

Enhancing Learners' Lexical Resources In IELTS Task 2 Writing: The Role Of Noticing Hypothesis, Model Essays, And Productive Vocabulary Knowledge

Nguyen Long Quoc¹, Le Ha Van²

¹FPT University, Ho Chi Minh, Vietnam.

²FPT University, Ho Chi Minh, Vietnam.

ABSTRACT

The purpose of this study was to determine if the combination of the noticing hypothesis and model essays had a substantial impact on learners' lexical resource performance in IELTS Task 2 writing, as well as to investigate the predictive function of productive vocabulary knowledge in this process. Using a quasi-experimental design with the participation of 32 students, it was discovered that, despite not playing a significant role in boosting learners' vocabulary scores in academic essay writing compared to the conventional method, the noticing-model essays helped learners raise their LR score to as high as 0.69 in just two weeks. In addition, it was found that productive vocabulary knowledge did not significantly affect IELTS Task 2 writing LR scores. Therefore, teachers are encouraged to use this innovative technique in their lectures, with or without the combination of teacher and peer feedback. Another educational implication was that teachers should not rush to assess their students' performance but should wait until students have had sufficient exposure to model essays and adequate practice before timed examinations. This study also included several limitations and recommendations for future research.

Keywords: Lexical resources, IELTS writing task 2, noticing hypothesis, model essays, vocabulary knowledge.

I. Introduction

Obtaining a sufficient IELTS band score (International English Language Test System) has become one of the primary goals of many L2 learners, especially those in Asia, who hope to be admitted to schools in English-speaking countries such as the US, the UK, Australia, and New Zealand. Nonetheless, it is a challenging task for such learners to gain the desired score without frequently and properly practicing the language. Of all the skills of English, academic writing has been regarded as the most difficult. As a result, many researchers have studied the best ways to help learners write more effectively. The most prevalent and traditional way is to rely on teachers' feedback (Yang et al., 2006; Séror, 2011;

Zhao, 2010). However, as it is highly time-consuming for teachers to give sufficient feedback on every single student's writing (Lee, 2003), the focus has been placed on peer feedback (e.g., Uymaz, 2019; Susanto & Hidayati, 2020; Ruru & Sulisty, 2020). Although it proved beneficial to students' writing, peer feedback can be subjective and lengthy (Braine, 2003; Rollinson, 2005). Another method is to let the learner review his or her own paragraphs or essays for improvements (Yang et al., 2006; Lin, 2009; Chen, 2010).

One more kind of feedback that has received attention from researchers recently is the use of model essays as an instrument (e.g., Qi & Lapkin, 2001; Hanaoka, 2007; Bagheri & Zare, 2009;

Tieu & Baker, 2022). Learners are given native speakers' essays which they use to compare and contrast their own writing. Then, they talk to their teacher or friends about what they notice or learn from the model essays before being asked to revise their own writing. It was shown that learners did improve their writing score, primarily the lexical aspect (Tieu & Baker, 2022; Qi & Lapkin, 2001); however, the number of studies on this method is still limited. Further, whether learners' productive vocabulary knowledge played any role in writing improvements still remains unclear. These deficiencies highlight the importance of further research to clarify the doubt.

2. Literature review

2.1 IELTS Task 2 Writing, Academic Module

Task 2 Writing in IELTS requires test-takers to write an academic essay of at least 250 words on a given topic in 40 minutes (IELTS Academic 17, 2022). While the topic can be anything, the types of essays include, but not limited to, causes and effects, causes and solutions, advantages and disadvantages, opinions, and discussion of both views. The essays are graded based on four main criteria: Task Response (TR – ideas), Lexical Resources (LR – vocabulary), Grammatical Range and Accuracy (GR – grammar), and Cohesion-Coherence (CC – organization). Further details can be found in the public version of Task 2 Writing band descriptors (n.d.) (See Appendix). The lowest score is 0, and the highest score is 9.0. If the score is a decimal, it can be rounded to 0 or .5 (i.e., 6.10 to 6.0 and 6.3 to 6.5).

