

Bridging the gap: Analysis of Olympiad Writing for sustainable ESP curriculum design

Anita Dwi Hapsari¹, Irmayani², and Syifa' Khuriyatuz Zahro³

^{1,2,3}English Language Education, Universitas Islam Darul 'Ulum, Lamongan, Indonesia

Abstract. The mismatch between classroom-based English instruction and students' real-world communicative demands highlights the need for empirically grounded curriculum design. This study aims to provide empirical insights for English for Specific Purposes (ESP) curriculum development by examining authentic argumentative writing produced by high school students in an English Olympiad. Employing a descriptive quantitative approach, the study analyzes 20 writing scripts using an analytic scoring rubric adapted from CEFR descriptors, focusing on task achievement, coherence, lexical resources, and grammatical accuracy. The findings reveal that although students demonstrate strong ideational creativity and critical thinking, their written performance is constrained by limited lexical resources and frequent grammatical inaccuracies. These results indicate a significant gap between students' conceptual understanding and their linguistic realization in academic writing. The analysis of argumentative writing performance therefore provides valuable insights into students' specific linguistic deficiencies and latent learning needs. The study proposes that such empirical evidence can inform the design of a responsive and sustainable ESP curriculum that supports the development of critical thinking, persuasive communication, and academic literacy required for higher education and future professional contexts.

Keywords: Needs Analysis; ESP Curriculum; Argumentative Writing; Sustainable English Education

1 Introduction

The persistent gap between English language instruction delivered in classroom settings and the communicative competencies required in real-world academic and professional contexts remains a central challenge in English language teaching (Hyland, 2016). In many educational contexts, particularly in secondary education, the focus on textbook-driven, generalized English instruction often fails to equip students with the specific linguistic skills necessary for higher education, competitive environments such as academic Olympiads, and future career demands (Anthony, 2018). This disconnect underscores an urgent need for curriculum frameworks that are not only responsive to learners' immediate academic goals but also sustainable in fostering long-term communicative competence. In this study, a

¹ Corresponding author: anitadwi@unisda.ac.id

sustainable ESP curriculum is defined as an instructional framework that not only addresses learners' immediate academic needs but also equips them with transferable language competencies, autonomous learning strategies, and adaptive literacy skills that remain relevant across evolving disciplinary and professional contexts. Unlike a short-term test-preparation model, a sustainable ESP curriculum emphasizes long-term learner development, reusability of skills, and the ability to continue learning independently beyond the course. English for Specific Purposes (ESP) has emerged as a pedagogical approach that directly addresses this need by prioritizing learner-centered, needs-based curriculum design (Hutchinson & Waters, 1987). However, the effectiveness of an ESP curriculum is contingent upon the rigor and empirical basis of the needs analysis that informs it (Nation & Macalister, 2010).

A robust needs analysis serves as the cornerstone of any ESP curriculum, providing critical insights into the target situation, learners' existing competencies, and the specific linguistic deficiencies that must be addressed (Brown, 2016). Traditionally, needs analysis has relied on surveys, interviews, and stakeholder perceptions. While valuable, these methods can sometimes overlook the nuanced, performance-based evidence of learners' actual capabilities (Serafini et al., 2015). To design a truly responsive and sustainable ESP curriculum, it is essential to complement subjective data with objective analyses of authentic learner output. One rich source of such evidence is the academic writing produced by students in high-stakes, competitive settings like English Olympiads. These contexts demand higher-order cognitive skills, such as critical thinking and persuasive argumentation, alongside advanced linguistic accuracy (Cumming, 2013). Analyzing performance in such settings can therefore reveal not only what students can do but also the specific constraints that prevent them from fully realizing their conceptual potential.

This study is predicated on the belief that an in-depth analysis of authentic argumentative writing can provide the empirical foundation needed to bridge the gap between generalized instruction and specific learner needs. By focusing on writing scripts from an English Olympiad, this research aims to move beyond a surface-level assessment of language proficiency to a granular understanding of students' strengths and weaknesses in academic writing. The argumentative genre is particularly revealing, as it requires the integration of logical reasoning, structural coherence, precise lexical choices, and accurate grammatical structures—skills that are indispensable for success in university and professional environments (Grabe & Kaplan, 2014; Wingate, 2012).

