalities (oral and written). The empirical findings illuminate a striking and theoretically profound modality-split. Results from data analysis suggest that the intervention may have led to the dismantling of barriers that obstruct successful L2 development. By forcing learners into the "noticing" function of output and subjecting them to the real-time, peer-driven negotiation of meaning, the SJE pathway may have catalyzed massive, equitable improvements in academic speaking across all proficiency levels. Within the oral domain, interaction may have served as a dynamic equalizer, thus suggesting that polytechnic students are highly capable of engaging in sophisticated academic discourse when provided with structured, communicative scaffolding.

On the other hand, academic writing proved to be highly resistant to short-term, generalized improvement due to constraints of high element interactivity and the architectural limits of human working memory. The data starkly illustrates the limitations of human working memory: while high-proficiency learners have leveraged their automatized foundational skills to thrive and expand their written repertoire, lower-proficiency learners succumbed to cognitive overload, trapping them in the extraneous load of L1-to-L2 translation. Furthermore, the significant association between asynchronous digital engagement (Padlet) and writing gains – but not speaking gains – reinforced the possibility of TAP and Modality Effect, proving that specific cognitive pathways must be explicitly and visually targeted to yield specific linguistic transfers.

To conclude, this research has provided for a vital, empirically grounded blueprint for educational reform in second language learning within the TVET sector. It has established the possibility that short-term, cognitively demanding interventions may be useful in triggering acquisition of academic literacy and oral confidence. However, it also serves as a critical warning that the psycholinguistic burden of academic writing requires prolonged, differentiated, and modality-specific support. Moving forward, the parameters and limitations established by this pilot study provide the methodological progression criteria that are necessary for a full-scale, active-control RCT study. By continuing to refine, differentiate, and scale these digital and cognitive interventions, educators and policymakers can ensure that polytechnic graduates possess the advanced linguistic capital required not just for entry-level technical employment, but also for upward academic mobility in the globalized knowledge economy.

Author Contributions: The research study was carried out successfully with contributions from all authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

Al-Fattah, N. (2018). Cognitive load theory in the context of second language academic writing. *Higher Education Pedagogies*, *3*(1), 385–402. https://doi.org/10.1080/23752696.2018.1513812.

Allaire, J. (2012). RStudio: integrated development environment for R. *Boston, MA*, *770*(394), 165-171.

Aqila, A. P. (2025). TBLT Strategies on Students' Speaking Achievement: An Experimental Study. *JET (Journal of English Teaching)*, *11*(2), 185–197. https://doi.org/10.33541/jet.v11i2.6911.

Ben-Shachar, M. S., Lüdeck, D., & Makowski, D. (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. *Journal of Open Source Software*, *5*(56), 2815. https://doi.org/10.21105/joss.02815.

Bramão, I., & Johansson, M. (2018). Neural Pattern Classification Tracks Transfer-Appropriate Processing in Episodic Memory. *Eneuro*, *5*(4), ENEURO.0251-18.2018. https://doi.org/10.1523/ENEURO.0251-18.2018.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.

Cummins, J. (2008). BICS and CALP: Empirical and Theoretical Status of the Distinction. In N. H. Hornberger (Ed.), *Encyclopedia of Language and Education* (pp. 487–499). Springer US. https://doi.org/10.1007/978-0-387-30424-3_36.

Dixon, C. N., & Nessel, D. D. (1983). *Language experience approach to reading (and writing): Language-experience reading for second language learners*. Alemany Press.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third). Sage.

Gamer, M., Lemon, J., & Singh, I. F. P. (2019). *irr: Various Coefficients of Interrater Reliability and Agreement*. https://doi.org/10.32614/CRAN.package.irr.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Lenth, R. V., & Piaskowski, J. (2025). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. https://doi.org/10.32614/CRAN.package.emmeans.

Long, M. H. (1996). The Role of the Linguistic Environment in Second Language Acquisition. In W. C. Ritchie & T. K. Bhatia (Eds), *Handbook of Second Language Acquisition* (pp. 413–468). Academic Press. https://doi.org/10.1016/B978-012589042-7/50015-3.

Ministry of Education Malaysia. (2018). *English Language Education Reform in Malaysia: The Roadmap 2015-2025* (Second Edition). Ministry of Education Malaysia.