2.2 Productive vocabulary knowledge

Productive vocabulary knowledge is the ability to utilize a word appropriately in oral and written contexts (Crow, 1986). Numerous L2 vocabulary researchers have placed emphasis on the possible connection between vocabulary knowledge and the capacity for speaking, reading, writing, and listening in the target language (e.g., Stæhr, 2009;

Oya, Manalo & Greenwood, 2009; Uchihara & Saito, 2019). Read (2020) asserts that several assessments can be used to measure productive vocabulary knowledge, including the Vocabulary Levels Test and the Productive Vocabulary Levels.

2.3 The noticing hypothesis

In 1983, Schmidt laid the groundwork for this hypothesis when positing that learners needed to be drawn attention to the target language, use it, and test it against native speakers' version for knowledge gains. This is because "input does not become intake for language learning unless it is noticed, that is, consciously registered" (Schmidt, 1983, p. 271). Schmidt's hypothesis received strong support from many researchers in second language acquisition (Schmidt & Frota, 1986; Swain & Lapkin, 1985; Leow, 2018). Despite the fact that noticing is crucial to SLA, little research has been done on how it affects writing, particularly in IELTS academic essays (Abe, 2009; Tieu & Barker, 2022). Further research is required in light of this.

2.4 The output hypothesis

Swain first developed and discussed the output hypothesis for second language learning in 1985 and then expanded it in the next two decades (1995, 1998, 2005). This hypothesis consists of three essential components, the first one, noticing, occurs as learners attempt to use the target language from which they can identify the issues that impede them from producing the desired meaning. This discrepancy encourages students to pay special attention to the necessary tools of expression that they need in order to properly and accurately convey the intended message. The second function, hypothesis testing, is for students to check what they have learned previously. At this stage, feedback is crucial, whether from peers or teachers, because students require a benchmark against which they can measure their hypotheses (Swain, 1998). The last component is called "metalinguistic awareness,"

in which students reflect on their newly obtained knowledge by using it in a new context.

2.5 Past research on model essays and noticing

A number of empirical research has been conducted in various contexts to examine the effect of noticing-model essays on learners' writing in IELTS Task 2, especially the lexical dimension (Abe, 2009; Bagheri and Zare, 2009; Tieu & Baker, 2022). Results from these studies indicated that model essays and noticing had a positive impact on students' vocabulary use; nonetheless, there were several limitations that needed to be addressed.

In Japan, Abe (2009) set the stage for the noticing-model essay exploration by letting L2 learners write their own essays, read the native speakers' essays and compare the writings. Then, he interviewed seven participants to gain insights into which part of the model essay they noticed the most. Abe found that lexical items were the most attention-drawing aspect among the five categories (forms, vocabulary, content, discourse, and others). However, the author did not run any tests to prove the relationship between vocabulary gains and the amount of noticing. In other words, it still remains unknown whether there was a significant connection between the two variables.

In Iran, Bagheri and Zare (2009) investigated the topic much further by dividing 65 undergraduates into three groups: the control group (L2 intermediate level), the experimental group 1 (intermediate L2 level), and the experimental group 2 (L2 advanced level). While the control group learned writing in the ordinary way, the experimental groups were given model essays to read and asked to rewrite their own essays based on the input they received from the native speakers'. Results showed that the experimental groups outperformed the control group in all four aspects of IELTS Task 2 writing, and that the advanced students scored higher than the intermediate partners.

In the context of Vietnam, Tieu and Baker (2022) were the first to officially examine the relationship between model essays, noticing, and essay writing. They conducted a quasi-experimental study on 33 learners in a language center, who were put into two groups: the control group and the experimental group. Whereas the control group took lessons of writing in a normal way, those in the experimental group worked on native speakers' essays and revised their own essays after discussing with their teacher about their most favorable part in the model writing. Results from the pre-test as well as post-test revealed that the experimental group did better than the control group in terms of vocabulary, grammar, organization, and content. Nonetheless, due to Covid-19 pandemic, the post-test was delayed for three months after the treatment, which might have affected the research's outcomes as learners could have been exposed to other kinds of input and output.