While previous ESP needs analysis studies have predominantly relied on surveys, interviews, and stakeholder perceptions (Brown, 2016), these methods may not fully capture learners' actual linguistic performance. Recent scholarship calls for the integration of direct learner output as a complementary data source to enhance the validity and specificity of needs identification (Serafini et al., 2015; Long, 2005). However, empirical studies that systematically analyze authentic learner writing to inform ESP curriculum design remain scarce, particularly in secondary school contexts and competitive academic settings such as English Olympiads. Therefore, this study aims to conduct a comprehensive needs analysis for ESP curriculum development by examining the argumentative writing of high school students participating in an English Olympiad. Employing a descriptive quantitative approach, the study analyzes 20 writing scripts using an analytic scoring rubric adapted from the Common European Framework of Reference for Languages (CEFR) descriptors (Council of Europe, 2001). The analysis focuses on four key dimensions: task achievement, coherence and cohesion, lexical resources, and grammatical accuracy. By identifying the specific linguistic deficiencies and latent learning needs evident in students' written performance, this research seeks to provide actionable insights for designing an ESP curriculum that is not only responsive to current academic demands but also sustainable in supporting students' long-term development of critical thinking, persuasive communication, and academic literacy.

This paper will first review the relevant literature on ESP curriculum design and needs analysis, then detail the methodology, present the findings from the analysis of the Olympiad writing scripts, and finally discuss the implications for developing a sustainable ESP curriculum.

2 Method

This study employed a descriptive quantitative approach to analyze argumentative writing scripts produced by high school students in an English Olympiad. The focus was on measuring specific dimensions of writing performance to identify linguistic deficiencies and inform ESP curriculum design.

2.1 Design

A descriptive quantitative research design was utilized, as the study aimed to systematically describe the characteristics of students' argumentative writing performance across four analytical categories: task achievement, coherence and cohesion, lexical resources, and grammatical accuracy. This design allowed for the objective quantification of writing quality using a structured scoring rubric, enabling the identification of patterns and areas of weakness within the sample (Creswell & Creswell, 2018). No experimental treatment was applied; the research focused solely on analyzing authentic learner output in a naturalistic competitive context.

2.2 Participant/ Population and Sample

The participants were 20 high school students from various schools across East Java, Indonesia, who advanced to the semifinal round of a regional English Olympiad. From a larger pool of applicants, all participants initially took a multiple-choice test; only the 20 students with the highest scores qualified for the semifinal, during which they produced the argumentative writing scripts analyzed in this study. The participants represented different grade levels (X, XI, and XII) and had received formal English instruction at the secondary level for periods ranging from one to three years. Their selection to the semifinal round indicates a level of English proficiency typically ranging from intermediate to upper-intermediate, as commonly observed in such competitive academic events. Their ages ranged from 16 to 18 years.

The sample size of 20 participants is considered adequate for the purposes of this descriptive quantitative study, as it aligns with commonly accepted practices in in-depth writing performance analysis using analytic rubrics (Weigle, 2002). Additionally, the sample represents the entire population of semifinalists available from the Olympiad, making it a total sampling rather than a sampled subset. This allows for a complete rather than partial representation of high-achieving writers in this competitive context.

The classification of participants' English proficiency as intermediate to upper-intermediate, was based on their performance on the initial multiple-choice test, which was designed to align with CEFR B1 and B2 descriptors. Students who scored above the cut-off point (top 20 scores) were assumed to possess at least B1 (intermediate) proficiency, with the highest scorers approaching B2 (upper-intermediate). No external proficiency test (e.g., TOEFL, IELTS) was administered, so this categorization is indicative rather than definitive.

A key limitation of this study is that the sample consists exclusively of semifinalists, who represent higher-achieving English learners. The findings may therefore not be generalizable to the broader population of high school students, particularly those with lower proficiency levels who did not qualify for the semifinal round. Future research should include a more diverse range of proficiency levels to capture a fuller spectrum of ESP needs.

2.3 Instrument/ Data collection

The primary instrument for data collection was an analytic scoring rubric adapted from the descriptors of the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001). The rubric assessed student writing across four domains:

1. Task achievement: the extent to which the writing addressed the prompt, presented a clear stance, and fulfilled the communicative purpose.
2. Coherence and cohesion: the logical organization of ideas, use of paragraphing, and effective use of cohesive devices.
3. Lexical resources: the range and appropriateness of vocabulary, including precision, variety, and control of word choice.
4. Grammatical accuracy: the syntactic complexity, error frequency, and overall control of grammatical structures.