Moore, P. J. (2018). Task-Based Language Teaching (TBLT). In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (1st edn, pp. 1–7). Wiley. https://doi.org/10.1002/9781118784235.eelt0175.

Muñoz, D. A. A., Fernanda Tobón, M., & Cely Betancourt, B. L. (2025). The Advantages of using English as the Medium of Instruction for Globally Oriented Professional Training. A Theoretical Review. *Ciencia Latina Revista Científica Multidisciplinar*, *9*(4), 8376–8398. https://doi.org/10.37811/cl_rcm.v9i4.19422.

Parveen, N. (2025). Evidence-Based Practices of Cognitive Load Management to Enhance Learning. *Psychological Studies*, *70*(2), 374–386. https://doi.org/10.1007/s12646-025-00841-6.

Razkane, H., Ben Hammou, S., & Sayeh, A. Y. (2025). A constructivist inquiry into English-Medium Instruction in Moroccan higher education. *Training, Language and Culture*, *9*(3), 71–83. https://doi.org/10.22363/2521-442X-2025-9-3-71-83.

Rodríguez, A. S. (2010). *The Influence of Cross-Linguistic Input and L2 Proficiency on L2 Reading Comprehension Among Spanish-Speaking Adults Reading Comprehension Among Spanish-Speaking Adults Learning English as a Second Language Learning English as a Second Language* [CUNY Academic Works].

Saad, S. N. (2024). *Student's Guide Book : Diploma In Information Technology (DIT). Version 1.0 2024*. Jabatan Teknologi Maklumat & Komunikasi, Politeknik Muadzam Shah, 21.

Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics*, *11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129.

Sugiarni, S., Tahrun, T., & Firdaus, M. (2025). Padlet and process writing: A collaborative way for improving students' writing achievement. *Celtic : A Journal of Culture, English Language Teaching, Literature and Linguistics*, *12*(1), 369–383. https://doi.org/10.22219/celtic.v12i1.40607.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds), *Input in Second Language Acquisition* (pp. 235–253). Newbury House Publishers.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford University Press.

Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, *22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5.

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/A:1022193728205.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed). Pearson Education.

Tan, Y. S., & Raman, S. R. (2007). Problems and Challenges of Learning Through a Second Language: The Case of Teaching of Science and Mathematics in English in the Malaysian Primary Schools. *Kajian Malaysia*, *25*(2), 29–54.

Teemant, A., & Pinnegar, S. A. (2019). Understanding BICS and CALP: Variability Summary A. In B. Allman (Ed.), *Principles of Language Acquisition* (p. 160). EdTech Books.

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., Robson, R., Thabane, M., Giangregorio, L., & Goldsmith, C. H. (2010). A tutorial on pilot studies: The what, why and how. *BMC Medical Research Methodology*, *10*(1), 1. https://doi.org/10.1186/1471-2288-10-1.

Thompson, J. (2024). The Global Significance of English Language Education in the 21st Century. *Journal of Foreign Language Education and Technology*, *9*(3).

Veas Aguirre, S. J., & Pacheco Vasquez, E. A. (2025). Effect of technology-mediated task-based language teaching on EFL learners' linguistic and affective outcomes: A systematic review: Efecto de la enseñanza de idiomas basada en tareas y mediada por tecnología en los resultados lingüísticos y afectivos de los estudiantes de Inglés como Lengua Extranjera: una revisión sistemática. *Boletín Científico Ideas y Voces*, *5*(3). https://doi.org/10.60100/bciv.v5i3.265.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, H. R., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., ... & Revelle, W. (2012). A language and environment for statistical computing. *Vienna, Austria*.

Wickham, H., Francois, R., Henry, L., Muller, K., & Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation*. https://doi.org/10.32614/CRAN.package.dplyr.

Wickham, H., Vaughan, D., & Girlich, M. (2024). *tidyr: Tidy Messy Data*. https://doi.org/10.32614/CRAN.package.tidyr.

Xie, Y. (2015). *Dynamic documents with R and Knitr* (Second edition). CRC Press/Taylor & Francis.

Yao, G., & Fan, L. (2025). Cognitive load scale for AI-assisted L2 writing: Scale development and validation. *Frontiers in Psychology*, *16*, 1666974. https://doi.org/10.3389/fpsyg.2025.1666974.