In brief, the literature review highlighted two major gaps in previous studies. First, the role of model essays and noticing hypothesis in promoting learners' lexical resources in IELTS Task 2 writing was still unclear. Second, no studies ever examined the mediating effect of productive vocabulary knowledge in this process. Therefore, this present study intended to address these deficiencies by investigating the relationship between noticing, model essays, and academic writing in IELTS Task 2 and by exploring whether productive vocabulary knowledge might have any predicting role in students' second language acquisition, particularly lexical improvements.

Consequently, the research mainly seeks answers to the following questions:

1. Does noticing and model essays improve learners' lexical performance in IELTS Task 2 writing more than the conventional way?

2. Does productive vocabulary knowledge predict learners' lexical performance in IELTS Task 2 writing?

This research is vital for a number of reasons. First, by clarifying the uncertainties in earlier studies, it adds to the body of evidence on utilizing model essays and noticing as a feedback tool to enhance students' performance on IELTS Task 2 Writing. Second, it assists ESL/EFL instructors and students in deciding whether to use this strategy or not when preparing for the IELTS. Additionally, it sheds light on the role of productive vocabulary knowledge in this process, showing students and teachers to what extent they should expect lexical gains after the treatment.

3. Methodology

3.1 Participants

Thirty-two students (aged 18 to 22) at two language centers in Vietnam volunteered to participate in the study. Their English proficiency was supposed to be around B1 according to the CEFR (Common European Framework of Reference for Languages), as their IELTS scores ranged from 4.5 to 5.0. Sixteen of them (10 males, six females) from the same center were chosen as the control group, and the other 16 (9 males, seven females) from the other center were put into the experimental group. The participants took IELTS courses at the language centers and studied writing lessons twice a week, with one and a half hours for each session. Although they came from two different language centers, they studied IELTS with the same teacher because he was teaching IELTS at both places.

3.2 Research design and instruments

The present study adopted the quasi-experimental design, including a pre-test, a post-test, and a productive vocabulary test. While the control group (n=16) learned writing lessons in a conventional way with teacher feedback and peer feedback, the experimental group (n=16) used

model essays and noticing as a feedback instrument. All participants took the same tests, and the treatment was the only different thing between the two groups.

The productive vocabulary test

The Productive Levels Test (Laufer & Nation, 1999), which includes terms with frequency levels ranging from 2,000 to 10,000 as well as academic words from the University Word List, was used to estimate participants' knowledge of productive vocabulary. In this test, students had to use the given first letter to create the target word, which was placed in a certain situation. For example, participants were required to type "motive" to get the full score of the following sentence: The suspect had the opportunity and m_____ to commit the crime.

The pre-test

In the pre-test, all participants were given a topic (about educational issues) taken from the book IELTS Academic (2022) and asked to write an essay of at least 250 words in response to the question in 40 minutes. During the test, they were not allowed to use any dictionaries or have any discussions. After the allotted time, their essays were collected and given back to the researchers.

The treatment

As mentioned, the control group underwent normal teaching and writing corrections, whereas the experimental group was given native speakers' essays as models, which were taken from the book "High-scoring IELTS Writing – Model Answers" (Fang & Wang, 2012) and the ebook "The Key to IELTS Writing Task 2" by Cullen (2020). The chosen essays were about the same topic as the one in the pre-test. After reading the model essays, the participants in the experimental group were asked to highlight the parts they found useful or interesting. Then, they worked with a partner and exchanged what they

liked about the model essays before revising their own essays.

The post-test

In the post-test, all participants were required to write another essay on a similar topic to the one in the pre-test. The writing question was taken from the book "High-scoring IELTS Writing – Model Answers" (Fang & Wang, 2012). All the testing conditions (time limit, supervision, material use) were the same as those in the pre-test.