Each domain was scored on a scale of 0–5, with detailed descriptors for each score point. Table 1 below presents the analytic scoring rubric used in this study, adapted from CEFR descriptors (Council of Europe, 2001) for the four domains. Each domain was scored on a 0–5 scale, where 0 = performance below the minimum standard, 1 = limited, 2 = moderate but weak, 3 = adequate, 4 = good, and 5 = excellent. The detailed descriptors for each score level are as follows:

Table 1. Analytic Scoring Rubric Adapted from CEFR Descriptors

Score	Task Achievement	Coherence & Cohesion	Lexical Resources	Grammatical Accuracy
5	Fully addresses prompt, clear stance, well-supported arguments	Logical organization, smooth transitions, effective paragraphing	Wide range of precise, natural vocabulary; academic style	Consistently accurate grammar; complex sentences used appropriately
4	Addresses prompt clearly, stance present, mostly supported	Generally logical, some cohesive devices used appropriately	Good range of vocabulary; occasional imprecise choices	Mostly accurate; few minor errors (e.g., article, preposition)
3	Addresses prompt but stance could be clearer; basic support	Adequate organization; some awkward or missing transitions	Adequate vocabulary for basic communication; some repetition	Adequate control; errors do not impede understanding
2	Partially addresses prompt; stance unclear or weak support	Poor organization; limited or inappropriate use of cohesive devices	Limited vocabulary; frequent repetition or word choice errors	Frequent errors that sometimes impede understanding
1	Minimally addresses prompt; no clear stance or support	No logical order; difficult to follow	Very limited vocabulary; mainly basic words	Pervasive errors that severely impede understanding
0	No response or completely off-topic	No coherence	Essentially no control	No evidence of grammatical control

This rubric was used by both raters to score each of the 20 writing scripts. The descriptors were derived from CEFR B1–B2 level performance indicators, which correspond to the proficiency range of the participants.

The rubric was pilot-tested on five writing samples not included in the final analysis to ensure inter-rater reliability, resulting in a Cohen's kappa coefficient of 0.85 across two raters, indicating strong agreement.

2.4 Data Analysis

The 20 argumentative writing scripts were collected from the Olympiad organizers with appropriate permissions. Each script was anonymized and assigned a code. Two trained raters independently scored each script using the analytic rubric. The scores were then tabulated and analyzed using descriptive statistics, including means, standard deviations, and frequency distributions for each of the four scoring domains (Pallant, 2020). The mean scores across domains were compared to identify relative strengths and weaknesses.

Detailed performance data. To support the interpretation of findings, the analysis presents not only overall mean scores but also the frequency distribution of scores across each domain (e.g., number of scripts receiving scores of 0–5 per category). This allows for a more nuanced understanding of how many students performed at each proficiency level.

Representative writing excerpts. In addition to quantitative scores, qualitative annotations were made during scoring to capture recurring patterns in lexical and grammatical errors. Representative excerpts from student scripts illustrating specific deficiencies (e.g., word choice errors, subject-verb agreement, tense inconsistency) are included in the results section to provide concrete evidence of the identified weaknesses. These excerpts are presented anonymously and have been lightly edited only for readability without changing the original grammatical or lexical errors.

Additionally, qualitative annotations were made during scoring to capture recurring patterns in lexical and grammatical errors, which were later categorized to provide deeper insights into the specific linguistic deficiencies. The results were then interpreted in light of the needs analysis framework for ESP curriculum design, aligning with the study's objective to generate empirically grounded recommendations.

3 Results and Discussion

3.1 Results

The analysis of 20 argumentative writing scripts using the CEFR-aligned analytic rubric yielded descriptive statistics for the four assessed domains: task achievement, coherence and cohesion, lexical resources, and grammatical accuracy. Table 2 presents the mean scores, standard deviations, and score ranges for each domain.

Table 2. Descriptive Statistics of Argumentative Writing Performance (N=20)

Domain	Mean	Standard Deviation	Range (Min–Max)
Task Achievement	3.75	0.72	2.5–4.5
Coherence and Cohesion	3.20	0.80	2.0–4.0
Lexical Resources	2.45	0.76	1.5–3.5
Grammatical Accuracy	2.30	0.81	1.0–3.5

The highest mean score was observed in task achievement ($M = 3.75$, $SD = 0.72$), indicating that students generally succeeded in addressing the essay prompt, presenting a clear position, and fulfilling the basic communicative purpose of argumentative writing. In contrast, lexical resources ($M = 2.45$, $SD = 0.76$) and grammatical accuracy ($M = 2.30$, $SD = 0.81$) received the lowest mean scores, suggesting that students faced considerable difficulties in vocabulary range, precision, and grammatical control. Coherence and cohesion ($M = 3.20$, $SD = 0.80$) fell in the mid-range, revealing moderate ability to organize ideas and use cohesive devices, though with notable inconsistency.