3.3 Data collection procedure

The whole procedure took place within two weeks. On the first day, all participants were asked to take the pre-test in 40 minutes and then take the productive vocabulary test in 10 minutes. Two days later, in the second writing lesson, students in the control group gave feedback on their partners' essays before receiving further feedback from their teacher. However, the experimental group read the model essays, noticed how the target language was used, and talked to one another about what they learned

from the model writing. During the pair discussion, any questions, if any, regarding the model essays were consulted with the teacher. In the following week, the third writing lesson, all participants were asked to rewrite their original essay (in the pre-test) based on either peer/teacher feedback (control group) or model essays (experimental group). In the fourth writing lesson, all participants took the post-test in 40 minutes, also under strict supervision. All the data were collected, graded, and stored for later analysis.

The pre-test and post-test essays were graded separately by the two researchers who were experienced in IELTS teaching for more than five years based on the IELTS Task 2 Writing Rubrics, Public Version (n.d.). As the main focus of the present study was on vocabulary, only scores of lexical resource (LR) criterion were discussed. The two raters then reviewed all the scores collectively and discussed every disagreement until they both came to an agreement on the final scores.

The data collection was summarized as follows:

Table 1. Data collection procedure

Day	Control Group	Experimental Group
1	Took the pre-test	
3	Gave feedback on another's essay. Received further feedback from teacher.	Read the model essays. Analyzed what could be learned. Shared ideas with a partner. Asked teacher for clarification, if any.
5	Rewrote the essay based on the feedback.	Rewrote the essay based on the model essays.
8	Took the post-test	

3.4 Data analysis

All data (test scores) were imported into Microsoft Excel prior to the analyses in SPSS (Statistical Packages for Social Sciences) version

27.0 and in R (R Core Team, 2022). Initially, several Shapiro-Wilk tests were run in order to check the distribution types of the test scores, which were demonstrated as follows:

Table 2. Normality tests

Group	Category	W	p
Control (n=16)	Lexical Resources (Pre-test)	0.788	<0.01
	Lexical Resources (Post-test)	0.748	<0.01
	Productive Vocabulary test	0.896	0.068
Experimental (n=16)	Lexical Resources (Pre-test)	0.796	<0.01
	Lexical Resources (Post-test)	0.827	<0.01
	Productive Vocabulary test	0.901	0.084

Table 2 revealed that the pre-test and post-test scores of participants in both groups were not normally distributed (all p values under .01). However, the productive vocabulary test in the control group ($W = .896$, $p = .068$) and in the experimental group ($W = .901$, $p = .084$) were of normal distribution.

As a result, Mann-Whitney U tests were used to compare the scores between the two groups, and Wilcoxon signed-rank tests were employed to compare the scores in the pre-test and post-test within each group. Additionally, as scores in the productive vocabulary test were normally distributed, the Linear Mixed Model (LMM) was adopted as this method could deal with data of non-normality and cover the individual

differences that other parametric tests like ANOVA or MANOVA might overlook (Noris, 2015). In other words, the LMM could yield more reliable outcomes (Linck & Cunnings, 2015). In this study, scores of lexical resources were treated as the dependent variable, while groups (control vs. experimental), productive vocabulary scores, and time (pre-test, post-test) were regarded as fixed effects. The participants were considered as random effects.

Consequently, the fitted model was (run via the lmerTest package in R)

$$\text{Scores} \sim \text{Group} + \text{Time} + \text{Productive Test} + (1|\text{Participant})$$

4. Results

Descriptive statistics

Table 3. Descriptive statistics for the three tests

Group	Test	Mean	SD	95% CI
Control (n=25)	Pre-test	4.69	0.704	4.31 – 5.06
	Post-test	4.94	0.574	4.63 – 5.24
	Productive vocabulary	31.5	10.328	26.0 – 37.0
Experimental (n=27)	Pre-test	4.75	0.683	4.39 – 5.11
	Post-test	5.44	0.814	5.0 – 5.87
	Productive vocabulary	31.94	9.740	26.75 – 37.13

Table 3 illustrates the descriptive data for the tests, and it could be inferred that students in both

groups achieved higher scores in lexical resources. Nonetheless, whether these

differences were significant or not required further analyses.