To illustrate the quantitative findings, representative excerpts from student writing scripts are presented below. These excerpts are reproduced exactly as written, including original

errors, to provide concrete evidence of the recurring deficiencies identified in the analysis. Examples of limited lexical resources include:

- [1] *"The government should make a rule to force people to care environment."* (inappropriate word choice: "force" instead of "encourage"; missing preposition "for")
- [2] *"Social media is good. Social media helps us. But social media also bad."* (repetition of basic vocabulary; lack of cohesive variety)
- [3] *"Students who want to be success must study hard."* (incorrect word form: "success" instead of "successful")
- [4] Subject-verb agreement: *"The arguments from the writer is very strong."* (should be "are")
- [5] Tense consistency: *"Students who join Olympiad usually prepare hard. Last year, they practice every day."* (should be "practiced")
- [6] Article usage: *"She is a best student in the class."* (should be "the best student")
- [7] Sentence fragment: *"Because the government does not care. Many people suffer."* (run-on/fragment issue)

These examples are representative of patterns observed across multiple scripts. The full set of annotated errors was used to inform the interpretation of the quantitative scores.

Qualitative annotations further revealed specific patterns. In lexical resources, frequent issues included reliance on basic vocabulary, repetition of words, inappropriate word choices, and limited use of academic or topic-specific terms. In grammatical accuracy, recurring errors involved subject-verb agreement, verb tense consistency, article usage, and sentence fragments. Despite these linguistic constraints, many scripts demonstrated well-developed ideas, logical reasoning, and creative arguments, which aligned with the relatively higher task achievement scores.

3.2 Discussion

The findings reveal a notable gap between students' conceptual abilities and their linguistic realization in argumentative writing. While students exhibited strong ideational creativity and critical thinking—evident in their ability to articulate clear positions and develop logical arguments—their performance was significantly constrained by limited lexical resources and frequent grammatical inaccuracies. This discrepancy aligns with previous research indicating that learners often possess higher-order cognitive skills but lack the linguistic repertoire to express them effectively in academic writing (Cumming, 2013; Wingate, 2012).

The finding that students demonstrate strong critical thinking but weak linguistic realization is consistent with the work of Cumming (2013), who reported that L2 writers in assessment contexts often produce conceptually rich but linguistically flawed texts. However, unlike Cumming's study which focused on university-level integrated writing tasks, the present study finds that this gap persists even among competitively selected secondary school writers, suggesting that the issue may begin earlier in the educational pathway.

Compared to traditional ESP needs analysis studies that rely on self-reported data (e.g., Brown, 2016), the present study's use of authentic writing samples revealed a different profile of needs. For example, while surveys might indicate that students perceive vocabulary as a challenge, the specific patterns of lexical errors identified here (e.g., inappropriate word choice, repetition, word form errors) provide a more precise target for instruction. This aligns with Serafini et al. (2015), who argued that performance-based needs analysis yields more specific and actionable data than perception-based methods alone.

Furthermore, compared to previous analyses of argumentative writing in EFL contexts (e.g., Wingate, 2012), which emphasized structural coherence as the primary challenge, the

present study found that task achievement and coherence were relatively stronger, whereas lexical and grammatical accuracy were weaker. This difference may reflect the specific preparation and selection mechanisms of the Olympiad context, which likely emphasizes argumentation structure over language precision. Such contextual variation underscores the importance of conducting locally grounded needs analysis rather than applying generic findings across settings.

The relatively higher scores in task achievement suggest that the students understood the rhetorical demands of argumentative writing, such as taking a stance and supporting it with reasons. This reflects the emphasis on critical thinking and argumentation in the Olympiad preparation process. However, the low scores in lexical resources and grammatical accuracy point to a foundational deficiency in language proficiency that impedes effective communication. These results echo the findings of Grabe and Kaplan (2014), who emphasize that advanced literacy requires not only cognitive skills but also automaticity in vocabulary and grammatical structures.

From the perspective of needs analysis for ESP curriculum design, these results provide empirical evidence of specific learning needs. The linguistic deficiencies identified—particularly in vocabulary and grammar—represent the *lacks* that must be addressed to bridge the gap between students' current performance and the target competencies required for higher education and professional contexts (Hutchinson & Waters, 1987). Moreover, the strengths in task achievement and critical thinking suggest that a sustainable ESP curriculum can build upon these existing assets while systematically scaffolding lexical and grammatical development.