Research question 1: Does noticing and model essays improve learners' lexical performance in IELTS Task 2 writing more than the conventional way?

Table 4. Comparisons of the pre-test and post-test scores

	Mean Difference	Z/ U	p
Control group (CG) Post-test – Pre-test	0.25	Z = -1.155	0.248
The experimental group (EG) Post-test – Pre-test	0.69	Z = -3.317	0.001
Pre-test: CG – EG	-0.06	U = 121	0.772
Post-test: CG – EG	-0.5	U = 86	0.067

It could be seen from table 4 that there was no significant difference in the vocabulary score between the pre-test and post-test of the control group ($Z = -1.155, p = .248$), while those in the experimental group scored substantially higher in the post-test than in the pre-test ($Z = -3.317, p = .001$). However, results of the Mann-Whitney U tests indicated that there were no significant differences in the vocabulary scores between the two groups, either in the pre-test ($U = 121, p = .772$) or in the post-test ($U = 86, p = .067$).

Therefore, it could be concluded that although the noticing-model essays helped learners improve their LR score in IELTS Task 2 Writing up to .69 (more than 0.5 band score), the differences, when compared to the conventional way, were not significant.

Research question 2: Does productive vocabulary knowledge predict learners' lexical performance in IELTS Task 2 writing?

An independent-sample t-test of the productive vocabulary scores between the two groups showed that there was no significant difference ($t(30) = -.123, p = .903$).

Table 5. The LMM fixed effects

	β	SE	t	p
Intercept	4.494	0.391	11.505	<.001
Experimental Group	0.280	0.218	1.284	0.209
Time	0.469	0.127	29.0	<.001
Productive vocabulary	0.003	0.011	0.239	0.812

It was evident from table 5 that the productive vocabulary knowledge did not have any predicting role in learners' LR scores in IELTS

Task 2 Writing ($\beta = .003, SE = .01, t = .0239, p < .812$).

5. Discussion

After the quantitative analyses, there were two major findings that needed to be mentioned.

First, although noticing-model essays helped learners achieve a higher score in lexical resource criterion in IELTS Task 2 writing (up to .69 band score), this difference was not significant when compared to the results from the group trained conventionally. This finding was in contrast with previous studies (Bagheri and Zare; Tieu & Baker, 2022). There were two possible explanations. For one thing, the number of participants in the two groups was too small, with only 16 people, which might not be enough to generate a significant value. Another point was that in the research by Tieu and Baker (2022), the delay of the post-test could have affected the results as the participants might have been exposed to other types of input and output themselves. One more feasible element was that the amount and time of noticing participants in the experimental group (in the present study) might not be sufficient, which meant more practice was needed before a notable outcome could be achieved (Schmidt, 1985).

Second, it was found that productive vocabulary knowledge did not have any impact on participants' lexical resource scores in IELTS Task 2 writing. This was not in line with previous research on productive vocabulary knowledge and skills in the English language (Manalo & Greenwood, 2009; Uchihara & Saito, 2016). It could be due to the fact that participants' scores of productive vocabulary tests were quite similar between the two groups, limiting their effect on their essays. Another plausible explanation was that productive vocabulary knowledge could not be fully utilized under time pressure (only 40 minutes) in the pre-test and post-test.

Although not significantly better than the conventional way of writing corrections and feedback, the role of noticing-model essays should not be overlooked because it did improve learners' scores in lexical resources. Therefore, it

is recommended that teachers and schools implement this method in their lessons to not only reduce the heavy workload for teachers but also let students learn something new on their own. Moreover, this method can also be combined with the conventional one to make the lessons more interesting and less repetitive. Additionally, teachers should measure students' progress after a sufficient period (i.e., a month or two) to assure that learners are adequately exposed and draw attention to comprehensible input in model essays. As for productive vocabulary knowledge, if the test results among students are not significantly different, teachers can choose not to invest time in discovering this aspect any further.

6. Conclusion

The present study aimed to discover whether the combination of noticing hypothesis and model essays had any significant role in improving learners' performances in lexical resources in IELTS Task 2 writing and also to examine the predicting part of productive vocabulary knowledge in this process. Via a quasi-experimental design with the participation of 32 students, it was found that despite not having a significant role in boosting learners' vocabulary scores in academic essay writing when compared to the conventional method, the noticing-model essays still helped learners raise their LR to score up to .69, in just within two weeks. Another finding was that productive vocabulary knowledge did not substantially affect learners' LR scores in IELTS Task 2 writing. Therefore, teachers are recommended to implement this new technique, either with or without the combination of the teacher or peer feedback, in their lessons. Another pedagogical implication was that teachers should not rush to test their learners' performances, yet should wait for learners to be sufficiently exposed to model essays and have adequate practice prior to any assessments under time pressure. However, two limitations should be mentioned in this study. First, the number of

participants was too small, which might not entirely reflect the circumstances. Second, no qualitative data was employed to gain insights into learners' thoughts and attitudes toward the use of noticing-model essays. Consequently, future research is recommended to employ a mixed-methods approach with a large sample size to yield results with higher reliability.

References

1. ABE, M. (2009). Noticing in comparing own essay with model essay: An exploratory study of Japanese L2 writers. *The Bulletin of the Kanto-Koshin-Etsu English Language Education Society*, 23, 71–82.
2. Bagheri, M. S., & Zare, M. (2009). THE ROLE OF USING IELTS MODEL ESSAYS IN IMPROVING LEARNERS' WRITING AND THEIR AWARENESS OF WRITING FEATURES.
3. Braine, G. (2003). From a teacher-centered to a student-centered approach: A study of peer feedback in Hong Kong writing classes. *Journal of Asian Pacific Communication*, 13(2), 269–288.
4. Chen, C. W. (2010). Graduate students' self-reported perspectives regarding peer feedback and feedback from writing consultants. *Asia Pacific Education Review*, 11(2), 151–158.
5. Cullen, P. (n.d.). The Key to IELTS Writing Task 2. Cullen Education - Key to IELTS. Retrieved September 1, 2022, from <https://keytoielts.com/the-key-to-ielts-success-2/>
6. Fang, T., & Wang, D. F. (2012). High-scoring IELTS Writing Model Answers (Based on past papers). *Tổng hợp Tp Hồ Chí Minh; Nhân trí Việt*.
7. Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11(4), 459–479.
8. IELTS Academic 17. (2022). Cambridge University Press.
9. Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33–51.
10. Lee, I. (2003). L2 writing teachers' perspectives, practices and problems regarding error feedback. *Assessing Writing*, 8(3), 216–237.
11. Leow, R. P. (2018). Noticing hypothesis. *The TESOL Encyclopedia of English Language Teaching*, 1–6.
12. Lin, Y.-C. A. (2009). An examination of teacher feedback, face-to-face peer feedback, and Google Documents peer feedback in Taiwanese EFL college students' writing. Alliant International University, San Diego.
13. Linck, J. A., & Cummings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, 65(S1), 185–207.
14. Norris, J. M. (2015). Statistical significance testing in second language research: Basic problems and suggestions for reform. *Language Learning*, 65(S1), 97–126.
15. Oya, T., Manalo, E., & Greenwood, J. (2009). The influence of language contact and vocabulary knowledge on the speaking performance of Japanese students of English. *The Open Applied Linguistics Journal*, 2(1).
16. Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, 10(4), 277–303.
17. R Core Team. (n.d.). R: A language and environment for statistical computin. R