The study underscores the value of using authentic performance data, such as Olympiad writing scripts, as a source for needs analysis. Traditional methods like surveys may capture perceived needs, but they often fail to reveal the nuanced linguistic gaps that become evident only through direct analysis of learner output (Serafini et al., 2015). By integrating such empirical evidence, curriculum designers can develop ESP programs that are both responsive and sustainable—responsive because they target actual deficiencies, and sustainable because they leverage learners' existing cognitive strengths to foster deeper engagement with academic literacy.

Furthermore, the findings highlight the importance of integrating language instruction with content and critical thinking within ESP frameworks. A sustainable ESP curriculum should not treat vocabulary and grammar as isolated components but embed them within authentic academic writing tasks that require argumentation, persuasion, and critical analysis (Hyland, 2016). Such an approach would align with the principles of sustainable education, which emphasize the development of transferable skills that prepare learners for lifelong learning and adaptability in diverse communicative contexts.

4 Conclusion

This study analyzed argumentative writing scripts produced by high school students in an English Olympiad to identify specific linguistic deficiencies that can inform ESP curriculum design. The findings revealed that while students possess strong critical thinking and task achievement skills, their writing is constrained by significant deficiencies in lexical resources and grammatical accuracy. These results demonstrate a critical gap between conceptual understanding and linguistic realization, which must be addressed to prepare students for the academic and professional demands of higher education.

Within the scope of this study, which is limited to 20 semifinalist writing scripts from a regional Olympiad, the findings provide empirical evidence of learners' specific weaknesses in vocabulary and grammar. No claim is made that these findings are generalizable to all secondary school learners, nor that the study has developed a full ESP curriculum framework.

The study contributes to the field by demonstrating the utility of authentic writing performance as an empirical basis for needs analysis. By moving beyond subjective perceptions and analyzing actual learner output, educators can identify precise areas of need and design targeted, evidence-based interventions.

Based on the findings, several preliminary implications for a sustainable ESP curriculum can be offered, rather than a fully developed framework. As defined in the introduction, a sustainable ESP curriculum emphasizes long-term learner development, transferable skills, and autonomous learning. Accordingly, the following components warrant consideration in future curriculum design:

1. Integration of academic vocabulary development within argumentative writing tasks, focusing on high-frequency academic words and contextualized usage.
2. Grammar instruction embedded in meaningful writing contexts (e.g., teaching subject-verb agreement through analysis of student-generated claims), rather than as isolated drills.
3. Continued emphasis on critical thinking and task achievement, leveraging students' existing strengths while scaffolding language development.
4. Inclusion of self-access learning materials and peer feedback cycles to foster autonomous learning, aligning with the sustainability principle of long-term skill transferability.

However, the concept of sustainability in ESP curriculum remains underdeveloped in the existing literature and in this study. The present research only begins to operationalize sustainability by identifying learner needs that, if addressed, could lead to more durable competencies. Future research should further conceptualize and empirically test what constitutes a sustainable ESP curriculum, including longitudinal studies that track whether learners retain and transfer the targeted skills over time.

Given the study's limitations (small sample, exclusive use of semifinalists, and focus on a single writing genre), future research should expand the sample size, include multiple genres of writing, and incorporate a wider range of proficiency levels to provide a more comprehensive profile of learner needs. Longitudinal studies could also examine how ESP curricula informed by such needs analyses impact students' writing development over time.

Ultimately, bridging the gap between classroom instruction and real-world demands requires a commitment to empirically grounded, learner-centered curriculum design that empowers students to realize their full communicative potential.

References

- Anthony, L. (2018). *Introducing English for specific purposes*. Routledge.
- Brown, J. D. (2016). *Introducing needs analysis and English for specific purposes*. Routledge.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.)*. SAGE Publications.
- Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly*, 1–8. <https://doi.org/10.1080/15434303.2011.622016>
- Grabe, W., & Kaplan, R. B. (2014). *Theory and practice of writing: An applied linguistic perspective*. Routledge.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.

- Hyland, K. (2016). *Teaching and researching writing* (3rd ed.). Routledge.
- Long, M. H. (2005). *Second language needs analysis*. Cambridge University Press.
- Nation, I. S. P., & Macalister, J. (2010). *Language curriculum design*. Routledge.
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). Routledge.
- Serafini, E. J., Lake, J., & Long, M. H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for Specific Purposes*, 40, 11–26. <https://doi.org/10.1016/j.esp.2015.05.002>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
- Wingate, U. (2012). ‘Argument!’ Helping students understand what essay writing is about. *Journal of English for Academic Purposes*, 11(2), 145–154. <https://doi.org/10.1016/j.jeap.2011.11.001>