- Foundation for Statistical Computing, Vienna, Austria. Retrieved August 20, 2022, from <https://www.R-project.org>
18. Read, J. (2019a). Key issues in measuring vocabulary knowledge. In *The Routledge handbook of vocabulary studies* (pp. 545–560). Routledge.
 19. Read, J. (2019b). Key issues in measuring vocabulary knowledge. In *The Routledge handbook of vocabulary studies* (pp. 545–560). Routledge.
 20. Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, 59(1), 23–30.
 21. Ruru, T. A. N., & Sulisty, T. (2020). Peer review in writing activities: Outcomes and perceptions of EFL students. *Journal of Research on English and Language Learning (J-REaLL)*, 1(2), 127–132.
 22. Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. *Talking to Learn: Conversation in Second Language Acquisition*, 237, 326.
 23. Schmidh, R. (1983). Interaction, acculturation and the acquisition of communicative competence. *Sociolinguistics and Language Acquisition*, 137–174.
 24. Séror, J. (2011). Alternative Sources of Feedback and Second Language Writing Development in University Content Courses. *Canadian Journal of Applied Linguistics/Revue Canadienne de Linguistique Appliquée*, 14(1), 118–143.
 25. Susanto, S., & Hidayati, M. (2020). The Effect of Peer Review on Students' Motivation and Writing Ability. *IJOTL-TL: Indonesian Journal of Language Teaching and Linguistics*, 5(1), 1–14.
 26. Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in Second Language Acquisition*, 15, 165–179.
 27. Swain, M. (1995). Three functions of output in second language learning. *Principles and Practice in Applied Linguistics: Studies in Honor of HG Widdowson*, 125–144.
 28. Swain, M. (1997). The output hypothesis, focus on form and second language learning. *Applying Linguistics: Insights into Language in Education*, 1–21.
 29. Swain, M. (2005). The output hypothesis: Theory and research. In *Handbook of research in second language teaching and learning* (pp. 495–508). Routledge.
 30. Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371–391.
 31. Tieu, L., & Baker, J. R. (2022). Using model essays in conjunction with noticing as a feedback instrument in IELTS writing preparation. *Innovation in Language Learning and Teaching*, 1–13.
 32. Uymaz, E. (2019). The effects of peer feedback on the essay writing performances of EFL students. *International Journal of Curriculum and Instruction*, 11(2), 20–37.
 33. WRITING TASK 2: Band Descriptors (public version). (n.d.). IELTS. <https://www.ielts.org/-/media/pdfs/writing-band-descriptors-task-2.ashx>
 34. Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179–200.
 35. Zhao, H. (2010). Investigating learners' use and understanding of peer and

teacher feedback on writing: A comparative study in a Chinese English writing classroom. Assessing Writing,

15(1), 3–17.

Appendix



WRITING TASK 2: Band Descriptors (public version)

Band	Task response	Coherence and cohesion	Lexical resource	Grammatical range and accuracy
9	<ul style="list-style-type: none"> fully addresses all parts of the task presents a fully developed position in answer to the question with relevant, fully extended and well supported ideas 	<ul style="list-style-type: none"> uses cohesion in such a way that it attracts no attention skillfully manages paragraphing 	<ul style="list-style-type: none"> uses a wide range of vocabulary with very natural and sophisticated control of lexical features; rare minor errors occur only as 'slips' 	<ul style="list-style-type: none"> uses a wide range of structures with full flexibility and accuracy; rare minor errors occur only as 'slips'
8	<ul style="list-style-type: none"> sufficiently addresses all parts of the task presents a well-developed response to the question with relevant, extended and supported ideas 	<ul style="list-style-type: none"> sequences information and ideas logically manages all aspects of cohesion well uses paragraphing sufficiently and appropriately 	<ul style="list-style-type: none"> uses a wide range of vocabulary fluently and flexibly to convey precise meanings skillfully uses uncommon lexical items but there may be occasional inaccuracies in word choice and collocation produces rare errors in spelling and/or word formation 	<ul style="list-style-type: none"> uses a wide range of structures the majority of sentences are error-free makes only very occasional errors or inappropriacies
7	<ul style="list-style-type: none"> addresses all parts of the task presents a clear position throughout the response presents, extends and supports main ideas, but there may be a tendency to over-generalise and/or supporting ideas may lack focus 	<ul style="list-style-type: none"> logically organises information and ideas; there is clear progression throughout uses a range of cohesive devices appropriately although there may be some under-/over-use presents a clear central topic within each paragraph 	<ul style="list-style-type: none"> uses a sufficient range of vocabulary to allow some flexibility and precision uses less common lexical items with some awareness of style and collocation may produce occasional errors in word choice, spelling and/or word formation 	<ul style="list-style-type: none"> uses a variety of complex structures produces frequent error-free sentences has good control of grammar and punctuation but may make a few errors
6	<ul style="list-style-type: none"> addresses all parts of the task although some parts may be more fully covered than others presents a relevant position although the conclusions may become unclear or repetitive presents relevant main ideas but some may be inadequately developed/unclear 	<ul style="list-style-type: none"> arranges information and ideas coherently and there is a clear overall progression uses cohesive devices effectively, but cohesion within and/or between sentences may be faulty or mechanical may not always use referencing clearly or appropriately uses paragraphing, but not always logically 	<ul style="list-style-type: none"> uses an adequate range of vocabulary for the task attempts to use less common vocabulary but with some inaccuracy makes some errors in spelling and/or word formation, but they do not impede communication 	<ul style="list-style-type: none"> uses a mix of simple and complex sentence forms makes some errors in grammar and punctuation but they rarely reduce communication
5	<ul style="list-style-type: none"> addresses the task only partially; the format may be inappropriate in places expresses a position but the development is not always clear and there may be no conclusions drawn presents some main ideas but these are limited and not sufficiently developed; there may be irrelevant detail 	<ul style="list-style-type: none"> presents information with some organisation but there may be a lack of overall progression makes inadequate, inaccurate or over-use of cohesive devices may be repetitive because of lack of referencing and substitution may not write in paragraphs, or paragraphing may be inadequate 	<ul style="list-style-type: none"> uses a limited range of vocabulary, but this is minimally adequate for the task may make noticeable errors in spelling and/or word formation that may cause some difficulty for the reader 	<ul style="list-style-type: none"> uses only a limited range of structures attempts complex sentences but these tend to be less accurate than simple sentences may make frequent grammatical errors and punctuation may be faulty; errors can cause some difficulty for the reader
4	<ul style="list-style-type: none"> responds to the task only in a minimal way or the answer is tangential; the format may be inappropriate presents a position but this is unclear presents some main ideas but these are difficult to identify and may be repetitive, irrelevant or not well supported 	<ul style="list-style-type: none"> presents information and ideas but these are not arranged coherently and there is no clear progression in the response uses some basic cohesive devices but these may be inaccurate or repetitive may not write in paragraphs or their use may be confusing 	<ul style="list-style-type: none"> uses only basic vocabulary which may be used repetitively or which may be inappropriate for the task has limited control of word formation and/or spelling; errors may cause strain for the reader 	<ul style="list-style-type: none"> uses only a very limited range of structures with only rare use of subordinate clauses some structures are accurate but errors predominate, and punctuation is often faulty
3	<ul style="list-style-type: none"> does not adequately address any part of the task does not express a clear position presents few ideas, which are largely undeveloped or irrelevant 	<ul style="list-style-type: none"> does not organise ideas logically may use a very limited range of cohesive devices, and those used may not indicate a logical relationship between ideas 	<ul style="list-style-type: none"> uses only a very limited range of words and expressions with very limited control of word formation and/or spelling errors may severely distort the message 	<ul style="list-style-type: none"> attempts sentence forms but errors in grammar and punctuation predominate and distort the meaning
2	<ul style="list-style-type: none"> barely responds to the task does not express a position may attempt to present one or two ideas but there is no development 	<ul style="list-style-type: none"> has very little control of organisational features 	<ul style="list-style-type: none"> uses an extremely limited range of vocabulary; essentially no control of word formation and/or spelling 	<ul style="list-style-type: none"> cannot use sentence forms except in memorised phrases
1	<ul style="list-style-type: none"> answer is completely unrelated to the task 	<ul style="list-style-type: none"> fails to communicate any message 	<ul style="list-style-type: none"> can only use a few isolated words 	<ul style="list-style-type: none"> cannot use sentence forms at all
0	<ul style="list-style-type: none"> does not attend does not attempt the task in any way writes a totally memorised response 			